Modelling epidemic dynamics in exchange networks with constrained interactions and adaptive behaviour.

Mathieu Moslonka Lefebvre

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Ecole Doctorale ABIES
Agriculture
Agriculture
Alimentation
Food
Biologie
Biology
Environnement
Environment
Santé
Health

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : Mathématiques appliquées au vivant et à l'environnement

présentée et soutenue publiquement par

Mathieu MOSLONKA-LEFEBVRE

le jeudi 4 décembre 2014

Modéliser les dynamiques épidémiques dans des réseaux d'échanges avec interactions contraintes et comportements adaptatifs

Directeur de thèse : **M. Hervé MONOD**
Co-directeur de thèse : **M. Christopher A. GILLIGAN**
Co-encadrement de la thèse : **M. João A. N. FILIPE** et **Mme Elisabeta VERGU**

Jury

M. Bernard CAZELLES	Professeur des universités	EEM, Université Pierre et Marie Curie	Rapporteur
Mme Louise MATTHEWS	Senior research fellow	IBAHCM, University of Glasgow (UK)	Rapporteur
M. Olivier ALLAIS	Chargé de recherche	ALISS, INRA	Examinateur
M. Minus Van BAALEN	Chargé de recherche	EEM, CNRS	Examinateur
M. Michel De LARA	Professeur	CERMICS, Ponts ParisTech	Président
M. Hervé MONOD	Directeur de recherche	MaIAGE, INRA	Directeur de thèse
M. Alexandre PÉRY	ICPEF, HDR	ABIES, AgroParisTech	Examinateur
Mme Elisabeta VERGU	Chargée de recherche	MaIAGE, INRA	Co-encadrante de thèse

MaIAGE, INRA, France
Epidemiology and Modelling Group, University of Cambridge, UK

Au cosmos. Tout particulièrement,

à Djudju, étoile radiante,

à Anna, étoile résonante,

à Anton, notre étoile naissante.

Remerciements

Je tiens à exprimer ma reconnaissance au Ministère chargé de l'Agriculture, à l'Institut National de Recherche Agronomique, à l'AgroParisTech et à l'Académie d'Agriculture de France ainsi qu'à l'Université de Cambridge pour avoir soutenu mon projet de thèse. Mes remerciements vont également à l'Institut Henri Poincaré pour m'avoir permis de travailler dans d'excellentes conditions à proximité de mon domicile.

Chercher comporte des risques. Et pratiquer l'interdisciplinarité nécessite de faire preuve d'une certaine indiscipline. Je salue le dispositif de formation doctorale du corps des Ingénieurs des Ponts, Eaux et Forêts pour la réelle capacité de prise de risque qu'il m'a apporté, et cela en tout début de carrière. J'ai pu construire et conduire ma thèse comme je l'entendais ; en profitant pleinement du droit à l'erreur, en m'autorisant des digressions, en tissant librement des réseaux de collaborations. Avec cette expérience, j'ai aussi acquis la conviction que la possibilité de construire et conduire librement un projet de recherche peut apporter quelque chose d'utile à la société.

Je remercie chaleureusement mon équipe d'encadrement pour m'avoir constamment soutenu de part et d'autre de la Manche : Hervé Monod et Elisabeta Vergu à l'INRA de Jouy-en-Josas ; Chris Gilligan et João Filipe à l'Université de Cambridge. Ce projet de thèse a été l'occasion d'initier un partenariat fructueux entre l'Unité Mathématiques et Informatiques Appliquées du Génome à l'Environnement (MaIAGE) et l'*Epidemiology and Modelling Group* de Chris Gilligan. J'espère qu'il se poursuivra à l'avenir.

Ce travail doctoral, d'apparence hétéroclite, reflète mon aspiration à une approche plus intégrative des sciences, tout particulièrement dans le domaine de la santé. L'étude de la santé humaine, animale et végétale, ainsi que de la résilience des écosystèmes et des systèmes financiers reste encore largement cloisonnée. Il me semble que les mathématiques proposent des outils mobilisables pour transcender les clivages de la santé. C'est tout le sens de mon travail, notamment des dernières approches que j'ai proposées pour mieux appréhender les interactions entre processus épidémiques et économiques.

La recherche contemporaine en « sciences dures » émane essentiellement de collectifs. Mon travail s'inscrit pleinement dans cette dynamique, et j'ai précisé en introduction de chaque chapitre les contributions de mes encadrants et collègues. Outre mes encadrants, je tiens ainsi à remercier Samuel Alizon, Gaël Beaunée, Catherine Belloc, Sebastian Bonhoeffer, Pauline Ezanno, Sébastien Geeraert, Tom Harwood, Patrick Hoscheit, Christel Kamp, Mike J. Jeger et Marco Pautasso. Sans eux, mes travaux n'auraient pu aboutir.

Certains de ces collègues m'ont particulièrement marqué et je tiens à rédiger quelques mots à ce sujet. Les conseils bienveillants de Samuel Alizon et Marco Pautasso ont été déterminants dans nombre

de mes choix scientifiques. Sans eux je n'aurai pas eu la même trajectoire professionnelle. Co-diriger le stage de recherche Sébastien Geeraert a été un privilège. Sa vivacité intellectuelle m'a impressionné. Les échanges et éclairages de Patrick Hoscheit ont été déterminants pour assurer une montée en puissance de nos travaux collectifs en épidémiologie économique. Sa rigueur mathématique et sa détermination forcent l'admiration.

Enfin, j'adresse mille mercis à ma famille et mes amis pour m'avoir soutenu constamment, mais aussi pour avoir nourri mon univers intérieur. Tout particulièrement, je remercie Anna et Anton. Avec eux, ce sera toujours plus qu'hier et moins que demain.

Modéliser les dynamiques épidémiques dans des réseaux d'échanges avec interactions contraintes et comportements adaptatifs

Résumé

Les échanges entre agents tels que les individus ou les entreprises couvrent de nombreux besoins tels que la reproduction et la recherche de profits, mais peuvent dans le même temps faciliter la transmission des maladies infectieuses. En conséquence, les épidémies sont susceptibles de causer des dommages au sein des populations biologiques et d'altérer ces mêmes échanges qui véhiculent les infections. Les modèles épidémiologiques intègrent de plus en plus les comportements d'aversion au risque en réaction aux épidémies. Ces comportements induisent une réduction des contacts infectieux entre agents. Toutefois, les comportements qui sous-tendent les échanges ne se réduisent pas à l'aversion au risque. Les échanges sont conditionnés par des motivations différentes et sont contraints en raison de leurs coûts et du caractère limité des ressources. Face à une émergence épidémique, les comportements adaptatifs se traduisent par des ajustements dans les actions des agents qui peuvent augmenter ou réduire leur exposition au risque d'infection. De multiples mécanismes adaptatifs liés à des processus sociaux et économiques contribuent à expliquer ces ajustements. Les contraintes d'interaction limitent la capacité des agents à échanger de diverses manières. Elles peuvent réduire la capacité d'un agent à interagir avec ses alter ego (contrainte de creux), le taux d'échange (contrainte de pondération), la direction des échanges (contrainte de direction) ou la fréquence de rencontre entre agents (contrainte de friction). Si la contrainte de creux a été largement étudiée, les implications épidémiologiques des trois autres contraintes restent mal comprises. Dans cette recherche, nous combinons analyses empiriques et explorations de modèles mathématiques pour étudier l'influence des contraintes d'interaction et des comportements adaptatifs sur la dynamique conjointe des échanges et de l'infection. Nous pouvons ainsi proposer des politiques de prévention et de maîtrise des épidémies véhiculées par les échanges. Nos études de cas incluent la dynamique des infections sexuellement transmissibles dans des réseaux de contacts sexuels et les épidémies propagées dans des marchés d'échanges d'animaux ou de plantes. Dans un premier temps, nous montrons que la superposition de contraintes d'interaction engendre une grande diversité de structures d'échanges et de dynamiques épidémiques. Nous identifions des conditions analytiques pour lesquelles les contraintes de pondération et de direction limitent la probabilité d'émergence et la sévérité des maladies infectieuses, et traduisons ces conditions théoriques en matière de politiques de santé. Ces résultats sont obtenus en supposant que les agents sont passifs. Dans un second temps, nous prenons également en considération les comportements adaptatifs rencontrés dans les marchés qui véhiculent des infections, et suggérons que la dynamique conjointe de l'infection et des échanges est limitée par la friction marchande. La contrainte de friction engendre un compromis entre fréquence et intensité des transactions commerciales, et peut amoindrir les épidémies de manière plus importante que l'aversion au risque. Nous évoquons finalement des mesures pratiques pour limiter la transmission des maladies infectieuses dans les marchés tout en minimisant les effets délétères sur le commerce. Notre thèse démontre l'importance d'adopter des approches interdisciplinaires pour relever les défis des émergences épidémiques imputables aux échanges. Les idées et outils développés peuvent être transposés à d'autres systèmes, par exemple en écologie ou en économie.

Mots clés : épidémiologie économique ; prévention des maladies infectieuses ; réponse comportementale ; réseaux complexes ; R_0

Modelling epidemic dynamics in exchange networks with constrained interactions and adaptive behaviour

Abstract

Exchanges among agents, typically individuals or companies, fulfil various needs such as reproduction and economic profit, but can also support infectious disease transmission, impacting biological populations and potentially altering the disease-conducive exchanges. Epidemiological models increasingly account for reductions in infectious contact, such as risk-aversion behaviour in response to pathogen outbreaks. However, behavioural responses in exchange dynamics are not limited to risk-aversion ; they are driven by different motivation and are constrained because resources are limited and exchanges are costly. Adaptive behaviour refers to change in agent behaviour, and potentially in exposure to risk, in response to disruption such as disease outbreak ; the change may be through adaptation of social or economic mechanisms. Interaction constraints limit in different ways the capacity of agents to interact. They can limit the interaction of a given agent with everybody else (sparseness constraint), the rate and direction of exchanges (weighting and directional constraints), or the rate of encounter between agents (frictional constraint). While the sparseness constraint has been exhaustively studied, the epidemiological consequences of the three other constraints are poorly understood. Here, we use a combination of empirical analyses and mathematical modelling approaches to explore the influence of interaction constraints and adaptive behaviour on the combined dynamics of exchange and infection. We suggest relevant policies to prevent and mitigate exchange-driven epidemics. The examples investigated are STI dynamics in sexual-contact networks, and epidemics in markets of animal livestock or ornamental plants. We show that differing patterns in agent contact structures and in epidemic dynamics arise when the networks are subject to combinations of interaction constraints. We identify analytical conditions when weighting and directional constraints limit the occurrence and severity of epidemic outbreaks, and translate these threshold conditions into disease-control strategies. These results hold in the case when agents are passive. Moreover, we account for adaptive behaviour encountered in markets that propagate infections and propose that the joint dynamics of markets and disease spread are limited by trade friction. This specific constraint creates a trade-off between the frequency and intensity of market transactions that can influence epidemics more strongly than risk-aversion. We finally suggest policy for limiting disease contagion in markets and minimise its adverse impact on trade. Our work demonstrates that the integration of differing standpoints at the crossroad of natural and social sciences is important in tackling the challenges posed by the emergence of exchange-driven epidemics. We believe our general approach can be transposed to other systems where agents exhibit adaptive behaviour and face interaction constraints, e.g. in ecology or in economics.

Keywords : behavioural response ; complex networks ; disease prevention ; economic epidemiology ; R_0

Table des matières

Remerciements **v**

Résumé **vii**

Abstract **ix**

Liste des figures **xv**

Liste des tableaux **xvii**

Synthèse **1**

General introduction **31**

I Exchange-driven epidemics : overview and challenges **35**

1 Understanding exchange-driven epidemics with simple models **39**
1.1 Modelling exchange-driven epidemics with the SIRS compartmental model 39
1.1.1 Model formulation 40
1.1.2 Model exploration 41
1.2 From the compartmental to the network-based standpoint 44
1.2.1 Opening the compartmental black box 44
1.2.2 The SIRS complete-network-based model 46
1.3 Comparison and assumptions of the compartmental and network-based standpoints 48
1.3.1 Comparison of the two frameworks 48
1.3.2 Key assumptions shared : homogeneous-mixing and passive host behaviour 49

2 Towards a mechanistic description of exchange-driven epidemics **51**
2.1 Accounting for interaction constraints 51
2.1.1 Constraints diversify epidemiological contact structures and raise questions 52
2.1.2 Quantifying constraints in data 56
2.1.3 Constraints are intricate with multiple trade-offs : a bottom-up approach 59
2.2 Accounting for adaptive host behaviour 61
2.2.1 Risk-aversion : a standard example of adaptive behaviour 61
2.2.2 Adaptive behaviour does not reduce to risk-aversion : the case of market behaviour . . . 62
2.2.3 Interaction constraints can be subsumed into adaptive host behaviour 65
2.3 Approaches used to relax assumptions of simple SIRS models 66
2.3.1 Generalising the network-based standpoint 66
2.3.2 Specifying the sub-components of the compartmental framework 69
2.3.3 Choosing the most parsimonious approach 69

II A path from epidemic models with interaction constraints, but passive agents. . . **73**

3 Weighting matters for epidemics. Application to STI dynamics on networks. 77
Abstract . . . 79
Authors summary . . . 80
3.1 Introduction . . . 80
3.2 Materials and Methods . . . 83
3.2.1 The configuration model . . . 83
3.2.2 The epidemiological model . . . 83
3.2.3 Analytical results and their validation . . . 85
3.2.4 Simulations on weighted networks . . . 87
3.2.5 Derivation of r_0 and R_0 . . . 87
3.3 Results . . . 89
3.3.1 Validation of the analytical model with simulations on networks . . . 89
3.3.2 Capturing epidemic characteristics (expressions for r_0 and R_0) . . . 92
3.3.3 Application of the model to epidemiological data . . . 93
3.4 Discussion . . . 96

4 Exchange directionality impacts transmission. The case of *P. ramorum* spread through trade of ornmentals. 99

5 Combining weight and directionality to understand exchange-driven epidemics. Market-epidemiological analyses of livestock movements as a case study. 103
Abstract . . . 106
5.1 Introduction . . . 107
5.2 Materials and Methods . . . 109
5.2.1 Trade networks and livestock-exchange . . . 109
5.2.2 Market-centric categorisation of economic agents for representing the trade networks . . . 111
5.2.3 Elaboration, choice, and evaluation of targeted control strategies . . . 112
5.3 Results and Discussion . . . 115
5.3.1 Analyses and representation of trade networks using market categories . . . 115
5.3.2 Identification of targeted control strategies . . . 119
5.3.3 Evaluation of the targeted control strategies using MCDA . . . 120
5.4 Conclusion . . . 124

III . . .to epidemic models with adaptive behaviour and constrained interactions **127**

6 A mechanistic model to explore the interplay between exchange and epidemic dynamics in markets with homogeneous agents. 131
Abstract . . . 134
6.1 Introduction . . . 135
6.2 Market-epidemiological modelling framework . . . 137
6.2.1 Overview . . . 137
6.2.2 A frictional-trade market (FTM) model . . . 139
6.2.3 Market-epidemiological (ME) model with risk aversion . . . 145
6.3 Insights on market dynamics out-of-steady-state equilibrium and their interaction with epidemics 148
6.3.1 Market dynamics without shocks - effect of trade friction . . . 149
6.3.2 Market dynamics with epidemic shocks . . . 153
6.4 Discussion . . . 157

7 Modelling trade and trade-driven epidemics in markets with heterogeneous agents 163
7.1 A modelling framework to shed lights on trade dynamics . . . 165
7.1.1 The heterogeneous Frictional Trade Market (hFTM) model . . . 165
7.1.2 Analytical insights on the hFTM model . . . 167
7.1.3 Reproducing trade dynamics with the hFTM framework . . . 170
7.1.4 Explaining observed trade dynamics . . . 176
7.2 Preventing epidemics by manipulating elementary friction . . . 181
7.2.1 The heterogeneous market-epidemiological (hME) model . . . 182
7.2.2 Preventing trade-driven epidemics by increasing friction : insights from percolation experiments . . . 184
7.2.3 Epidemic trajectories with modified microscopic frictions . . . 187

IV Discussion **191**

8 A tale of exchanges and epidemics : main insights and prospects 193
8.1 Key results and messages . . . 193
8.1.1 Overall findings . . . 194
8.1.2 Take-home messages . . . 198
8.2 Perspectives . . . 201
8.2.1 Deriving analytical tools for epidemics on weighted directed networks . . . 202
8.2.2 Investigation of global spillovers : the "too important to fail" phenomenon . . . 203
8.2.3 Applying the eco-epi frameworks to other contexts . . . 204
8.2.4 Tailoring the adaptive frameworks to endemic diseases . . . 205
8.2.5 Selecting models against data to assess mechanistic assumptions . . . 205
8.3 Conclusion . . . 207

Annexes **225**

A Research paper : *Weighting for sex acts to understand the spread of STI on networks* 225

B Supporting information for : *Epidemic spread on weighted networks* 235

C Research paper : *SIS along a continuum (SIS_c) epidemiological modelling and control of diseases on directed trade networks* 247

D Supporting information for : *Market analyses of livestock trade networks to inform the prevention of joint economic and epidemiological risk* 257

E Supporting information for : *Epidemics in markets with trade friction and imperfect transactions* 269

F Internship report : *Construction et implémentation d'un modèle d'épidémiologie économique fondé sur des réseaux dynamiques et adaptatifs* 317

Table des figures

1 Le modèle SIRS à compartiments . . . 6
2 Le modèle SIRS sur réseau complet . . . 7
3 Exemples types de contraintes d'interaction influençant les épidémies imputables aux échanges . . . 8
4 Les comportements adaptatifs des hôtes influencent la dynamique conjointe des échanges et des épidémies 9
5 Conséquences topologiques et épidémiologiques de la pondération d'un réseau de contacts sexuels empirique 12
6 Pondération du réseau suivant un modèle de configuration . . . 14
7 Catégorisation des agents économiques : polarité des flux et part des flux . . . 20
8 Réseaux commerciaux représentant les marchés d'échanges de bovins et porcins français . . . 21
9 Cadre général employé pour modéliser le système couplé marché - épidémie . . . 27
10 Le délai maximal de mise en oeuvre des mesures de contrôle permettant encore de prévenir les épidémies dépend de la fluidité du marché et de la présence d'autres voies de transmission . . . 28

1.1 The SIRS compartmental model . . . 43
1.2 The SIRS complete-network-based model . . . 50

2.1 Typical interaction constraints influencing exchange-driven epidemics . . . 53
2.2 Sexual-contact data suggest the existence of a constant weighting constraint per agent . . . 58
2.3 Livestock-exchange data imply the existence of friction . . . 59
2.4 Adaptive host behaviour influence joint exchange and epidemic dynamics . . . 64
2.5 Epidemics on sparse networks with contrasted degree distributions . . . 70

3.1 Weighting between contacts . . . 82
3.2 Dynamics of the number of infected hosts (I) on different types of networks . . . 91
3.3 Disease spread on a network inferred from data . . . 95

5.1 Categorisation of economic agents : flow polarity and flow share . . . 117
5.2 Market categories defined according to flow polarity and flow share and used to assess joint economic-epidemiological risk . . . 118
5.3 Trade networks of the cattle and swine livestock markets . . . 119
5.4 Multiple-criteria decision analyses (MCDA) of contrasting targeted control strategies in the cattle livestock market . . . 122
5.5 MCDA of contrasting targeted control strategies in the swine livestock market . . . 123

6.1 Joint market-epidemiological modelling framework . . . 138
6.2 The influence of frictional-trade on transient and long-term market dynamics . . . 152
6.4 Maximal delay in enforcement of regulation that still allows prevention of epidemics depending on market fluidity and other paths of transmission . . . 156

7.1 Market dynamics for contrasted scenarios on initial stocks of supply and demand 172
7.2 Comparison of equilibrium values of supply and demand for the four scenarios tested 173
7.3 Typical outcomes of scenario $\mathscr{A}$ 175
7.4 Equilibrium values of q_{ij} and q_{ij}^{new} as function of observed values $\hat{q}_{ij}$ 177
7.5 Agents' partitioning preferences as function of flow polarity and flow share 179
7.6 Convergence of the adaptive partitioning process 181
7.7 Evolution of the critical time T^* when we increase friction κ_{ij} 186
7.8 Size of the LSCC and $(Y + Z)$ as function of time for estimated friction 188
7.9 Size of the LSCC and $(Y + Z)$ as function of time for increased friction 189
7.10 Size of the LSCC and $(Y + Z)$ as function of time for decreased friction 190

Liste des tableaux

3.1 Notations used in chapter 3 85
3.2 r_0 vs. R_0 93

6.1 Relationship between terminology in the frictional-trade market (FTM) model and in economics 144
6.2 Overview of the models investigated in our study 148
6.3 Parameter values calculated for French cattle and swine markets 149

7.1 Linear regressions assuming reference partitioning 176
7.2 Linear regressions assuming adaptive partitioning 180

8.1 Pros and cons of neglecting or accounting for adaptive behaviour 200

Als das Kind Kind war,
wußte es nicht, daß es Kind war,
alles war ihm beseelt,
und alle Seelen waren eins.

Peter Handke, *Lied Vom Kindsein*

Synthèse

Cette synthèse résume l'organisation et les principaux apports de ma thèse à destination d'un public francophone. Dans la mesure où ce projet de doctorat a été élaboré et conduit en co-direction entre la France et le Royaume-Uni, l'anglais s'est imposé comme langue de rédaction pour la majorité du manuscrit. Quoique substantielle, cette synthèse n'apporte qu'une clé de lecture à ce travail et ne permet pas d'en prendre toute la mesure, tout particulièrement sur le plan des méthodes mobilisées pour parvenir aux résultats annoncés. Une lecture attentive des chapitres du manuscrit reste nécessaire pour apprécier la portée et la robustesse des analyses et recommandations proposées.

Introduction générale

Sous l'effet conjoint de la mondialisation et de l'accroissement démographique, les risques posés par la propagation des maladies infectieuses sont considérables, comme l'illustrent les risques d'émergence épidémique à partir de réservoirs naturels et de pandémie d'ampleur planétaire [Jones *et al.*, 2013]. Les hôtes vecteurs d'infection peuvent se déplacer sur des distances et à des volumes d'échange toujours plus élevés [Tatem *et al.*, 2006]. À la différence des maladies non-infectieuses, les infections sont transmises d'individu à individu. Bien qu'il existe de nombreuses voies de transmission, les échanges directs entre hôtes constituent des moteurs épidémiques majeurs, et la probabilité qu'un hôte sain contracte une infection croît avec le nombre d'hôtes infectieux dans son voisinage direct [Keeling & Rohani, 2008]. Les structures de contact entre hôtes qui résultent d'échanges sont généralement riches sur le plan topologique [Pellis *et al.*, 2014], recouvrent des échelles de temps et d'espace contrastées [Gog *et al.*, 2014], et peuvent être considérablement remaniées sous l'effet de facteurs comportementaux et socio-économiques [Klein *et al.*, 2007; Funk *et al.*, 2014; Perrings *et al.*, 2014]. Cette complexité implique que la compréhension des interactions entre les facteurs responsables des échanges et les facteurs gouvernant les dynamiques épidémiques constituent un front de science fertile. Outre la satisfaction intellectuelle apportée, défricher ce type de front de science permet de proposer des stratégies pour maîtriser les épidémies [Keeling & Rohani, 2008]. Dans ce contexte, mon travail doctoral contribue à éclairer les mécanismes rendant compte des *épidémies véhiculées par les échanges*, qualifiées également *d'épidémies imputables aux échanges*. Ce travail met ainsi en lumière des leviers susceptibles d'être actionnés pour limiter la transmission des infections causées par les échanges.

Les échanges qui résultent d'interactions générées à des fins anthropiques se révèlent cruciaux pour notre bien-être. Prenons appui sur deux exemples largement développés dans cette thèse. À l'échelle des individus, les actes sexuels contribuent, entre autres, aux activités récréatives et à la procréation [Foucault, 1984]. À l'échelle des entreprises, les livraisons de bétail constituent une composante clé des filières agricoles et alimentaires [Vernon, 2011]. Ces deux cas pratiques correspondent à des échanges bénéfiques réalisés entre *agents*, agents que nous définissons dans ce travail comme des êtres humains ou des unités contrôlées par des êtres humains telles que des entreprises agricoles. Cette définition des agents peut être rapprochée de celle adoptée dans la mise en oeuvre de modèles numériques qualifiés de « systèmes multi-agents » (SMA). Pour ces derniers, un agent se comprend comme une entité autonome, réelle ou virtuelle, cherchant à accomplir ses propres objectifs et qui est capable d'agir sur son environnement et d'interagir avec d'autres agents. Les SMA permettent notamment de rendre compte de certains phénomènes économiques et épidémiques à travers d'algorithmes de décision visant à reproduire avec un degré de réalisme élevé le comportement des agents réels [voir Tesfatsion, 2001; Durham & Casman, 2012, et Guesnerie [1996] cité dans Callon [1998]]. Les SMA peuvent aussi rendre compte d'arbitrages inter-temporels complexes rendus par des agents économiques. Par contraste, notre travail mobilise des agents plus simples dans leurs comportements, ce qui à l'avantage de permettre la mise en oeuvre d'analyses mathématiques plus poussées que pour les SMA.

Bien que les échanges apportent des bienfaits indéniables, ils peuvent infecter des agents qui deviennent ainsi des hôtes infectieux. Ce mécanisme de contagion peut maintenir des épidémies localement et induire des contaminations à longue distance. Les épidémies véhiculées par les échanges recouvrent entres autres le Virus de l'immunodéficience humaine à travers les relations sexuelles [May & Anderson, 1987], la Fièvre aphteuse et la Fièvre catarrhale ovine à travers les mouvements de bétail [Ferguson *et al.*, 2001a], ainsi que la propagation de *Phytophthora ramorum* à travers des livraisons de plantes d'ornement [Pautasso & Jeger, 2008]. Aux interfaces de la santé humaine et de la santé animale, un exemple d'épidémie véhiculée par les échanges d'animaux est donné par la zoonose imputable à la bactérie *Escherichia coli* souche O157 qui infecte conjointement les hommes et les bovins. Le système digestif des bovins constitue le principal réservoir d'*E. coli* O157, chez qui la bactérie est asymptotique. Chez l'homme, une infection par *E. coli* O157 induit de sérieux troubles gastro-intestinaux et peut dans certains cas se révéler mortelle. Une infection humaine est généralement contractée par la consommation de viande ou d'eau contaminée, ou par le contact avec des excréments de bétail dans l'environnement [Matthews *et al.*, 2013]. Ces quelques exemples illustrent les impacts majeurs des épidémies véhiculées par les échanges, et cela pour un large spectre d'hôtes.

Réciproquement, les maladies infectieuses et les mesures de contrôle associées peuvent nuire aux agents, avec des répercussions sur la dynamique des échanges [Funk *et al.*, 2009]. À titre d'exemple, en 2001, une grande épidémie de Fièvre aphteuse a durement frappé les éleveurs au Royaume-Uni. Le bilan établit en 2002 par la Cour des comptes britannique donne la mesure de la crise [National

Audit Office, 2002] : plus de 6 millions d'animaux ont été abattus. Une large majorité de ces abattages (80 %) n'étaient pas liés à des animaux dont la contamination était avérée. Deux causes expliquent ce paradoxe : les abattages à titre préventif (44 %) et les abattages d'animaux sains, mais immobilisés à cause des mesures de restriction des échanges (36 %). Ces choix de gestion ont été sévèrement critiqués par certains acteurs impactés.

En raison des nombreuses boucles de rétroaction à l'oeuvre, démêler les liens entre épidémies et échanges sous-jacents n'est pas chose aisée. Pour maîtriser les épidémies induites par les échanges sans déstabiliser ces derniers, il est capital de mettre à jour les mécanismes fondamentaux qui gouvernent la dynamique conjointe des échanges et de la transmission des maladies infectieuses. Déjà conséquent, le bilan de l'épidémie de Fièvre aphteuse de 2001 aurait pu l'être encore davantage sans le recours aux modèles mathématiques prédictifs utilisés pour guider la gestion de crise dès son apparition [Ferguson *et al.*, 2001b; Keeling *et al.*, 2001]. Plus généralement, comprendre et gérer les épidémies imputables aux échanges serait difficile, voire impossible, sans recourir aux modèles mathématiques [Anderson & May, 1991]. Puisque par définition les épidémies se propagent à l'échelle de populations, il est impossible de les étudier expérimentalement pour des raisons d'ordres éthique et pratique. Bien évidemment, les pathogènes responsables des épidémies et leurs dynamiques infectieuses au sein des hôtes infectés peuvent être analysés en laboratoire. Mais les épidémies ne sauraient se réduire à des considérations strictement biologiques. Outre la dispersion naturelle à travers l'environnement telle qu'elle est observée pour les pathogènes véhiculés par l'air, les processus infectieux impliquent fréquemment des interactions complexes entre hôtes comme les relations sexuelles ou les relations commerciales [Keeling & Eames, 2005]. Les interactions directes entre hôtes constituent des voies de transmission efficaces des pathogènes. La structure des contacts responsables de la propagation des épidémies, qu'elle soit induite par des mécanismes naturels ou sociaux, est habituellement qualifiée de *structure de contact épidémique*. Les modèles se révèlent très pratiques pour décrire avec justesse, comparer ou même inférer de telles structures de contact. Les modèles peuvent également simuler des processus infectieux propagés sur toute structure de contact, que cette dernière soit postulée ou déduite d'observations. Les données historiques d'émergences épidémiques ou d'échanges sociaux sont nécessaires pour construire, améliorer et sélectionner des modèles. Réciproquement, la collecte des données implique souvent un cadre conceptuel implicite qui peut être formalisé par un modèle.

Dans ce travail structuré en quatre parties principales, des modèles mathématiques et des analyses empiriques sont mis en oeuvre pour améliorer la compréhension des épidémies propagées par les échanges et proposer des politiques de contrôle adaptées.

La partie I dresse le portrait des épidémies imputables aux échanges. Les défis posés par ce type d'épidémies et les approches classiques mobilisées pour les étudier sont évoqués, tout particulièrement les *modèles à compartiments* [Kermack & McKendrick, 1927] et les *réseaux* [Strogatz, 2001]. Deux

lacunes de recherche sont identifiées, à savoir l'influence des *contraintes d'interaction* et du *comportement adaptatif (des hôtes)* sur la dynamique conjointe des échanges et des épidémies. L'expression comportement adaptatif désigne ici le changement de comportement des agents, potentiellement en vue de réduire leur exposition au risque, en réponse à des perturbations telles que des émergences épidémiques. La modification de comportement peut notamment résulter d'adaptations de nature économique ou sociale [Funk *et al.*, 2014]. Les contraintes d'interaction limitent de différentes façons la capacité des agents à interagir. Elles peuvent par exemple restreindre un agent dans sa capacité à échanger avec l'ensemble des autres agents (contrainte de creux [1]), le taux et la direction avec lesquels se produisent les échanges (contraintes de pondération et de direction respectivement), ou le taux de rencontre entre agents (contrainte de friction). Si la contrainte de creux a déjà été largement étudiée [Danon *et al.*, 2011], les conséquences épidémiologiques des trois entre contraintes, considérées séparément ou conjointement, restent moins bien comprises [Meyers *et al.*, 2006; Pellis *et al.*, 2014].

Organisées suivant une approche ascendante, les parties II et III présentent des recherches nouvelles qui contribuent à combler les lacunes de connaissance identifiées dans la partie I. Nous commençons par émettre l'hypothèse restrictive que les agents sont non-adaptatifs pour nous concentrer sur la seule étude des contraintes d'interaction et leurs conséquences sur les dynamiques épidémiques (partie II). Puis, dans le cas général, nous étudions l'influence des comportements adaptatifs avec interactions contraintes sur la dynamique conjointe des échanges et des épidémies (partie III).

Enfin, la partie IV récapitule et met en discussion les principaux apports de cette thèse, et esquisse des pistes de travail pour des travaux à venir.

Présentation de la partie I – épidémies imputables aux échanges : vue d'ensemble et défis

De prime abord, comprendre les épidémies imputables aux échanges n'est pas chose aisée. Il s'agit non seulement de rendre compte de la dynamique des échanges et de la dynamique infectieuse prises séparément, mais également des interactions entre ces deux processus dynamiques. En complément des analyses de données, les modèles mathématiques peuvent apporter un appui pour identifier les principaux mécanismes qui sous-tendent ces épidémies.

En nous nourrissant des corpus bibliographiques de l'épidémiologie et de l'économie, nous combinons approches empiriques et approches de modélisation dynamique pour explorer l'architecture des échanges et son influence sur la dynamique infectieuse. Réciproquement, nous étudions les boucles de rétroaction par lesquelles les épidémies peuvent influencer les échanges – soit ces mêmes échanges qui contribuent à leur propagation. Nous nous plaçons dans le cas général où les échanges sont bornés par des contraintes et où les hôtes potentiellement infectieux s'adaptent à la dynamique conjointe des

1. le terme est ici employé en référence aux matrices creuses (*sparse matrices* en anglais).

échanges et des épidémies. Nous présentons un faisceau d'arguments indiquant que la maîtrise des épidémies imputables aux échanges nécessite de mettre en oeuvre des approches multidisciplinaires, avec par ailleurs la prise en compte de points de vue contrastés suivant les acteurs concernés. Nous illustrons notre propos à travers deux types d'échanges véhiculant des épidémies : le commerce et le sexe.

La première partie (page 37), composée de deux chapitres, est organisée suivant une logique descendante. Au premier chapitre, nous apportons des éléments de contexte sur des modèles simples permettant de décrire de façon grossière les épidémies et les échanges sous-jacents. Nous présentons deux approches complémentaires largement employées en épidémiologie : les modèles par compartiments et les modèles mobilisant des réseaux. Les principaux enseignements tirés de l'étude comparative de ces modèles simples sont soulignés. Au deuxième chapitre, nous adoptons un point de vue mécaniste. Nous expliquons comment des modèles simples peuvent voir leur champ d'application élargi de manière à rendre compte de deux phénomènes contribuant à la dynamique conjointe des échanges et des épidémies : les contraintes d'interaction et les comportements adaptatifs. Pour chacun de ces deux phénomènes, nous mettons en exergue des questions de recherche et des enjeux pratiques, et décrivons succinctement les concepts et points de vue mathématiques adoptés pour les étudier. Par ailleurs, des états de l'art détaillés sont proposés pour chaque chapitre de résultats (voir parties II et III), où des problématiques scientifiques plus ciblées sont abordées.

Synthèse du chapitre 1 – comprendre les épidémies imputables aux échanges en mobilisant des modèles simples

Les épidémies imputables aux échanges constituent des phénomènes dynamiques complexes et potentiellement influencés par une myriade de facteurs, ce qui rend leur étude difficile. Dans un tel cas, des modèles mathématiques même simples peuvent être mobilisés. Ces modèles permettent de décrire des processus infectieux dynamiques véhiculés par des échanges et d'identifier des mécanismes potentiels qui sous-tendent ce type de transmission épidémique.

Dans le premier chapitre (voir page 39), un modèle épidémiologique parmi les plus simples est présenté et analysé : le *modèle SIRS à compartiments* (voir figure 1). Malgré sa simplicité, ce modèle apporte des éclairages intéressants sur les épidémies imputables aux échanges. Les hypothèses qui sous-tendent le modèle SIRS à compartiments sont ensuite explicitées en adoptant le formalisme de la théorie des réseaux. Le modèle SIRS à compartiments est mis en regard avec un cadre bien établi que nous qualifions de *modèle SIRS sur réseau complet* (voir figure 2). Dans ce dernier modèle, les échanges entre agents sont décrits par une architecture particulière connue sous le nom de réseau complet et où les infections successives correspondent à des événements aléatoires se produisant suivant la structure particulière de ce réseau. Enfin, il est montré que ces deux modèles produisent, en moyenne, des résultats identiques, et partagent des hypothèses mécanistes clés, à savoir le *mélange homogène* et le

comportement passif des hôtes (voir page 48). Les deux types de modèles sont mobilisés dans ce travail et peuvent être étendus pour rendre compte de mécanismes de contact inter-agents plus subtils (voir chapitre 2 à la page 51).

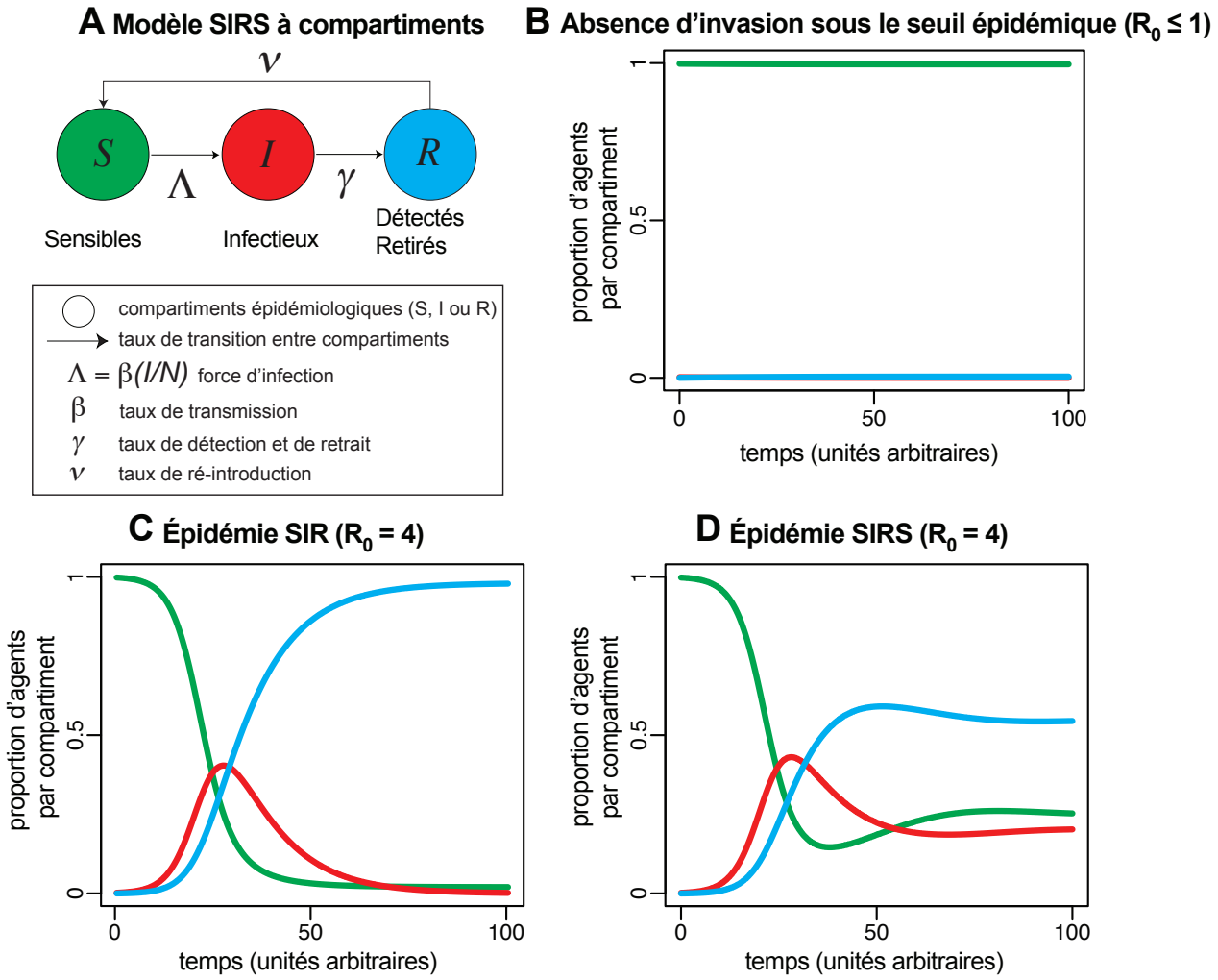


FIGURE 1 – **Le modèle SIRS à compartiments.** Structure du modèle (**A**) et simulations caractéristiques issues de ce dernier (**B-D**). Les agents sont sains et sensibles à l'infection (S, en vert), infectés et infectieux (I, en rouge) ou détectés et retirés (R, en bleu). Les agents sensibles deviennent infectieux au taux $\Lambda = \beta \cdot (I/N)$, où β représente le taux de transmission de l'infection et $N = S + I + R$ le nombre total d'agents supposé constant. Les agents infectieux sont détectés et retirés au taux γ. Les agents retirés sont ré-introduits au taux ν. R_0, ici donné par le rapport β/γ, représente le nombre de reproduction de base, soit le nombre de cas directement engendrés par un infectieux moyen sur l'ensemble de sa période infectieuse (ici $1/\gamma$), dans une population entièrement constituée d'agents sensibles ($I \approx N$). Le R_0 doit être strictement supérieur à 1 pour qu'une épidémie se produise. Les trajectoires épidémiques (**B-D**) représentent S/N (en vert), I/N (en rouge) et R/N (en bleu) en fonction du temps. Nous donnons à voir trois sorties caractéristiques de ce modèle SIRS : une absence d'épidémie (**B**, avec $\beta = 0.05 < \gamma = 0.1$), une épidémie de type SIR (**C**, avec $\beta = 0.4 > \gamma = 0.1$ et $\nu = 0$) et une épidémie du type SIRS (**D**, avec $\beta = 0.4 > \gamma = 0.1$ et $\nu = 0.04$). Les taux sont exprimés en unités arbitraires. Dans chaque cas, 20 agents sont infectés au début des simulations ($I(0) = 20$) dans une population totale de $N = 10000$ agents.

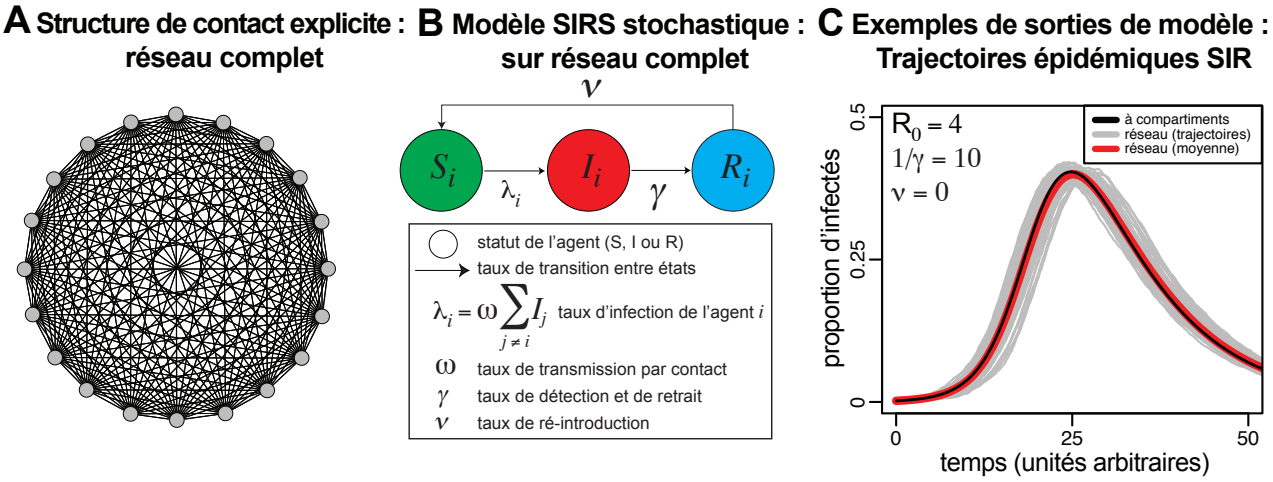


FIGURE 2 – **Le modèle SIRS sur réseau complet.** Le modèle SIRS sur réseau complet couple un réseau complet (**A**) à un processus stochastique d'épidémie sur réseau (**B**), permettant ainsi de mettre en lumière les hypothèses implicites qui sous-tendent le modèle SIRS à compartiments (**C**). Chaque agent, représenté par i prenant une valeur comprise entre 1 et N, est sensible ($S_i = 1$), infectieux ($I_i = 1$) ou détecté et retiré ($R_i = 1$). Par construction, $S_i + I_i + R_i = 1$, et nous négligeons les processus démographiques, ce qui implique une population totale d'agents N constante. Les agents sensibles deviennent infectés et infectieux au taux λ_i^c (1.3). Les agents infectieux sont détectés et retirés au taux γ. Les agents retirés sont ré-introduits au taux ν. $R_0 = \beta/\gamma = N\omega/\gamma$, représente le nombre de reproduction de base, où β est le taux de transmission du modèle SIRS à compartiments et ω une constante représentant le taux de transmission *par contact* du modèle SIRS sur réseau complet. On peut remarquer la correspondance $\beta = N\omega$ [Aparicio & Pascual, 2007]. Nous comparons les dynamiques du type SIR d'une simulation déterministe issue du modèle à compartiment d'une part (**C**, en noir) ; à 100 simulations stochastiques issues du modèle sur réseau complet (**C**, en gris pour les trajectoires individuelles et en rouge pour la moyenne des 100 simulations indépendantes). Le processus stochastique est mis en oeuvre avec l'algorithme direct de Gillespie [Keeling & Rohani, 2008]. Pour les deux modèles, 20 agents sont infectés au début des simulations ($I(0) = 20$; avec infection uniformément au hasard pour le modèle stochastique) avec $\gamma = 0.1$, $\nu = 0$ et $N = 10000$. Nous prenons les valeurs de β et ω pour lesquelles $R_0 = 4$. Les taux sont exprimés en unités arbitraires.

Synthèse du chapitre 2 – vers une description mécaniste des épidémies imputables aux échanges

Les modèles SIRS simples sont flexibles et apportent des éclairages d'intérêt sur les épidémies imputables aux échanges. Qu'ils soient à compartiments ou sur réseau, de tels modèles simples partagent deux mécanismes implicites pour rendre compte des dynamiques infectieuses, à savoir le mélange homogène et le comportement passif des hôtes. Prises ensemble, ces hypothèses correspondent à une structure de contact épidémique particulière : un réseau complet qui connecte en permanence l'ensemble des agents sensibles et infectieux. La nature panmixique de ces contacts reste pour l'essentiel invariante, y compris lorsque les agents sont détectés et retirés ou lorsqu'ils sont ré-introduits au sein du réseau (chapitre 1). L'intuition suggère que le mélange homogène et le comportement passif des hôtes constituent des mécanismes peu vraisemblables pour décrire de façon réaliste les épidémies imputables aux échanges.

Dans le deuxième chapitre (voir page 51), nous proposons une revue de littérature et détaillons les arguments empiriques allant dans le sens d'une réfutation des hypothèses de mélange homogène et

de comportement passif des hôtes. Dans une tentative pour apporter une description mécaniste des épidémies imputables aux échanges, nous nous focalisons sur deux mécanismes influençant la dynamique conjointe des échanges et des infections : les *contraintes d'interaction*, qui tendent à restreindre le mélange homogène entre agents et induisent par définition des interactions contraintes (figure 3) ; et les comportements adaptatifs des hôtes, qui évitent de recourir à l'hypothèse restrictive de comportement passif (figure 4). Nous expliquons ensuite comment les modèles SIRS simples, qu'ils soient à compartiments ou sur réseau, peuvent être généralisés pour rendre compte des contraintes d'interaction et des comportements adaptatifs et ainsi apporter des éclairages complémentaires (voir page 66). L'étude approfondie des contraintes d'interaction, ainsi que des comportements adaptatifs subsumant des interactions contraintes, constituent respectivement le coeur des parties II et III.

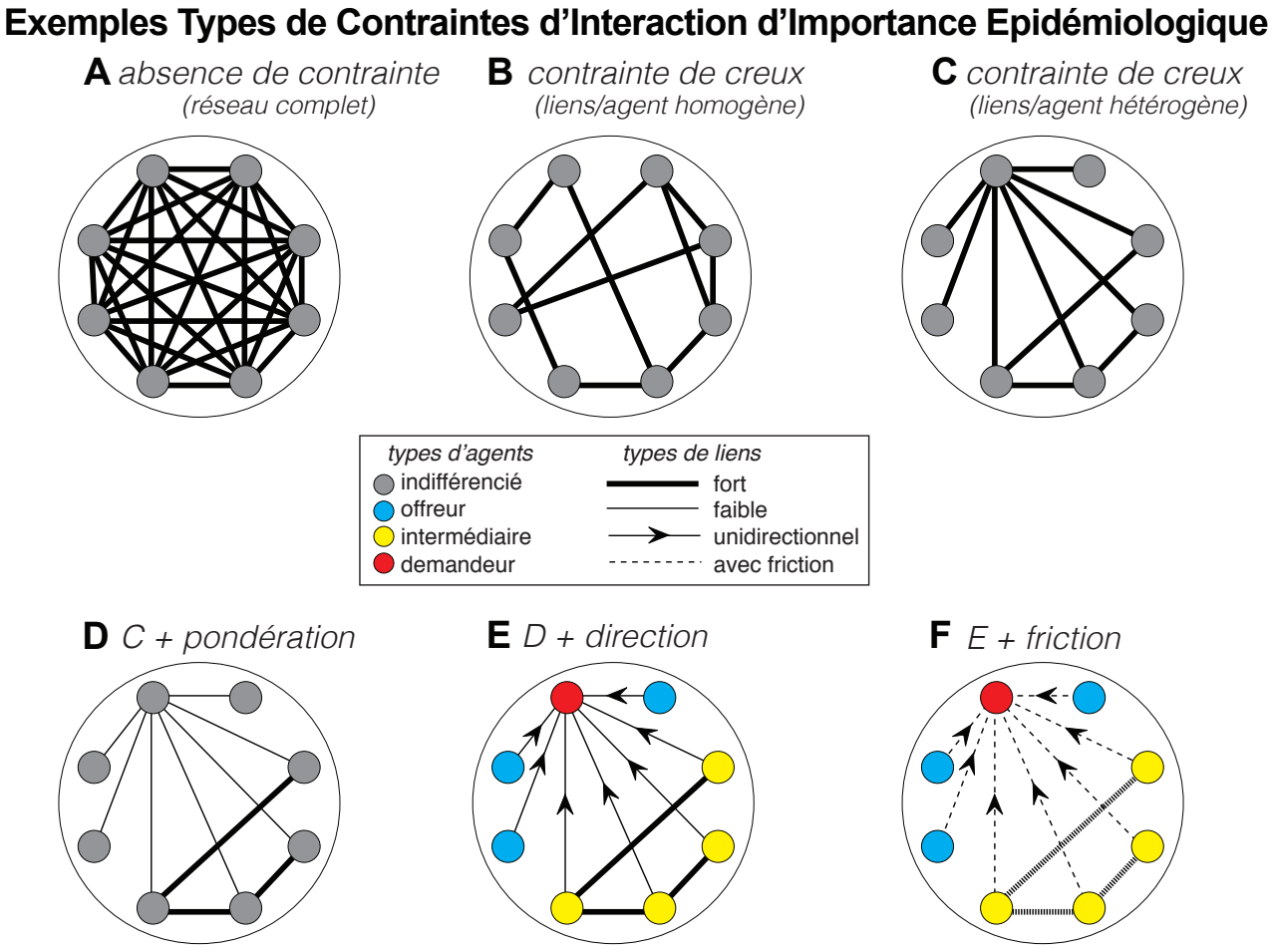


FIGURE 3 – **Exemples types de contraintes d'interaction influençant les épidémies imputables aux échanges.** À l'opposé de la structure de réseau complet qui sous-tend les modèles SIRS simples (**A**), les architectures des échanges observées présentent fréquemment de multiples *interactions contraintes*. Ces interactions contraintes résultent de *contraintes d'interaction* et contribuent à expliquer les chaînes de transmission infectieuses induites par les échanges. Des contraintes d'interaction types incluent : la *contrainte de creux*, où seulement une faible part de l'ensemble des liens possibles existe, ce qui conduit à des distributions des liens par agent homogènes (**B**) ou hétérogènes (**C**) ; la *contrainte de pondération*, avec des liens de poids hétérogènes (**D**) ; la *contrainte de direction*, où les directions des liens sont susceptibles de refléter différents types d'agents tels que les offreurs, les intermédiaires et les demandeurs au sein des marchés (**E**) ; et la *contrainte de friction*, qui implique que les liens n'existent qu'à certains moments dans le temps (**F**).

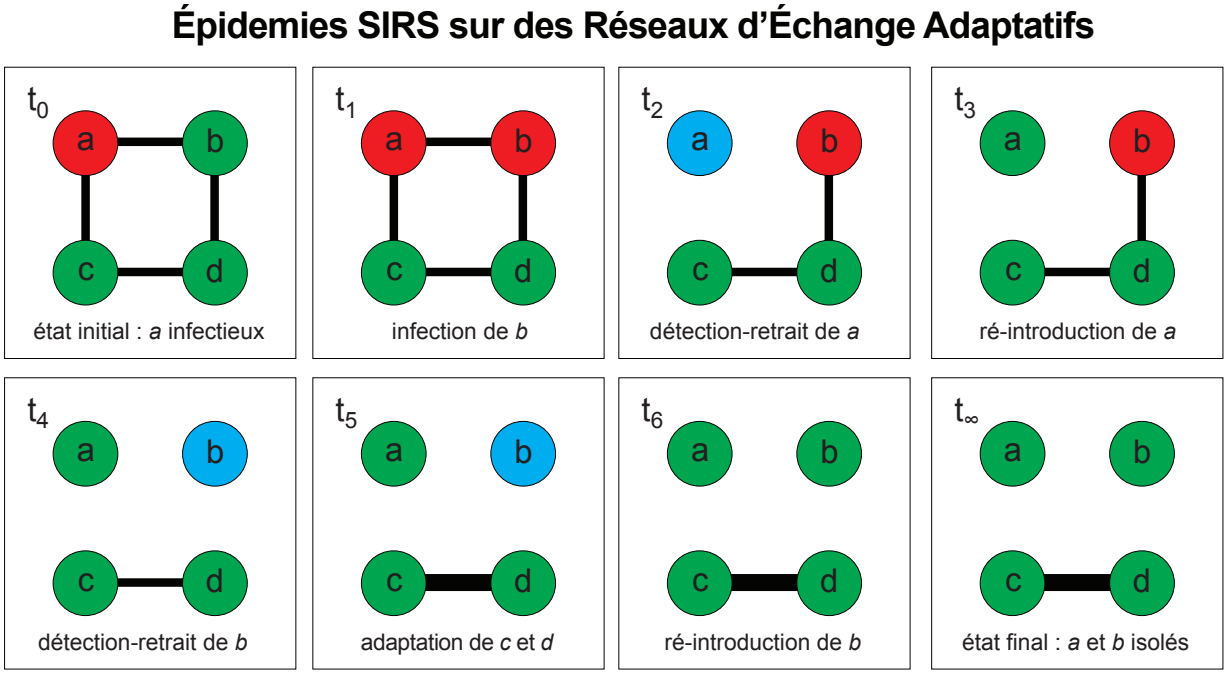


FIGURE 4 – **Les comportements adaptatifs des hôtes influencent la dynamique conjointe des échanges et des épidémies.** Nous considérons un processus infectieux propagé sur un réseau d'échange adaptatif composé de quatre agents a, b, c et d. Les huit panneaux (t_0 – t_∞) représentent un épisode complet d'émergence épidémique SIRS (statut S en vert, I en rouge, R en bleu), depuis une infection initiale de a au temps t_0, jusqu'à l'état final du système au temps t_6 – t_∞.

Présentation de la partie II – des modèles épidémiologiques rendant compte d'interactions contraintes, mais avec des agents passifs

Explorer l'influence de différentes structures de réseaux sur des processus épidémiques sus-jacents constitue un défi de longue haleine pour l'épidémiologie des réseaux [Strogatz, 2001]. Les structures de contact induites par les échanges se révèlent d'une grande richesse. Cette richesse s'explique notamment par les combinaisons des contraintes auxquelles sont confrontés les agents en interaction (section 2.1). On peut ainsi imaginer que ces contraintes jouent un rôle déterminant sur les épidémies. Les interactions contraintes résultent des limites auxquelles sont confrontés les agents dans leurs capacités d'échanges. Ces limites peuvent s'appliquer au nombre de partenaires par agent (contrainte de creux), aux taux et à la direction des échanges (contraintes de pondération et de direction) ou au taux de rencontre avec d'autres agents (contrainte de friction). Bien que la contrainte de creux ait été largement étudiée, l'influence épidémiologique des trois autres contraintes, ainsi que les interactions à l'oeuvre entre ces quatre contraintes, restent largement inconnues [Meyers *et al.*, 2006; Pellis *et al.*, 2014].

Dans la deuxième partie (page 75), nous analysons de manière approfondie la manière dont les contraintes d'interaction influent les épidémies avec un niveau d'échange agrégé maintenu constant. Nous nous plaçons dans le cas où les réseaux d'échanges ne sont pas affectés par les comportements

des agents. Cette hypothèse simplificatrice nous apporte davantage d'éclairages mathématiques sur les facteurs structurels qui gouvernent la dynamique infectieuse. L'identification des facteurs structurels déterminant la transmission des épidémies joue un rôle clé pour la mise en oeuvre de politiques de contrôle. Nous commençons par étudier les implications épidémiologiques de la contrainte de pondération, contrainte qui se révèle essentielle pour comprendre la dynamique des infections sexuellement transmissibles (chapitre 3). Dans un second temps, nous explorons l'influence de la contrainte de direction sur les processus infectieux, avec une focale sur les épidémies imputable aux échanges marchands de plantes d'ornement (chapitre 4). Enfin, nous montrons que l'analyse combinée des contraintes de pondération et de direction peut se révéler fructueuse pour évaluer et prévenir les risques économiques et épidémiologiques affectant les marchés d'échange de bétail (chapitre 5). Dans la mesure où l'appréhension de la contrainte de friction repose sur une modélisation explicite de la dynamique des échanges, cette contrainte est analysée dans le cas où les agents présentent des comportements adaptatifs (partie III).

Synthèse du chapitre 3 – pondérer les réseaux pour comprendre les épidémies ; l'exemple des IST.

Le troisième chapitre (voir page 77) présente les implications épidémiologiques de la contrainte de pondération. Ce chapitre se fonde sur deux articles de recherche publiés dans les journaux *Journal of Theoretical Biology* [Moslonka-Lefebvre *et al.*, 2012a] et *PLoS Computational Biology* [Kamp *et al.*, 2013]. Seul l'article [Kamp *et al.*, 2013] est détaillé dans le corps de texte puisque cette étude a été intégralement conduite dans le cadre la présente thèse. Il convient de souligner que Christel Kamp a joué un rôle déterminant dans la conduite de cette dernière étude, puisqu'elle y a apporté les concepts et analyses mathématiques qui en forme la quintessence. Mes principales contributions à [Kamp *et al.*, 2013] sont les suivantes : la motivation épidémiologique de l'étude, l'analyse des données de sondage, la comparaison d'approches contrastées pour générer des réseaux pondérés à partir de distributions de probabilité conjointes en nombre de liens par agent et en poids des agents, et la mise en oeuvre d'explorations numériques de processus infectieux sur réseaux pondérés pour confirmer les prédictions analytiques de l'étude.

Dans [Moslonka-Lefebvre *et al.*, 2012a] (voir Figure 5 et annexe A), nous montrons que la prise en compte du caractère pondéré des relations sexuelles se révèle cruciale pour comprendre la dynamique des infections sexuellement transmissibles (IST) à l'oeuvre sur des réseaux. Les réseaux de contacts sexuels humains présentent une structure hétérogène où quelques rares individus ont de nombreux partenaires tandis que la grande majorité des individus ont peu de partenaires. La théorie des réseaux prédit que la propagation des IST sur de tels réseaux devrait présenter certaines propriétés remarquables telles qu'une propagation rapide et un seuil épidémique pratiquement nul. Toutefois ces propriétés prédites ne sont pas corroborées par les données épidémiologiques. Les modèles de ré-

seaux actuels font l'hypothèse simplificatrice que le risque de transmission d'une IST est constant pour chaque couple de partenaires sexuels. Cette hypothèse n'est pas réaliste puisqu'elle implique que l'activité sexuelle est proportionnelle au nombre de partenaires et que les individus ont le même niveau d'activité avec chacun de leurs partenaires. Nous développons un cadre théorique nous permettant de pondérer n'importe quel réseau de contacts sexuels sur la base d'hypothèses biologiques en cohérence avec les données disponibles. Nos résultats indiquent que la propagation des IST sur les réseaux pondérés avec notre approche perd les propriétés associées aux réseaux hétérogènes non-pondérés, ce qui va dans le sens des données et d'études antérieures. De façon plus générale, notre étude suggère qu'altérer la distribution des poids par agent sur un réseau donné sans modifier le poids agrégé de ce réseau peut avoir des conséquences déterminantes sur les dynamiques épidémiques.

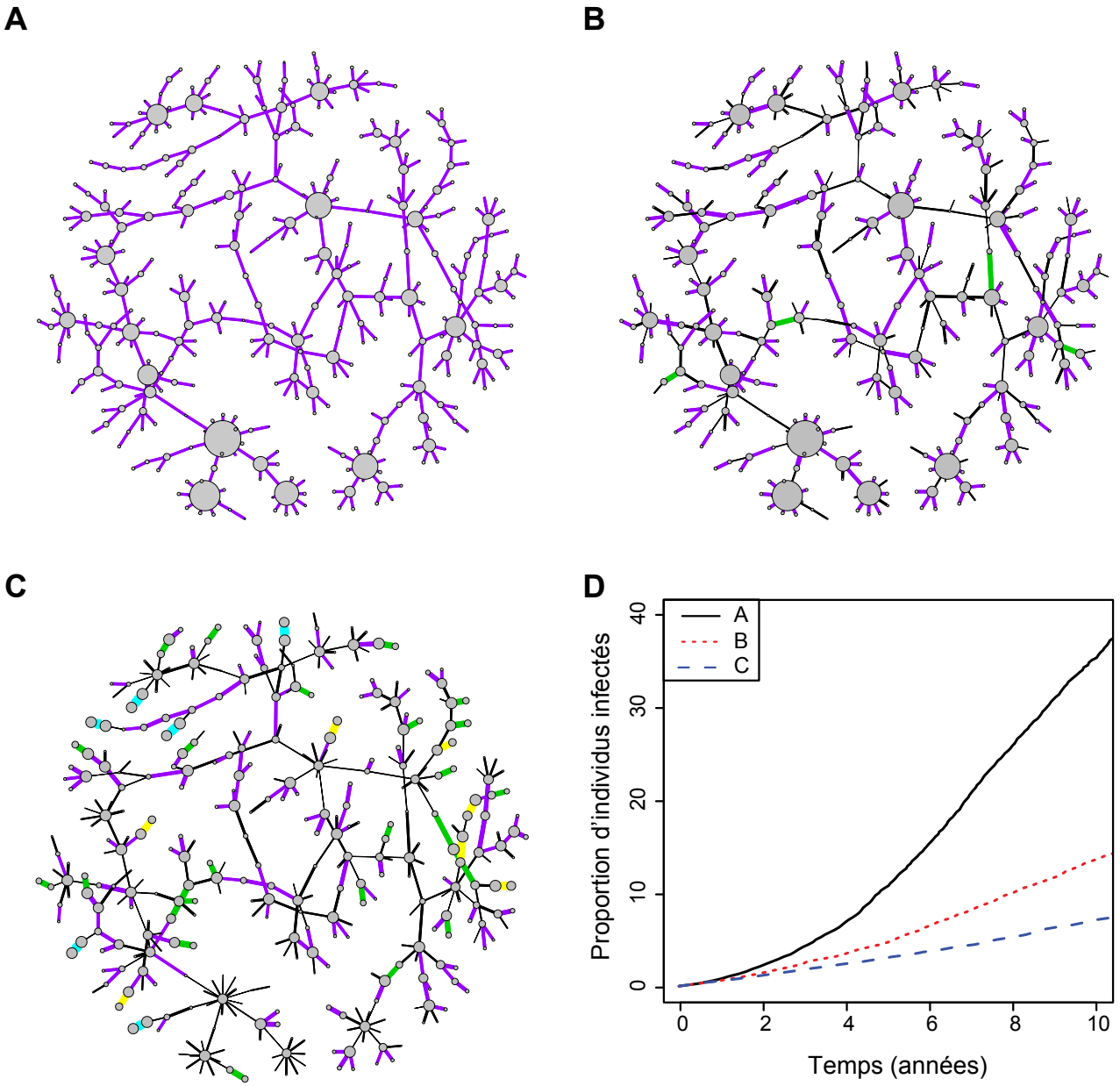


FIGURE 5 – **Conséquences topologiques et épidémiologiques de la pondération d'un réseau de contacts sexuels empirique. A**) Réseau de contact empirique obtenu sur la base des hypothèses de pondération implicites des modèles de réseaux usuels, à savoir avec un nombre d'actes sexuels par individu proportionnel au nombre de partenaires par individu et une activité identique pour chaque relation. **B**) Comme pour **A**), mais avec un nombre d'actes sexuels par relation tiré uniformément au hasard. **C**) Comme pour **A**), mais avec un nombre d'actes sexuels par individu croissant avec le nombre de partenaires par individu et avec un effet de saturation tel qu'observé dans les données, et un nombre d'actes sexuels par relation tiré uniformément au hasard. Pour les panneaux A, B, C, le nombre total d'actes sexuels agrégés sur le réseau est maintenu constant, le diamètre des noeuds (l'épaisseur des liens respectivement) représente le nombre d'actes sexuels réalisé par un individu par semaine (le nombre d'actes sexuels réalisés par un couple par semaine respectivement). Le code couleur des liens est le suivant : moins de 2 (en noir), entre 2 et 4 (pourpre), entre 4 et 6 (vert), entre 6 et 8 (jaune) et plus de 8 (cyan) actes sexuels par semaine. **D**) Médiane de la proportion d'individus infectés en fonction du temps suivant 1000 répétitions indépendantes d'un modèle épidémiologique de type SI appliqué aux réseaux A (en noir), B (en rouge) et C (en bleu). Les méthodes et les choix de paramètres sont détaillés en annexe A (voir en particulier la figure 2).

Dans [Kamp *et al.*, 2013], nous confirmons analytiquement les résultats numériques publiés dans [Moslonka-Lefebvre *et al.*, 2012a] et nous proposons de nouveaux outils mathématiques pour rendre

compte avec une bonne approximation d'épidémies SIR pour tout réseau pondéré. En bref, nous adoptons le même formalisme que les modèles dits « de configuration », ce qui implique qu'un réseau pondéré est uniquement décrit par sa distribution de probabilité conjointe P_{kl} en degré k (par exemple le nombre de partenaires sexuels par individu) et en poids l (par exemple le nombre d'actes sexuels par individu par unité de temps). Une illustration de ce formalisme est donnée à la figure 6. Nous pouvons ensuite décrire des épidémies de type SIR sur l'ensemble de leurs trajectoires temporelles en résolvant un système d'équations différentielles partielles de fonctions génératrices de probabilités dérivées de P_{kl}. Cette nouvelle approche est donc applicable aux données issues de sondages, puisque ces dernières nous donnent accès à P_{kl}. Ce type de données présente l'avantage d'être aisément accessibles et plus représentatives en comparaison des données de réseaux empiriques de contacts sexuels obtenues par des études de traçage des contacts (section 2.3). En outre, cette approche nous permet de calculer des estimations analytiques du nombre de reproduction de base (R_0) et du taux initial de croissance épidémique (r_0). En particulier, nous trouvons qu'une approximation de R_0 est données par :

$$R_0 \approx \frac{\pi \langle l/k \rangle}{\pi \langle l/k \rangle + \gamma} \left(\frac{\langle kl \rangle}{\langle l \rangle} - 1 \right) , \tag{1}$$

où π est la probabilité de transmission par événement d'interaction, $\langle l/k \rangle$ est le nombre moyen d'événements d'interaction par contact par unité de temps, γ le taux de rétablissement, $\langle l \rangle$ le nombre moyen d'événements d'interaction par agent par unité de temps et $\langle kl \rangle$ est le second moment de la probabilité de distribution conjointe P_{kl}. En remarquant que le taux de transmission par contact ω (see subsection 1.2.1) est donné par $\omega = \pi \langle l/k \rangle$, nous pouvons reformuler l'expression de R_0 :

$$R_0 \approx \frac{\omega}{\omega + \gamma} \left(\frac{\langle kl \rangle}{\langle l \rangle} - 1 \right) . \tag{2}$$

Si les poids par agent sont constants ($l \propto 1$ in (2)) – une hypothèse corroborée en première approximation par les données de sondages (voir Fig. 2.2 en page 58) – R_0 est donné par :

$$R_0 \approx \frac{\omega}{\omega + \gamma} \left(\langle k \rangle - 1 \right) \leq \frac{\omega}{\omega + \gamma} \left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right) , \tag{3}$$

où le membre de droite est la valeur de R_0 pour laquelle le réseau n'est pas pondéré ($l \propto k$ in (2), ce qui est identique à l'expression de référence de la littérature (voir équation 2.2 en page 68). L'inégalité entre les deux membres de l'inéquation est cohérente puisque $\langle k \rangle^2 \leq \langle k^2 \rangle$, avec égalité seulement dans le cas où la variance de k, donnée par $\langle k^2 \rangle - \langle k \rangle^2$, est négligeable, c'est-à-dire lorsque les réseaux sont strictement réguliers. La relation (3) confirme analytiquement les résultats rapportés dans [Moslonka-Lefebvre *et al.*, 2012a], et démontre l'importance de considérer de concert les contraintes de creux et de pondération.

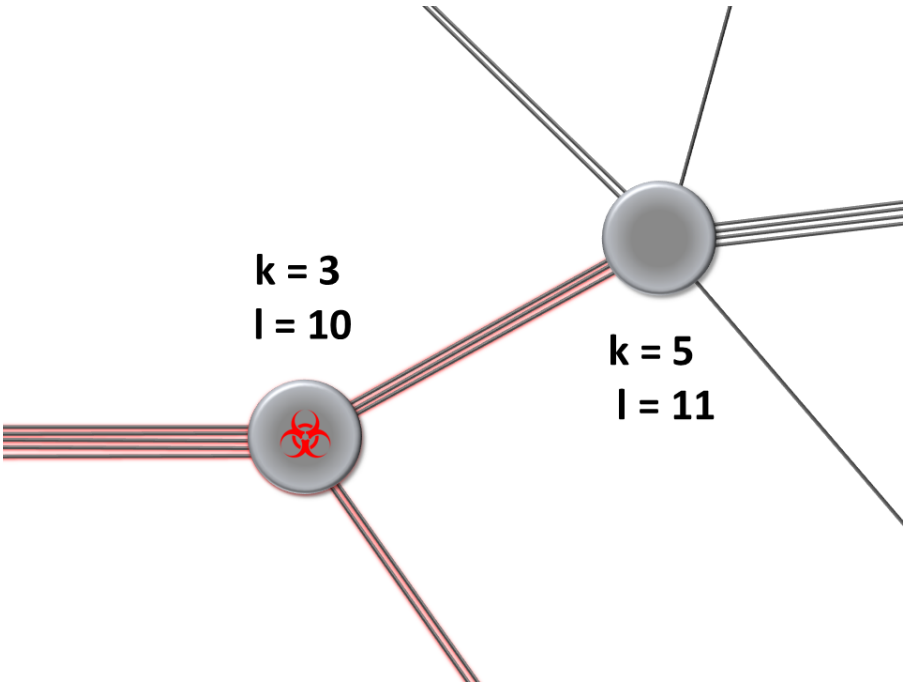


FIGURE 6 – **Pondération du réseau suivant un modèle de configuration.** Le nombre de partenaires sexuels d'un individu et le nombre d'actes sexuels par unité de temps qu'il ou elle assigne à chacun / chacune de ses partenaires constituent deux facteurs de risque épidémique. Dans le modèle décrit dans [Kamp *et al.*, 2013], chaque individu dispose de l événements d'interaction par unité de temps (actes sexuels par semaine par exemple) qu'il ou elle distribue selon ses k contacts (nombre de partenaires sexuels par exemple). À l'échelle de la structure de contact épidémiologique, ces variables sont décrites par la distribution de probabilité conjointe P_{kl} d'identifier un individu avec k contacts et l événements d'interaction par unité de temps.

Pris ensemble, et à rebours de nombre de résultats classiques de la littérature [par exemple Albert *et al.*, 2000] [Moslonka-Lefebvre *et al.*, 2012a; Kamp *et al.*, 2013] suggèrent que les mesures de prévention épidémiques concentrant les ressources sur les individus les plus connectés du réseau n'est pas nécessairement la stratégie optimale pour la réduire la propagation des IST. Par contraste, des campagnes de prévention « grand public » (par exemple des spots TV pour promouvoir l'usage du préservatif) peuvent se révéler d'un rapport efficacité – coûts plus favorables pour ce type d'infections.

Synthèse du chapitre 4 – la direction des échanges influence la transmission ; le cas de la propagation de *P. ramorum* au sein des marchés de plantes

Le quatrième chapitre (voir page 99) souligne l'importance épidémiologique de la contrainte de direction et résume un article de recherche publié dans le journal *Mathematical Biosciences* [Moslonka-Lefebvre *et al.*, 2012b]. Les résultats détaillés sont disponibles à l'annexe C mais ne sont pas présentés dans le corps de texte puisque ce travail n'a pas été initié dans le cadre du présent projet de thèse.

Les réseaux dirigés constituent un outil plus approprié que les réseaux non-dirigés pour décrire les flux commerciaux. Par exemple, le commerce des plantes d'ornement est mieux décrit par des réseaux dirigés puisque les connexions des producteurs aux intermédiaires et aux utilisateurs finaux sont bien plus probables que dans la direction opposée [Pautasso *et al.*, 2010]. La plupart des études se sont

concentrées sur les réseaux non-dirigés, avec une connexion symétrique des noeuds. Le recours aux réseaux dirigés reste limité puisque les matrices d'adjacence qui les décrivent sont asymétriques dans le cas général, ce qui rend leur étude plus ardue [Newman *et al.*, 2001; Meyers *et al.*, 2006; Park & Kim, 2006]. Un modèle épidémiologique adapté aux réseaux dirigés apporterait une représentation plus réaliste des épidémies se propageant sur des systèmes asymétriques tels que les marchés. Dans [Moslonka-Lefebvre *et al.*, 2012b], nous introduisons et explorons un modèle décrivant la propagation d'un processus infectieux sur des réseaux dirigés et non-pondérés. Nous qualifions ce nouveau cadre théorique de *modèle SIS sur un continuum* (SIS_c). Le modèle SIS_c décrit l'état infectieux de chaque noeud d'un réseau dirigé suivant un continuum d'états entre l'état totalement sensible et l'état totalement infecté. Ce cas de figure se révèle réaliste pour des épidémies affectant le commerce des plantes, puisque les pépinières végétales et les points de vente peuvent être totalement indemne d'un pathogène donné, totalement infestées par ce même pathogène, où être infectées à un niveau intermédiaire entre ces deux extrêmes [Pautasso & Jeger, 2008]. En utilisant les mêmes notations que pour la partie I, nous présentons le modèle SIS_c en temps continu. Les agents infectés transmettent l'infection au taux de transmission par contact ω et le stock infectieux est retiré et éliminé au taux γ (voir section 2.2.2. de l'annexe C). Soit $Y_i(t)$ la variable représentant le nombre total de produits infectés au sein de l'agent i au temps t. Le modèle SIS_c décrit l'évolution temporelle de Y_i suivant un système d'équations différentielles ordinaires :

$$\frac{dY_i}{dt} = \omega \sum_{j \neq i} w_{ji} Y_j - \gamma Y_i \ , \tag{4}$$

où, $w_{ji} = 1$ si j est connecté à i, et $w_{ji} = 0$ sinon. On peut noter que $w_{ji} \neq w_{ij}$ dans le cas général (voir aussi la sous-section 2.3.1). Le modèle SIS_c peut s'appliquer à de nombreuses entités propagées, telles que l'information, les maladies et les rumeurs véhiculées entre les individus ou les organisations. Dans le cas présent, nous nous focalisons sur la propagation des maladies dans les réseaux commerciaux horticoles telles que l'épidémie de *Phytophthora ramorum* [Jeger *et al.*, 2007; Moslonka-Lefebvre *et al.*, 2011]. Dans un tel système, un processus épidémique de type SIS se révèle particulièrement adapté puisqu'une immunisation est impossible : tant qu'un producteurs de plantes continue d'échanger des végétaux, il s'expose au risque d'être infecté, y compris après une éradication complète du stock infecté et l'adoption de pratiques visant à prévenir des infections ultérieures.

Nous commençons par évaluer le risque épidémique global associé à un réseau dirigé. Comme pour un réseau pondéré (chapitre 3), un réseau dirigé peut être caractérisé par sa distribution de probabilité conjointe $P_{k^{in}k^{out}}$, où k_i^{in} et k_i^{out} représentent respectivement le nombre de partenaires entrants et sortants (degrés entrant et sortant) de l'agent i. Le nombre de reproduction de base R_0 [2] associé au modèle SIS_c peut être estimé à partir de l'hypothèse usuelle de champ moyen qui négligent

2. Pour le cas où les agents sont soit entièrement susceptibles, soit entièrement infectés, nous pouvons postuler, par analogie avec (3.4a)–(3.4b), que R_0 est donné par $R_0 \approx \frac{\omega}{\omega + \gamma} \left(\frac{\langle k^{in} k^{out} \rangle}{\langle k \rangle} - 1 \right)$.

les corrélations du type degré – degré entre pairs d'agents [Restrepo *et al.*, 2007] (see section 3.1 of appendix C for further details) :

$$R_0 \approx \frac{\omega}{\gamma} \frac{\left\langle k^{in} k^{out} \right\rangle}{\langle k \rangle} \ , \tag{5}$$

où $\langle k \rangle$ est le degré moyen (les moyennes des degrés entrants et sortants doivent être identiques), et $\left\langle k^{in} k^{out} \right\rangle$ est le second moment de la distribution de probabilité conjointe $P_{k^{in}k^{out}}$. L'équation (5) montre que les corrélations entre les degrés entrants et sortants ($\left\langle k^{in} k^{out} \right\rangle$) font croître le R_0. Dans le cas extrême où les degrés entrants et sortants de chaque agent i sont identiques ($k_i^{in} = k_i^{out} = k_i$ dans (5)), R_0 s'exprime :

$$R_0 \approx \frac{\omega}{\gamma} \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} \ , \tag{6}$$

ce qui est identique à la relation (3.4a), c'est-à-dire la formule de R_0 pour des raisons non-dirigés du type « fully mixed ».

Par contraste, dans le cas intermédiaire des réseaux dirigés non corrélés (lorsque $\left\langle k^{in} k^{out} \right\rangle = \left\langle k^{in} \right\rangle \left\langle k^{out} \right\rangle = \langle k \rangle^2$ dans (5)), nous remarquons :

$$R_0 \approx \frac{\omega}{\gamma} \langle k \rangle \leq \frac{\omega}{\gamma} \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} \ , \tag{7}$$

ce qui implique un risque d'invasion plus faible pour les réseaux dirigés non corrélés [3] en comparaison des réseaux dirigés corrélés positivement. Pour un nombre total fixé de liens dirigés, nos résultats suggèrent que les réseaux dirigés sont associés à un risque d'invasion plus faible que les réseaux non dirigés (où par construction les liens dirigés sont strictement symétriques pour chaque pair de noeuds, ce qui induit une corrélation maximale entre degrés entrants et sortants à l'échelle des agents). Ces prédictions théoriques sont corroborées par les résultats numériques rapportés dans [Moslonka-Lefebvre *et al.*, 2009].

Pour un réseau dirigé donné, le modèle SIS_c apporte également des éclairages d'intérêts pour hiérarchiser les risques épidémiques à l'échelle des agents exposés. Suivant Pautasso *et al.* [2010], nous qualifions les agents de producteurs, intermédiaires et détaillants en fonction de la différence relative entre leurs degrés entrants et sortants. Plus généralement, la position hiérarchique d'un agent i au sein d'un réseau dirigé, c'est-à-dire la position de i dans la chaîne d'approvisionnement, est quantifiée par un index hiérarchique Δ_i comme suit :

$$\Delta_i = \frac{k_i^{in} - k_i^{out}}{k_i^{in} + k_i^{out}} \ , \tag{8}$$

où par construction $-1 \leq \Delta_i \leq -1$. Les producteurs correspondent à des exportateurs nets (cas des Δ_i négatifs), les détaillants à des importateurs nets (Δ_i positifs), et les intermédiaires à des agents sans prépondérance marquée pour l'export ou l'import (valeurs de Δ_i centrées sur zero), avec un rôle

3. Le risque serait encore plus faible pour les réseaux présentant des corrélations négatives entre les degrés entrants et sortants à l'échelle des agents.

de relai entre les deux catégories précédentes. En d'autres termes, nous faisons l'hypothèse que les directions relatives des liens reflètent la spécialisation des agents dans la chaîne d'approvisionnement (Fig. 3E). Il est à noter que nous avons adapté la définition (24) dans l'annexe C pour faciliter la comparaison avec la notion de polarité des flux définie dans le chapitre 5. La relation 5 suggère qu'une stratégie de contrôle efficace consisterait à réduire les niveaux de corrélations élevés entre liens entrants et sortants en ciblant préférentiellement les mesures de police sanitaire sur les intermédiaires. Cette prédiction théorique a été confirmée numériquement dans [Pautasso *et al.*, 2010] et dans la section 3.2 de l'annexe C.

Synthèse du chapitre 5 – combiner pondération et direction des échanges pour évaluer conjointement les risques économiques et sanitaires ; application aux mouvements de bétail.

Le cinquième chapitre (voir page 103) présente un article intitulé *Market analyses of livestock trade networks and prevention of joint economic and epidemiological risks*, en cours de resoumission – proposée par le comité éditorial – pour *Journal of the Royal Society Interface*. En bref, nous décrivons et analysons les mouvements de bétail par des réseaux pondérés et dirigés. En effet, le nombre total d'animaux transportés par pair d'agents sur une période de temps donnée est associé à une variabilité inter-paires importante [par exemple Vernon, 2011], et les animaux sont majoritairement transportés de sites vendeurs vers des sites acheteurs [par exemple Lal Dutta *et al.*, 2014]. Nous proposons une nouvelle classification des agents qui rend compte conjointement du poids des sites (contrainte de pondération, voir chapitre 3) et de la direction d'ensemble des échanges (contrainte de direction, voir chapitre 4). Cette classification nous permet d'identifier des stratégies présentant des rapports efficacité – coût avantageux pour prévenir les épidémies tout en limitant les effets adverses sur le commerce qui sous-tend la propagation de ces mêmes épidémies.

Le commerce du bétail joue un rôle clé pour les secteurs agricoles et agro-alimentaires, mais constitue également une voie de transmission des maladies infectieuses [Vernon, 2011]. L'architecture des mouvements d'animaux est extrêmement riche et peut être décrite par des réseaux de contacts potentiellement infectieuses [Vernon & Keeling, 2009]. En permettant le calcul de déterminants structurels du risque épidémique, de tels modèles de réseaux se révèlent d'une grande valeur pour identifier les sites agricoles susceptibles de jouer un rôle prépondérant dans la propagation des épidémies [Bajardi *et al.*, 2012]. Des exemples de facteurs de risque incluent le nombre total de partenaires commerciaux d'un agent [le degré total, voir par exemple Rautureau *et al.*, 2011] et le nombre de sites de regroupements d'animaux susceptibles d'êtres contaminés en suivant les liens temporellement compatibles à partir d'un site agricole donné [la chaîne infectieuse sortante, voir par exemple Noremark *et al.*, 2011]. Dans le cas général, toute mesure de centralité des réseaux appliquée à un site donnée peut être mobilisée comme un indicateur du risque de contagion et / ou de transmission. Pour prévenir et atténuer les

épidémies, il apparaît naturel de recommander une surveillance préférentielle des sites les plus centraux des marchés de bétail. Toutefois, concentrer les mesures de prévention sur les acteurs commerciaux les plus importants peut se révéler coûteux, ce qui peut rendre ces stratégies inapplicables en pratique. Les plus grands sites agricoles impliquent en effet de très nombreux animaux à diagnostiquer, ce qui induit des coûts plus élevés pour le gestionnaire de risque, et leur déstabilisation, par exemple par excès de précaution, peut induire des déstabilisations économiques majeures [Paton *et al.*, 2009]. L'évaluation conjointe des risques économiques et épidémiologiques constitue une lacune d'importance dans nos connaissances.

Dans le présent travail, nous cherchons à identifier des stratégies de prévention épidémique efficaces tout en en minimisant les distorsions de marché et les coûts d'intervention. Pour ce faire, nous introduisons et calculons des indicateurs de risques économiques et épidémiologiques. Nous développons une catégorisation adaptée aux marchés pour classer les agents suivant des critères économiques. Ces critères diffèrent des critères de centralité structurale traditionnellement utilisés en théorie des réseaux. Dans notre cadre d'étude, chaque site i correspond à un agent économique caractérisé par deux quantités, à savoir la *polarité des flux*, notée fp_i, et *la part des flux*, notée fs_i, qui sont définies de la façon suivante :

$$\begin{aligned} \mathrm{fp}_i &= \frac{\phi_i^{\mathrm{in}} - \phi_i^{\mathrm{out}}}{\phi_i^{\mathrm{in}} + \phi_i^{\mathrm{out}}} \ , \\ \mathrm{fs}_i &= \frac{\phi_i^{\mathrm{in}} + \phi_i^{\mathrm{out}}}{\sum_j (\phi_j^{\mathrm{in}} + \phi_j^{\mathrm{out}})} \ , \end{aligned} \tag{9}$$

où ϕ_i^{in} et ϕ_i^{out} représentent respectivement les volumes commerciaux achetés et vendus par l'agent i sur une période de temps donnée. La polarité des flux quantifie la *position dans la chaîne d'approvisionnement* (Fig. 7A,C), et constitue à ce titre une extension de l'index hiérarchique (4.5) aux réseaux dirigés avec des poids hétérogènes. La part des flux est un indicateur classique de *leadership de marché* pour évaluer les parts de marchés détenues par un agent donné via des études marketing [Kotler & Armstrong, 2013]. La part des flux peut s'interpréter comme le poids relatif d'un agent dans un marché donné (Fig. 7B,D). Nous définissons des *catégories de marché* en combinant la polarité des flux et la part des flux et étudions les réseaux formés par ces catégories (Fig. 8).

En analysant deux jeux de données traçant les mouvements de bovins et porcins en France, nous montrons que les catégories de marché correspondent à des groupes contrastés d'agents économiques (Fig. 7C,D). Ces catégories sont relativement stables dans le temps et peuvent se révéler utiles pour évaluer conjointement les risques économiques et épidémiologiques. Bien que les profils de risques par catégorie marchande soient similaires pour les deux marchés, nos analyses suggèrent que le commerce des bovins présente un niveau de risque plus élevé que le commerce des porcins (voir par exemple Fig. 8), ce qui est bien en phase avec les analyses d'une étude antérieure [Rautureau *et al.*, 2012].

Nous comparons enfin des stratégies contrastées de prévention épidémique. Des stratégies statiques et dynamiques sont comparées, aussi bien sur la base de données historiques que de données en temps réel. Les stratégies de prévention que nous décrivons sont mises en oeuvre par une autorité publique qui cherche à protéger préférentiellement des agents appartenant à certaines catégories de marchés à risque. Toutefois, ce qui fait qu'une stratégie est jugée préférable dépend de l'acteur qui conduit l'évaluation. En effet, la mise en oeuvre d'une mesure de protection est coûteuse et peut occasionner des distorsions de marché. Nous confrontons donc deux points de vue lorsque nous évaluons les stratégies de prévention : celui de l'autorité publique et celui du marché. Pour ce qui concerne le point de vue exclusif de l'autorité publique, si nous émettons l'hypothèse que les coûts d'intervention sont proportionnels au nombre d'agents économiques à cibler, la meilleure stratégie consiste à protéger préférentiellement les agents avec les parts de flux les plus élevées, ce qui constitue une recommandation usuelle de la littérature pour des systèmes similaires [Albert *et al.*, 2000]. Par contraste, quand le point de vue du marché est également pris en compte, ou bien quand les coûts de prévention sont plutôt proportionnels au volume total échangé comme c'est probablement le cas pour les campagnes de vaccination, il apparaît préférable de protéger en priorité les agents mineurs sur le plan économique, tout particulièrement pour le marché bovin français.

Prise dans son ensemble, notre étude suggère que de multiples points de vue devraient être confrontés lors de l'évaluation de stratégies de prévention. En effet, les échanges, les contaminations infectieuses imputables à ces échanges et la mise en oeuvre de politique de contrôle induisent des impacts et des décisions contrastées suivant les agents considérés. Ce résultat présente une portée générale pour des études épidémiologiques et écologiques visant à optimiser des interventions en vue de maintenir des (éco-)systèmes sains et diversifiés.

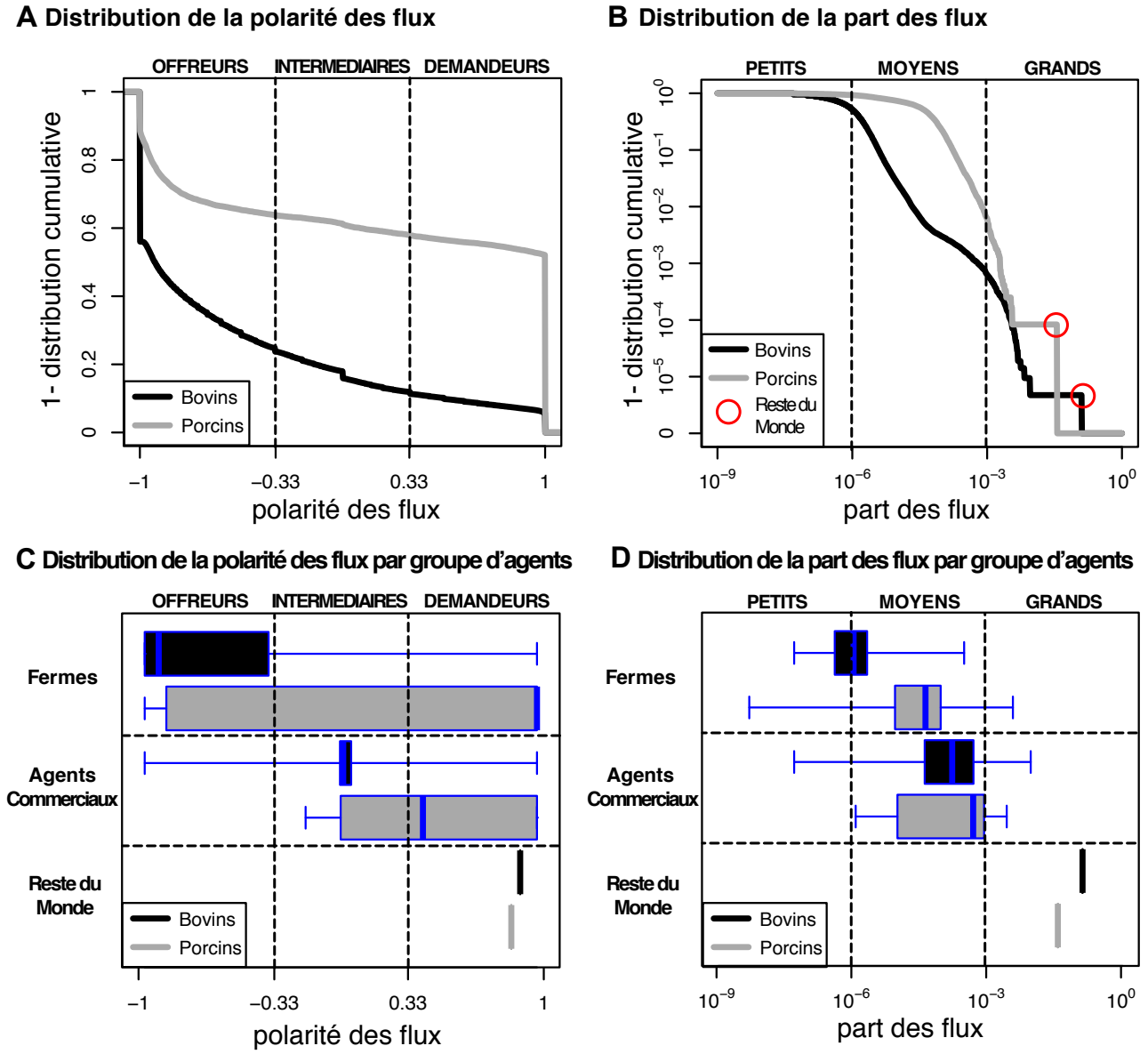


FIGURE 7 – **Catégorisation des agents économiques : polarité des flux et part des flux. A)-B)** Distributions de la polarité des flux et de la part des flux dans les marchés bovins et porcins français. La polarité des flux et la part des flux d'un agent économique (5.1) quantifient respectivement sa *position dans la chaîne d'approvisionnement* et son *leadership de marché*. Sur un plan qualitatif, la polarité des flux peut être utilisée pour définir des *offreurs (O)*, *intermédiaires (I)* et *demandeurs (D)*. Ces catégories d'agents correspondant respectivement à des valeurs négatives, symétriquement distribuées autour de zéro, et positives de polarités des flux. De manière analogue, la part des flux peut être utilisée pour définir des *petits (P)* (valeurs faibles), *moyens (M)* (valeurs intermédiaires) and *grands (G)* acteurs (valeurs élevées). **C)-D)** Distributions de la polarité des flux et de la part des flux par groupe d'agents dans les marchés bovins et porcins. Les groupes correspondent à des ensembles de *fermes* ou à des *agents commerciaux* français, ou bien à une entité unique agrégeant l'ensemble des agents situés en dehors de la France, c'est à dire le *reste du monde*. Les polarités des flux et les parts des flux sont calculées en 2009 pour les bovins et en 2010 pour les porcins.

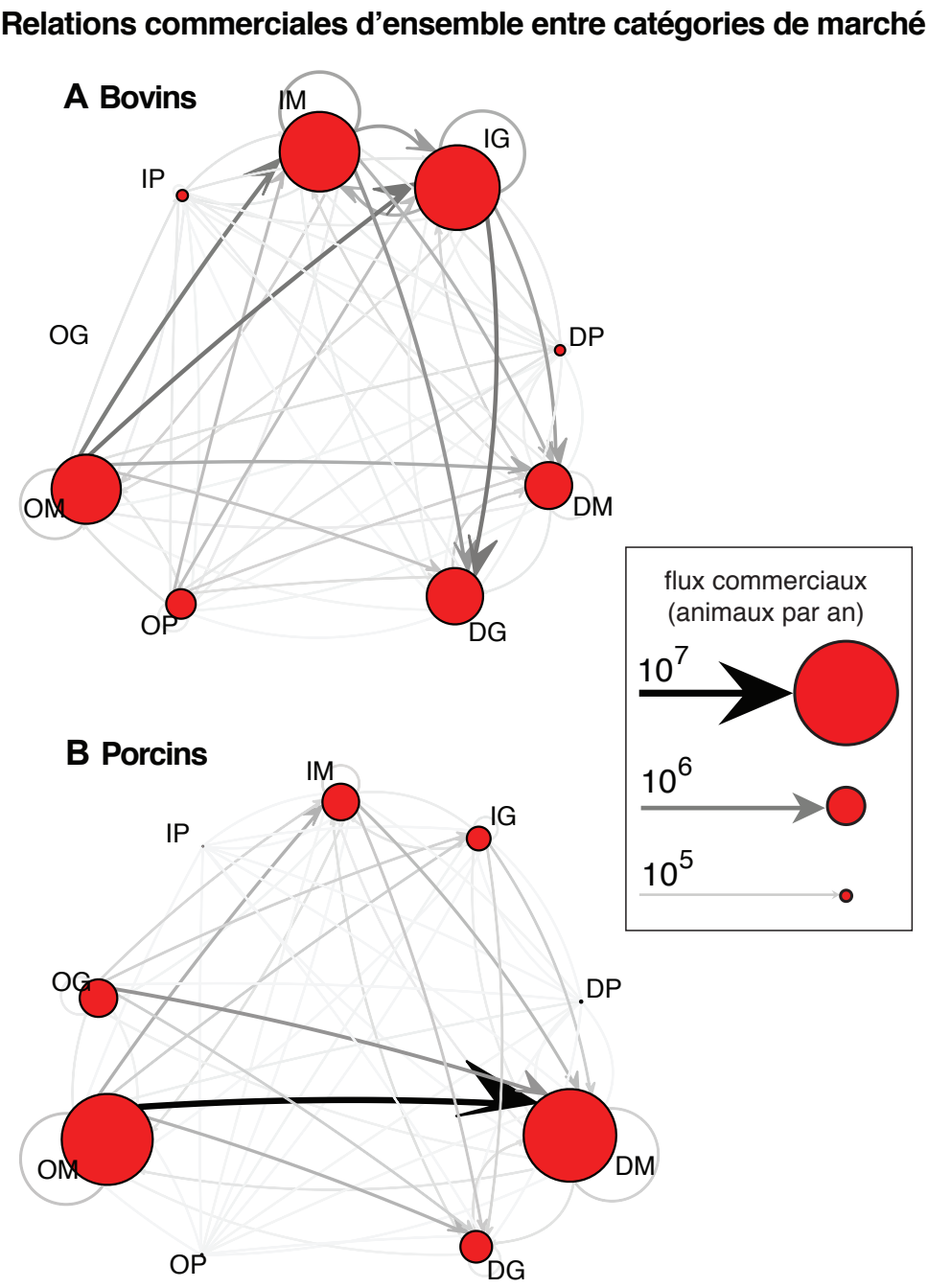


FIGURE 8 – **Réseaux commerciaux représentant les marchés d'échanges de bovins (A) et porcins (B) français.** Les données sont agrégées annuellement (2009 pour les bovins et 2010 pour les porcins). Chaque noeud regroupe l'ensemble des agents correspondant à chaque catégorie de marché. Ces catégories de marchés sont ici discrétisées suivant une grille en 3-par-3 reprenant les notations et les valeurs seuils en polarité des flux et en part des flux de la Fig. 7. La taille de chaque noeud est proportionnelle au flux annuel obtenu par sommation sur l'ensemble des agents dans chaque catégorie agrégée. Les épaisseurs et niveaux de gris des liens dirigés (flèches) sont proportionnels aux flux annuels entre chaque paire de catégories agrégées. Ces flux sont obtenus par sommation sur l'ensemble des agents de catégories aggégées impliquées dans des mouvements entrants ou sortants d'animaux.

Présentation de la partie III – des modèles épidémiologiques intégrant comportements adaptatifs et interactions contraintes

Les contraintes d'interactions auxquelles font face les agents exercent une influence fondamentale sur les dynamiques épidémiques (partie II). Ces contraintes émanent du fait que les échanges sont coûteux et que les ressources sont limitées, et impliquent qu'un agent ne peut interagir en permanence et uniformément avec chacun des autres agents d'un système. Toutefois les contraintes n'expliquent pas comment les échanges se produisent en premier lieu. Par ailleurs, tout choc déstabilisant les échanges est susceptible de modifier les interactions entre agents suivant des mécanismes adaptatifs.

Pour faciliter l'étude des interactions contraintes et de processus infectieux sus-jacents, nous avons émis l'hypothèse que les agents étaient passifs. Clairement, cette hypothèse n'est ni en adéquation avec les données disponibles, ni en accord avec le bon sens (section 2.2). En réalité, les échanges sont motivés par de nombreuses raisons, et les épidémies imputables aux échanges induisent non seulement des dommages biologiques et économiques aux agents contaminés, mais peuvent également altérer les échanges véhiculant des infections à travers des réponses comportementales complexes. À titre d'exemple, lorsque les agents font face à une épidémie, ils peuvent prendre des mesures de protection individuelles et collectives. Ainsi, certains éleveurs peuvent spontanément décider de vacciner leurs troupeaux à titre préventif. Si la couverture vaccinale résultant d'un tel processus se révèle insuffisante, les autorités publiques nationales et internationales peuvent mettre en oeuvre des incitations ou une réglementation pour accroître le nombre total de troupeaux vaccinés [Klepac *et al.*, 2013]. Les modèles épidémiologiques prennent de plus en plus en compte le comportement d'aversion au risque en réponse aux émergences de pathogènes [Funk *et al.*, 2010]. Toutefois les comportements des agents ne sauraient être réduits à l'aversion aux risques du fait d'interdépendances complexes impliquant des processus adaptatifs de nature sociale ou économique. De telles adaptations sont observées par exemple pour les marchés véhiculant des épidémies. Lorsque le marché est déstabilisé par une émergence majeure – donc susceptible d'engendrer une crise économique – certains agents peuvent être tentés de prendre des risques supplémentaires pour générer davantage de profits ou bien pour compenser les pertes affectant leurs partenaires commerciaux en réponse à des infections et des mesures de police sanitaire.

Dans la troisième partie (page 129), nous cherchons à éclairer la dynamique conjointe des échanges et des épidémies. Nous abandonnons donc l'hypothèse restrictive d'agents passifs pour prendre en compte des situations où les épidémies peuvent également déstabiliser les réseaux d'échange qui véhiculent les infections en modulant les mécanismes adaptatifs qui gouvernent ces échanges. Plus précisément, nous nous intéressons au cas des comportements adaptatifs complexes rencontrés dans les marchés et que nous intégrons dans des modèles couplés du type « marché - épidémie » (ME). Nous considérons également le cas particulier des comportements adaptatifs de marché sans épidémie pour déterminer leurs influences sur la dynamique des échanges. Les contraintes d'interaction sont dorénavant subsumées

sous les comportements adaptatifs, et les comportements sont décrits par des processus dynamiques résultant des décisions commerciales des agents. Il s'en suit que nous pouvons explorer les implications économiques et épidémiologiques d'une contrainte particulière, à savoir la contrainte de friction ou friction commerciale (Fig. 3F). Cette friction engendre un compromis entre fréquence et intensité des transactions marchandes. En prenant pour base le modèle SIRS à compartiments (voir partie I et section 1.1), nous commençons par concevoir et explorer un modèle ME à l'échelle macroéconomique avec une application aux mouvements de bétail en France (chapitre 6). Le modèle ME macroscopique suggère que la friction exerce une influence déterminante sur la dynamique conjointe du commerce et de l'infection. Pour confirmer ces premiers résultats, nous étendons le cadre macroscopique à l'échelle microéconomique. Nous pouvons ainsi rendre compte de réseaux dynamiques de contacts potentiellement infectieux entre agents économiques (chapitre 7). Grâce à la richesse des données disponibles concernant le commerce de bovins, nous pouvons clarifier la signification de certains paramètres et variables d'état des deux modèles ME. Nous pouvons également étudier l'influence des hétérogénéités élémentaires en friction, ainsi que les impacts des désordres économiques globaux sur la dynamique épidémies et des échanges. Nous soulignons que nos deux modèles ME, qui décrivent des processus marchands se déroulant à des niveaux macroscopique et microscopique, reposent sur plusieurs idées phares rapportées dans la partie II, en particulier les mécanismes socio-économiques présentés dans l'annexe A, ainsi que les catégories de marché définies dans le chapitre 5.

Synthèse du chapitre 6 – un modèle mécaniste pour explorer les interactions dynamiques entre échanges et épidémies dans des marchés avec des agents homogènes.

Le sixième chapitre (voir page 131) présente un article en révision mineure pour *Journal of Theoretical Biology* et intitulé *Epidemics in markets with trade friction and imperfect transactions* [Moslonka-Lefebvre *et al.*, 2013].

Bien qu'il soit admis que le commerce peut véhiculer des épidémies ainsi que d'autres types d'invasions biologiques [Noremark *et al.*, 2011], l'interaction de ces processus avec la dynamique des marchés reste mal comprise. Les marchés peuvent propager des maladies infectieuses entre agents (par exemple des fermes) à travers des échanges de « biens » contaminés (par exemple des animaux). Réciproquement, la dynamique de marché est influencée par des comportements adaptatifs complexes des agents. Ces comportements peuvent faire suite à des mesures collectives de police sanitaire ou à une prise de conscience individuelle de la menace épidémique par les agents [Funk *et al.*, 2009]. Les modèles épidémiologiques prennent de plus en plus en compte le phénomène de réduction des contacts potentiellement infectieux, par exemple le comportement d'aversion au risque en réponse à l'émergence épidémique d'un pathogène [Funk *et al.*, 2010]. Néanmoins, les réponses de la dynamique de marché ne sauraient se réduire à la simple aversion au risque, puisque les ressorts des comportements marchands sont gouvernés par des motivations complexes (voir section 2.2).

Ici, nous cherchons à modéliser l'influence globale de la dynamique de marché sur la dynamique des épidémies, et en retour, l'influence de la dynamique infectieuse sur la dynamique de marché. Nous développons donc un modèle couplé du type « marché – épidémie » (ME) à l'échelle macroéconomique pour modéliser la dynamique des échanges marchands et des épidémies véhiculées par ces échanges (Fig. 9A). Pour représenter ce processus, nous relions un modèle décrivant un marché, c'est-à-dire un système économique, et un modèle décrivant un processus épidémique. La dynamique de chaque modèle existe en soi et peut-être étudiée indépendamment. Les épidémies peuvent se produire au sein de populations d'hôtes distincts des marchés. Les marchés fonctionnent quant à eux souvent sans propager des émergences épidémiques à travers les liens commerciaux qu'ils impliquent. Néanmoins, en construisons un système qui relie les dynamiques de ces deux sous-systèmes, nous pouvons étudier leurs interactions. Nous étudions le modèle ME en adoptant une approche ascendante : nous commençons par analyser un nouveau modèle de marché qui prend explicitement en compte les contacts entre agents en l'absence d'épidémie (Fig. 9B), puis nous explorons un modèle ME intégratif où épidémies et commerce s'influencent mutuellement (Fig. 9A-C).

En supposant initialement une absence d'épidémie, nous développons un modèle qualifié de *frictional-trade market (FTM) model* (modèle de marché d'échanges avec friction en français). Le modèle FTM décrit la dynamique de marché où un seul type de biens est échangé contre de l'argent, avec un prix d'échange variable et unique. Par contraste avec des modèles économiques usuels [Mas-Colell *et al.*, 1995], le modèle FTM prend explicitement en compte les *événements d'interaction transitoire* entre agents économiques. C'est ce type d'événement qui va jouer un rôle déterminant dans le détermination de la structure de contact épidémiologique qui sous-tend le modèle ME analysé dans un second temps. Nous considérons un marché composé de N_S offreurs, N_D demandeurs, ainsi que $N_{S\cap D}$ intermédiaires, c'est-à-dire des agents simultanément offreurs et demandeurs. Nous supposons que les échanges commerciaux se produisent des offreurs en direction des demandeurs (en adéquation avec le chapitre 4) par le biais de *transactions*. Une *transaction* est un événement de livraison instantané d'un offreur vers un demandeur. À l'échelle du marché, les transactions se produisent à un certain taux Θ, *le taux de transaction*. Les transactions sont d'une importance extrême parce qu'elles forment une structure de contact dynamique entre les agents cherchant à échanger. Et c'est cette structure de contact qui va permettre la transmission des épidémies imputables aux échanges commerciaux. Lors d'une transaction, un *lot moyen* de q biens est échangé. Il s'en suit que le *flux marchand* Φ, qui représente le nombre de produits échangés des offreurs vers les demandeurs durant une période de temps divisée par cette même période de temps. Φ est donné par le produit $\Theta \times q$. Pour une valeur de Φ donnée, les quantités Θ et q sont donc reliées par un compromis. A titre d'exemple, les marchés d'échange bovins et porcins français ont à peu près le même Φ, mais en moyenne, q est bien plus élevé chez les bovins que chez les porcins (voir Fig. 2.3). Dans le modèle FTM, nous contrôlons le compromis entre

Θ et q par un paramètre appelé la *friction marchande*, noté κ. Un concept analogue de friction est employé de longue date en économie pour modéliser les marchés du travail [Pissarides, 2011]. Dans notre approche, la friction κ agrège un ensemble de contraintes telles que la recherche de partenaires commerciaux et la logistique de la livraison. Par construction, le paramètre κ gouverne le compromis entre fréquence et intensité des transactions marchandes : accroître la friction va diminuer le taux de transaction mais augmenter le nombre moyen de biens échangés par transaction (voir la contrainte de friction à la Fig. 3F). Le système d'équations régissant le modèle FTM est rendu clos en spécifiant de façon explicite les processus mécanistes qui sous-tendent les variables Θ and q. Nous supposons que q est déterminé par l'appariement [4] potentiellement imparfait entre les stocks de biens offerts et demandés sur le marché. Quant à Θ, nous supposons qu'il est fonction du prix actuel des biens et de la friction. Nous paramétrons le modèle FTM à partir d'une base de données française traçant les échanges de bovins et porcins.

Nous trouvons, pour ce qui concerne le modèle FTM, que la friction peut accroître de plusieurs ordres de grandeur le temps nécessaire pour que le marché atteigne son état d'équilibre sans nécessairement altérer la valeur à long terme du flux marchand. κ caractérise par conséquent la dynamique propre du marché, et notamment son temps de réponse caractéristique en cas de perturbation ; par exemple en cas d'émergence épidémique.

Pour explorer la manière dont marchés et épidémies s'influencent mutuellement, nous modélisation la propagation et le contrôle des maladies infectieuses au sein de marchés. En pratique, nous couplons le modèle SIRS à compartiments (section 1.1) au modèle FTM. Le modèle résultant de ce couplage est qualifié de modèle *« marché - épidémie »* (ME). De manière à comparer et inscrire notre approche dans la littérature, nous rendons compte dans notre modèle du comportement adaptatif d'aversion au risque (noté RA pour *risk aversion* en anglais). Bien que le modèle ME reste un modèle à compartiments, il présente des réponses qualitativement différentes de son sous-modèle SIRS strictement épidémiologique. Le modèle ME prend en compte deux types de boucles de rétroaction entre marchés et épidémies (Fig. 9A). Suivant la littérature, nous supposons que les épidémies déstabilisent les marchés à travers trois processus : l'aversion au risque ; le retrait des agents commerciaux infectieux ; et le rétablissement des agents retirés après la mise en oeuvre des mesures de police sanitaire. En retour, les marchés influencent les épidémies à travers la dynamique des échanges marchands, échanges qui émanent de la volonté des agents à échanger. Plus précisément, l'infection résulte d'échanges entre offreurs et demandeurs (chapitres 4 et 5). Ces échanges déterminent deux composantes du taux de transmission du processus infectieux imputable aux échanges marchands (voir la sous-section 1.2.1), à savoir le taux de contact ; et la probabilité d'infection par contact. D'un point de vue épidémiologique, une transaction constitue un contact élémentaire de la structure de contact épidémique sous l'hypothèse

4. concept de *matching* en anglais.

que le commerce constitue la seule et unique voie de transmission. Il en découle que le taux de contact est donné par le taux de transaction moyen divisé par le nombre de demandeurs actifs dans le marché, c'est-à-dire par Θ divisé par le nombre de demandeurs qui sont sensibles ou infectés. La probabilité d'infection par contact est une fonction croissante de q, qui représente l'intensité épidémiologique d'une transaction donnée. Il s'en suit que l'influence de la friction marchande sur les épidémies n'est pas évidente à déterminer *a priori*, puisque la friction module la balance entre les variables Θ et q, qui toutes deux jouent positivement sur le taux de transmission.

Nous montrons que lorsque le commerce constitue la voie dominante de transmission, la friction de marché peut constituer un déterminant majeur des épidémies, et cela bien plus que l'aversion au risque. Dans le cas particulier où le commerce constitue l'unique route de transmission, le nombre de reproduction de base du modèle ME (R_0) est nécessairement inférieur à :

$$R_0 \leq \frac{1}{\gamma} \frac{N_{S \cap D} \Phi^*}{\kappa N_D N_S} \ , \tag{10}$$

où γ est le taux de retrait du marché, et Φ^* représente le flux marchand à l'équilibre temporel du système. La relation (10) implique l'existence d'un niveau critique de friction au-dessus duquel les épidémies ne peuvent se produire. Ce résultat suggère que les marchés à frictions élevées ne peuvent véhiculer des épidémies. En particulier, le commerce de porcins ($\kappa \approx 70$) serait bien moins en capacité de maintenir des épidémies que le marché bovins ($\kappa \approx 3$) ; une prédiction par ailleurs corroborée par d'autres études [Rautureau *et al.*, 2012]. En outre, (10) suggère que le délai maximal de retrait d'agents infectés qui permet encore une atténuation des épidémies croît avec la friction (Fig. 10A). Nos résultats ne dépendent pas des fonctions particulières choisies pour le modèle de transmission et sont robustes à la présence de voies de transmission non-marchandes (Fig. 10B). En matière d'appui aux politiques publiques de sécurité sanitaire, nos résultats suggèrent qu'une mesure adéquate pour minimiser les processus contagieux sans pour autant déstabiliser les marchés serait de mettre en oeuvre des incitations pour augmenter la friction. Un accroissement de la friction se traduit en effet par des transactions moins fréquentes et de plus grands volumes, ce réduit la capacité du marché à véhiculer des épidémies sans nécessairement impacter le flux marchand global.

Le modèle ME macroscopique suggère que la friction marchande exerce une influence déterminante sur la dynamique conjointe des échanges et de l'infection. Par souci de simplicité, il est supposé ici que le processus de rencontre entre agents suit un processus de mélange homogène. Un modèle explicite rendant compte des échanges entre agents devrait décrire les transactions à l'échelle élémentaire des pairs d'agents décidant d'échanger des biens contre des liquidités. Une lacune de connaissance que nous cherchons à combler au chapitre 7 consiste à déterminer si des niveaux de frictions réalistes peuvent également atténuer des épidémies dans des modèles de marchés fondés sur des réseaux explicites.

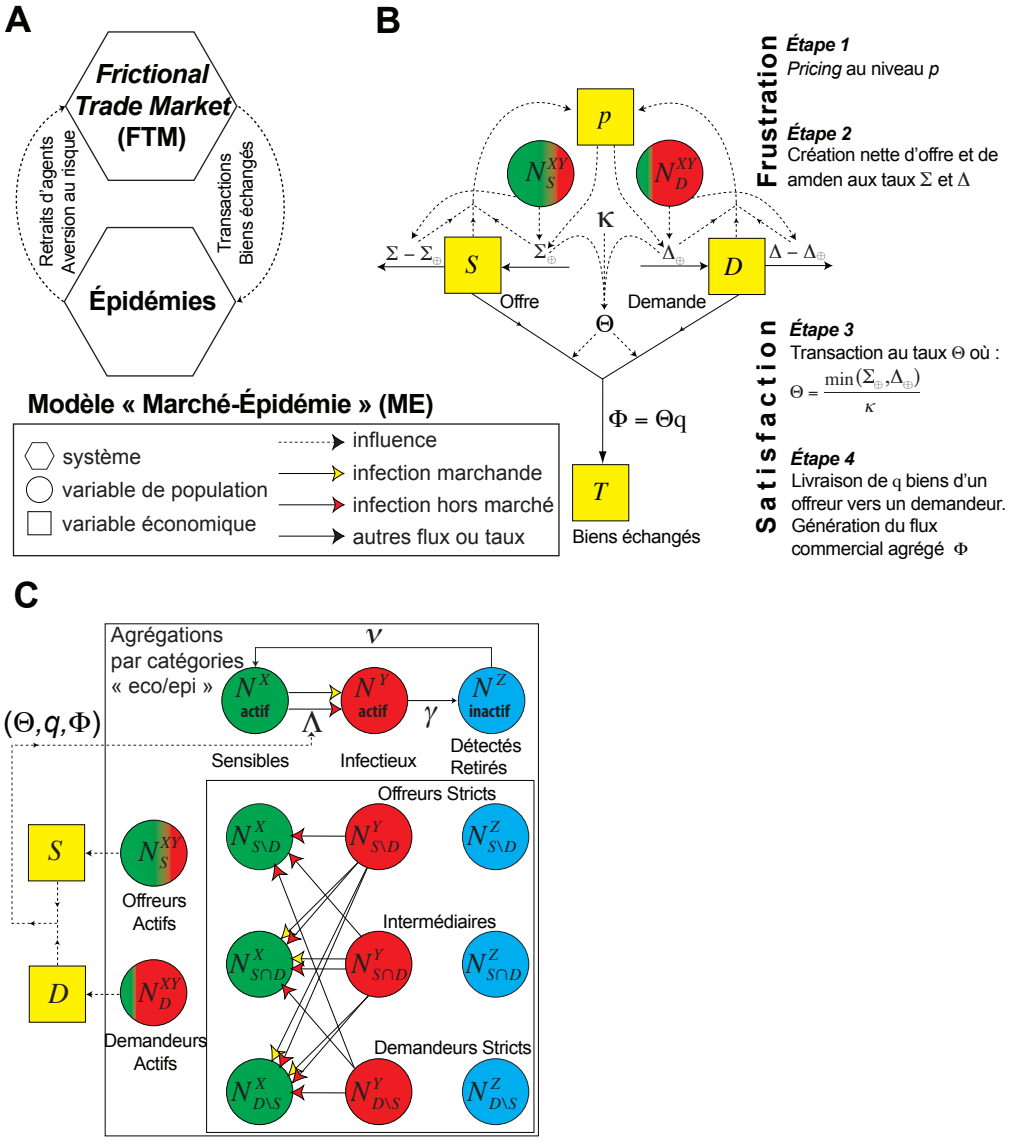


FIGURE 9 – **Cadre général employé pour modéliser le système couplé marché - épidémie. A**) Structure générale et interactions clés entre le sous-système marchand et le le sous-système épidémique. **B**) Composantes du frictional-trade market (FTM) model (modèle de marché d'échanges avec friction en français). **C**) Composantes du modèle « marché - épidémie » (ME). Les flèches jaunes (rouges) représentent la direction de la transmission de l'infection à travers la voie marchande (les voies non-marchandes). Les agents du marché constituent les hôtes du système. N_A représente le nombre d'agents du type A et peut varier dans le temps. $N_{S\setminus D}$, $N_{S\cap D}$ et $N_{D\setminus S}$ sont respectivement les nombres *d'offreurs stricts*, *d'intermédiaires* (offreurs et demandeurs) et *de demandeurs stricts*. D'un point du vue épidémiologique, chaque agent peut-être dans un état *Sensible* (X), *Infectieux* (Y), ou *Retiré* (Z). Les agents deviennent infectieux au taux $\Lambda(t)$; les agents infectés sont retirés au taux γ ; et les agents reviennent dans le marché (se rétablissement) au taux ν. Ici, un 'retrait' signifie qu'un agent infectieux est détecté et retiré du marché par le gestionnaire de risque et devient donc *inactif*. Chaque agent, qu'il soit offreur strict, intermédiaire ou demandeur strict, peut être dans n'importe lequel de ces états épidémiologiques. Par exemple N_S^X et N_S^Y représentent respectivement les nombres d'offreur sensibles et infectieux. Par conséquent, il y a $N_S^{XY} = N_{S\cap D}^{XY} + N_{S\setminus D}^{XY}$ offreurs actifs. Les stocks d'offre et de demande accumulés notés respectivement S et D quantifient la *volonté d'échanger* ou *frustration* des offreurs et demandeurs. Le stock d'offre agrège l'ensemble des biens disponibles à la vente (biens réels), tandis que le stock de demande représente le nombre total de biens que les acheteurs souhaitent acquérir (biens virtuels). Le *flux commercial* Φ émane des stocks d'offre et de demande et agrège l'ensemble des échanges entre agents offreurs et demandeurs qui acceptent d'échanger une part de leurs stocks respectifs au prix p. Les transactions se produisent au *taux de transaction* Θ, et à chaque transaction, un *nombre moyen de biens* q est échangé. Lorsque le marché atteint sont équilibre temporel, q est égal à κ, *le coefficient de friction*. Si le flux marchand est maintenu au niveau d'équilibre Φ^*, alors la transaction des biens se produit à une fréquence moindre $\Theta^* = \Phi^*/\kappa$ quand le nombre moyen de biens échangés par transaction $q = \kappa$ est augmenté (voir la section 'Results' du chapitre 6).

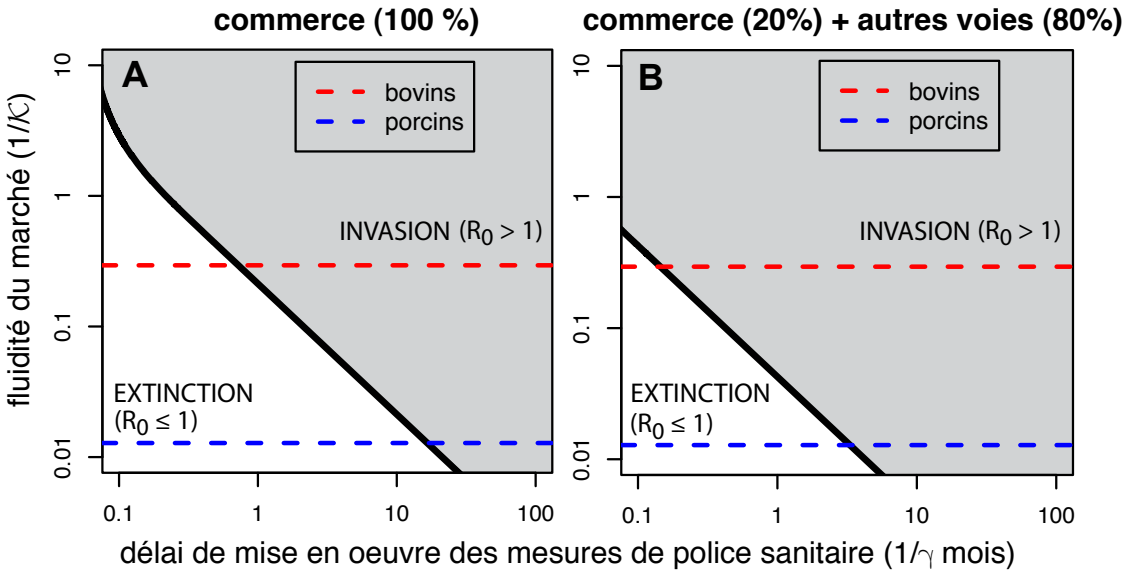


FIGURE 10 – **Le délai maximal de mise en oeuvre des mesures de contrôle permettant encore de prévenir les épidémies dépend de la fluidité du marché et de la présence d'autres voies de transmission.** Probabilité d'invasion P_I pour différents niveaux de fluidité de marché $1/\kappa$, délais de mise en oeuvre des mesures de contrôle $1/\gamma$ et intensité des autres voies de transmission (**A-B**). Pour chaque panneau, la courbe noire représente l'équation $R_0 = 1$ qui sépare l'espace $(1/\kappa, 1/\gamma)$ en deux sous-espaces : l'aire sous (au-dessus de) la courbe conduit à une extinction (invasion) de la maladie infectieuse, c'est-à-dire que $P_I = 0$ ($P_I = 1$). L'épidémie est soit uniquement causée par le commerce (la courbe noire est donnée par $R_0^{tr}(\kappa_c, \gamma_c) = 1$ $R_0 = 1$ dans (6.24) ; **A**) ou par le commerce et d'autres voies de transmission (la courbe noire est donnée par : $R_0^{tr}(\kappa_c, \gamma_c) = 0.2$ $R_0 = 0.2$ dans (6.24) ; **B**), où $1/\gamma_c$ est le délai maximal de mise en oeuvre de mesures de police sanitaire correspondant au niveau de friction critique $1/\kappa_c$, la seule valeur de $1/\kappa$ pour laquelle $R_0 = 1$. La ligne rouge (bleue) en tiret représente le niveau estimé de la fluidité du marché bovin français (du marché porcin français). Les autres paramètres sont identiques à la Fig. 6.3.

Synthèse du chapitre 7 – modéliser le commerce et les épidémies imputables aux échanges au sein de marchés avec agents hétérogènes

En lien direct avec le chapitre 6, le septième chapitre (voir page 163) expose des approches préliminaires visant à répondre à des questions spécifiques qui se posent lorsque diverses sources d'hétérogénéités dans les caractéristiques des agents et de leurs interactions sont prises en compte. La majorité des résultats que nous présentons ont été obtenus par Sébastien Geeraert dans le cadre d'un stage de Master dont j'ai assuré la co-direction (annexe F, en français). Mes contributions personnelles sont les suivantes : la formulation des modèles d'échanges commerciaux et des épidémies imputables à ces échanges, la conception des expériences numériques pour combler des lacunes de connaissances, et des propositions d'améliorations de ces modèles à la lumière du travail de Sébastien Geeraert.

Bien qu'il apporte des enseignements utiles sur les mécanismes économiques susceptibles de jouer un rôle déterminant dans la transmission des épidémies imputables aux échanges, tout particulièrement la friction marchande, le modèle *frictional trade market* (FTM) et le modèle « marché – épidémie » (ME) explorés au chapitre 6 reposent sur l'hypothèse forte que les offreurs (demandeurs) sont tout identiques à un offreur représentatif (demandeur représentatif). Quoique nécessaire pour développer une intuition

d'ensemble sur l'influence mutuelle des dynamiques infectieuses et marchandes, supposer que les marchés sont composés d'agents homogènes n'est clairement pas corroboré par les données commerciales [e.g. Atalay *et al.*, 2011; Vernon, 2011]. En réalité, les agents économiques et leurs interactions peuvent montrer des différences marquées en degré, en poids, dans la direction d'ensemble des échanges ou en friction des liens commerciaux. Un cas type de ces sources d'hétérogénéité est donné par les marchés d'échange de bétail (chapitres 5 et 6). A l'interface de l'économie et de l'épidémiologie, nous identifions deux lacunes de connaissance susceptibles d'être comblées par la prise en considération des propriétés hétérogènes observées dans les marchés : *i*) la découverte de mécanismes comportementaux plausibles pour expliquer et prédire la dynamique des échanges commerciaux tels qu'on peut les constater pour les échanges de bovins ; *ii*) déterminer dans quelle mesure un marché présentant des frictions diversifiées peut véhiculer des infections.

Dans le septième chapitre, nous avons pour ambition de combler les deux lacunes de connaissance identifiées en couplant processus commerciaux et processus épidémiologiques à l'échelle microéconomique / microscopique. Par conséquent, nous généralisons les modèles FTM et ME pour rendre compte des échanges commerciaux et des épidémies imputables aux échanges au sein de marchés avec des agents hétérogènes. Les cadres théoriques résultants sont qualifiés de modèles *FTM hétérogène* (hFTM) et *ME hétérogène* (hME). Lorsqu'ils sont confrontés à des données empiriques d'échanges de bovins à l'échelle des agents individuels, les modèles hFTM et hME apportent des éclairages sur des mécanismes comportementaux putatifs susceptibles d'expliquer la dynamique des échanges commerciaux observés dans les données (section 7.1). Nos résultats préliminaires indiquent par ailleurs que les frictions semblent bien jouer un rôle déterminant sur la dynamique conjointe des échanges et des épidémies à l'échelle microscopique (section 7.2).

Présentation de la partie IV et synthèse du chapitre 8 – Discussion

« On ne peut avoir le beurre et l'argent du beurre ». Ce dicton s'applique parfaitement aux échanges entre agents tels que les individus et les entreprises. Les échanges couvrent en effet des besoins essentiels comme la reproduction ou la production de richesses, mais ces bienfaits impliquent des efforts importants de la part des agents et comportent par ailleurs des risques (partie I). Les échanges étant coûteux, les agents font face à des contraintes d'interaction et doivent opérer des arbitrages. Ces contraintes induisent des structures de contact très complexes et diversifiées, structures qui constituent la condition des échanges et parfois également la cause de la transmission des maladies infectieuses (partie II). Les émergences épidémiques qui résultent des échanges peuvent provoquer des dommages considérables aux agents et à leurs biens. En retour, ces dommages peuvent avoir des retombées potentielles y compris sur d'autres agents en réponse à la déstabilisation des échanges, la mise en oeuvre de mesures de police sanitaire et au déclenchement de réponses comportementales adaptatives complexes (partie III).

Dans un monde de plus en plus global, les échanges se font plus fréquents et recouvrent des zones géographiques de plus en plus larges. Il s'en suit que l'exposition d'ensemble des agents aux risques infectieux devrait croître dans les décennies à venir. Pour des raisons identiques, et parce que les décisions qui influent sur la transmission des épidémies imputables aux échanges sont prises par des agents différents et à des échelles contrastées, les types et nombres d'agents touchés par des infections et / ou impliqués dans la gestion des épidémies en résultant devrait également croître.

Dans la discussion (page 193), nous commençons par récapituler les principaux résultats, puis les enseignements plus généraux tirés de ce travail qui contribuent à apporter des réponses aux défis croissants que posent les épidémies imputables aux échanges (section 8.1). Si les résultats ont déjà été largement évoqués dans la présente synthèse francophone, on peut mentionner trois messages de portée générale (voir page 198 pour une présentation plus détaillée). En premier lieu, nous soulignons l'intérêt de mettre en oeuvre une méthodologie duale pour étudier un objet complexe tel que les épidémies imputables aux échanges. Ce type de méthodologie mobilise conjointement des modèles très simples d'une part, et plus réalistes d'autre part, et cela afin de tirer parti des avantages respectifs des deux approches. En second lieu, nous préconisons de coupler davantage processus naturels et processus sociaux au sein d'approches intégratives pour rendre compte de façon mécaniste des échanges et des épidémies. En troisième lieu, nous présentons les éclairages apportés par la mise en oeuvre d'approches comparatives où des points de vue d'agents contrastés sont pris en compte.

Puis, nous présentons cinq perspectives pour dépasser certaines limites de nos approches, et par ailleurs appliquer nos modèles intégratifs à de nouveaux systèmes et de nouveaux fronts de recherche (voir section 8.2 pour une présentation détaillée) : *i*) la conception d'outils analytiques adaptés pour rendre compte d'épidémies propagées sur des réseaux pondérés et dirigés composés d'agents supposés passifs ; *ii*) la recherche de politiques de prévention épidémique efficaces qui minimiseraient en même temps les effets adverses sur les marchés dans des systèmes dynamiques avec agents adaptatifs et hétérogènes ; *iii*) l'application des approches intégrées du type « marché – épidémie » à d'autres contextes, en particulier la propagation de *P. ramorum* dans les systèmes agronomiques, forestiers et naturels, ainsi que la propagation des infections sexuellement transmissibles *iv*) l'ajustement de notre cadre adaptatif au cas des maladies endémiques où les agents anticipent les infections ; *v*) tirer pleinement partie des données disponibles pour faire de la sélection de modèles et valider ou infirmer certains mécanismes putatifs proposés dans le cadre de ce travail.

Nous concluons ce travail par une série de remarques de portée générale (section 8.3). En particulier, nous insistons sur les limites inhérentes aux modèles, tout particulièrement pour rendre compte de comportements humains qui sont, par nature, très complexes. Nous soulignons enfin l'intérêt de conduire des recherches interdisciplinaires.

General introduction

In an increasingly globalised and populated world, the spread of infectious diseases presents considerable risks including epidemic emergence from natural reservoirs and global pandemics [Jones *et al.*, 2013]. Infection-conducing hosts can move swiftly over greater distances and in greater volumes than ever before [Tatem *et al.*, 2006]. In contrast with non-infectious diseases, infections are transmitted from individual to individual. Though diverse routes of transmission exist, direct exchanges among hosts constitute key drivers of epidemics, and the chance for an healthy host to get infected tends to increase with the number of infectious hosts in its immediate vicinity [Keeling & Rohani, 2008]. Contact patterns among hosts resulting from exchanges are generally complex [Pellis *et al.*, 2014], involve multiple scales [Gog *et al.*, 2014], and can be considerably reshuffled by behavioural and socio-economic factors [Klein *et al.*, 2007; Funk *et al.*, 2014; Perrings *et al.*, 2014]. Due to this complexity, understanding the interrelationships between the factors driving exchanges and the dynamics of epidemics constitute key research questions. In addition to being intellectually challenging, addressing such questions is also of high practical importance, most importantly to suggest appropriate epidemic control measures [Keeling & Rohani, 2008]. In this context, our doctoral work contributes to shed light on potential mechanisms driving *exchange-driven epidemics* and discuss possible courses of action to limit infection transmission when it is attributable to exchanges.

Exchanges resulting from human-mediated interactions are crucial to our well-being. Let us take two examples exhaustively discussed in this doctoral work. At the scale of individuals, sexual intercourses contribute, among others, to recreation and procreation [Foucault, 1984]. At the scale of businesses, livestock deliveries constitute a key component of the agri-food sector [Vernon, 2011]. Both cases correspond to beneficial exchanges occurring among *agents*, which we define here as humans or units controlled by humans such as agricultural holdings. This definition of agents can be likened to the definition adopted for the implementation of numerical models known as "multi-agent systems" (MAS). For the latter models, an agent is understood as an autonomous entity, either real or virtual, willing to reach its own objectives and with capacity to exert an influence on its environment and to interact with other agents. MAS can reproduce certain economic and epidemiological phenomena through decision algorithms designed to reproduce the behaviour of real agents with a high degree of realism [see Tesfatsion, 2001; Durham & Casman, 2012, and Guesnerie [1996] cited in Callon [1998]]. For instance,

MAS can account for complex inter-temporal decisions made by economic agents. In contrast, our work relies on simpler agents' behaviour, which leads to systems more amenable to mathematical analyses compared to MAS.

Though exchanges are valuable, they can also infect agents who then act as infectious hosts, sustaining epidemics locally and promoting long-distance introductions. Exchange-driven epidemics include Human Immunodeficiency Virus through sexual encounters [May & Anderson, 1987], Foot and Mouth Disease and Bluetongue Diesease through movements of livestock [Ferguson *et al.*, 2001a] and the spread of *P. ramorum* through shipments of ornamentals [Pautasso & Jeger, 2008]. At the crossroad of human health and animal health, an example of epidemic driven by animal exchanges is the zoonosis caused by the bacterium *Escherichia coli* strain O157 which infects both humans and cattle. The main reservoir of *E. coli* O157 is the digestive system of cattle, where it remains asymptomatic. In humans, an infection by *E. coli* O157 induces serious gastrointestinal diseases and can sometimes lead to death. Human infections are generally resulting from the consumption of contaminated meat or water, or from contact with cattle faeces in the environment [Matthews *et al.*, 2013]. These few examples illustrate the major impacts of trade-driven epidemics on a wide host range.

In turn, infectious diseases and subsequent control measures can be detrimental to agents, thus affecting overall exchange dynamics [Funk *et al.*, 2009]. For instance, in 2001, British farmers were harshly hit by a major outbreak of Foot-and-Mouth Disease (FMD). The official assessment carried out in 2002 by the National Audit Office clearly attests the scope of the crisis [National Audit Office, 2002] : more than 6 million animals were culled, 80 % of which were not confirmed cases. Two causes explain this large figure : animals culled for disease prevention (44 %) and the slaughter of healthy animals that could not be moved due to exchange restrictions (36 %). These decisions were severely criticized by the stakeholders impacted by the 2001 FMD outbreak.

Due to the numerous causative factors and feedback loops at stake, disentangling the intricate relationship between epidemics and underlying exchanges remains a difficult task. To bring exchange-driven epidemics under control while maintaining efficient exchanges, it is of central importance to uncover the most important mechanisms leading to joint exchange and disease transmission dynamics. Though already important, the impacts of the 2001 outbreak of FMD could have gone worse if predictive mathematical models had not been implemented to guide risk managers from the beginning of the outbreak [Ferguson *et al.*, 2001b; Keeling *et al.*, 2001]. Plus généralement, comprendre et gérer les épidémies imputables aux échanges serait difficile, voire impossible, sans recourir aux modèles mathématiques [Anderson & May, 1991]. In the general case, understanding and managing exchange-driven epidemics would be difficult, if not impossible, without resorting to mathematical models [Anderson & May, 1991]. Since, by definition, epidemics operate at the scale of populations, studying them experimentally is impossible for both ethical and practical reasons. Of course, the underlying disease-conducing pathogens and their dynamics within infected hosts can be explored in the lab, but epidemics do not restrict

to purely biological considerations. Besides natural dispersion through the environment as observed for air-borne pathogens, infectious processes frequently involve complex interactions among hosts such as sexual intercourses or business relationships [Keeling & Eames, 2005]. Direct inter-host interactions constitute efficient transmission pathways for pathogens. The structure of epidemic-conducing contacts among hosts, whether induced by natural or social mechanisms, is usually referred to as the *epidemiological contact structure*. Models can prove very handy to adequately describe, compare or even infer such contact structures. Models may also simulate infectious processes propagated on any postulated or observable contact structure. Data on past outbreaks or social mixing patterns are required to build, improve and select models, but the same consideration holds true the other way around : data collection usually implies some sort of implicit underlying modelling framework.

Here, we present a combination of mathematical models and empirical analyses to increase our understanding of epidemics propagated through exchanges and suggest relevant disease control policies.

Our work is structured into four main parts.

Part I provides an overview of exchange-driven epidemics and associated challenges from a personal standpoint. Among others, we discuss typical questions of practical interest and classical approaches used to tackle them, most importantly *compartmental models* [Kermack & McKendrick, 1927] and *networks* [Strogatz, 2001]. We also identify two current research gaps, namely the influence of *interaction constraints* and *adaptive (host) behaviour* on joint exchange and epidemiological dynamics. Adaptive behaviour refers to change in agent behaviour, and potentially in exposure to risk, in response to disruption such as disease outbreak ; the change may be through adaptation of social or economic mechanisms [Funk *et al.*, 2014]. Interaction constraints limit in different ways the capacity of agents to interact. They can limit the interaction of a given agent with everybody else (sparseness constraint [5]), the rate and direction of exchanges (weighting and directional constraints), or the rate of encounter between agents (frictional constraint). While the sparseness constraint has been exhaustively studied [Danon *et al.*, 2011], the epidemiological consequences of the three other constraints, whether studied separately or together, remain less well understood [Meyers *et al.*, 2006; Pellis *et al.*, 2014].

Following a bottom-up approach, parts II and III expose novel research addressing the knowledge gaps identified in part I. We initially assume agents are non-adaptive to concentrate on the sole study of interaction constraints and their consequences on disease dynamics (part II). Then, we relax the initial assumption of non-adaptive agents to explore how adaptive behaviour with constrained interactions shape joint exchange and disease dynamics (part III).

Finally, part IV discusses the overall insights gained from this PhD project and evokes perspectives for future work.

5. the term “sparseness” refers to spare matrices.

Première partie

Exchange-driven epidemics : overview and challenges

Foreword to part I

Understanding exchange-driven epidemics appears highly challenging. Not only do we have to correctly account for exchange dynamics and infection dynamics separately, but also to understand how such dynamical processes mutually influence each other. Enriched by insights derived from data analyses, mathematical models can prove invaluable to uncover the likely mechanisms underlying exchange-driven epidemics.

Building upon the epidemiological and economic literature, we adopt a combination of empirical and dynamical modeling approaches to investigate the architecture of exchanges and its influence on infection dynamics. In turn, we also explore feedback loops where epidemics can influence the underlying exchanges. In the general case when exchanges are constrained and when infection-conducing hosts adapt themselves to disruptions in joint exchange and epidemic dynamics, we argue that tackling exchange-driven epidemics requires a combination of cross-disciplinary approaches and points of views. We illustrate our purpose with two types of exchanges known to drive epidemics : trade and sex.

This first part, made of two chapters, is organised from a top-down perspective. We start by providing background information on and present typical insights gained from the crude description of epidemics and underlying exchanges with simple models (chapter 1). We present two complementary approaches largely employed by the epidemiological community : the compartmental and the network-based frameworks. Then, we adopt a mechanistic point of view and explain how simple models can be extended to incorporate two particular phenomena contributing to exchange and epidemic dynamics : interaction constraints and adaptive behaviour (chapter 2). For each phenomenon at hand, we highlight key research questions of practical interest and briefly describe the particular concepts and mathematical standpoints adopted to address them. Note that we reserve detailed states of the art for each chapter of results (parts II and III), where we address specific research questions.

Chapitre 1

Understanding exchange-driven epidemics with simple models

Exchange-driven epidemics constitute complex dynamical phenomena potentially influenced by a plethora of causal factors, which render them difficult to study. In such cases, even simple mathematical models can describe dynamical infection processes conduced through exchanges and shed light on potential mechanisms underpinning disease transmission.

In this chapter, we start by presenting and analysing one of the simplest epidemiological model, the *SIRS compartmental model*, which provides interesting insights on exchange-driven epidemics. Then, we make explicit the assumptions underlying the SIRS compartmental model by adopting a *network-based* standpoint. We relate the SIRS compartmental model to a well-established framework, that we refer to as the *SIRS complete-network-based model*, where exchanges among agents are described by a particular architecture known as the complete network, and subsequent infections correspond to random events driven by the underlying network structure. Finally, we show the SIRS compartmental and complete-network-based models yield, on average, identical results, and share key mechanistic assumptions, namely *homogeneous-mixing* and *passive host behaviour*. Both frameworks are employed in this work and can be extended to account for richer contact mechanisms among agents (chapter 2).

1.1 Modelling exchange-driven epidemics with the SIRS compartmental model

Infectious diseases transmitted through exchanges are extremely diverse in the range of pathogens and hosts impacted. For instance, the Foot-and-Mouth Disease (FMD) and Human Immunodeficiency Virus (HIV) epidemics can be seen as two extreme cases, located at the opposite ends of the epidemic spectrum. FMD implies quick and bursting epidemics with short infectious periods (an infected remains infectious for about a week), exemplified by the 2001 outbreak in the UK [Ferguson *et al.*, 2001b]. In

contrast, the HIV pandemic appears rather slow-moving, and is associated with very long infectious periods (about a decade). When FMD quickly impacts all cattle located within a given farm, HIV slowly infects humans one after another, implying the corresponding disease-conducing agents correspond to farms [1] and individuals respectively. Though both FMD and HIV appear essentially transmitted through human-driven exchanges, it may seem difficult, at first sight, to describe them with a common modelling framework. It turns out it is actually feasible, and with a fairly simple model. In the 1920s, Kermack & McKendrick [1927] introduced a groundbreaking epidemiological model, nowadays referred to as the *SIR compartmental model*, which describes epidemiological transitions occurring at certain rates among compartments. Both the FMD (in countries where mass-culling is systematically applied) and HIV epidemics can be described, as a first approximation, by the SIR compartmental model [chapter 2 of Keeling & Rohani, 2008].

Here, we present and explore a widespread model, namely the SIRS compartmental model (Fig 1.1A), which extends the SIR compartmental model and allows to describe an even larger range of exchange-driven epidemics.

1.1.1 Model formulation

We now detail the ingredients composing the SIRS model.

We start by describing the *epidemiological compartments*. Agents, acting as hosts to a given disease-causing pathogen, are clustered into compartments S, I or R depending on their epidemiological status. Compartment S represent *susceptible* agents or *susceptibles*, i.e. agents not immune against the pathogen and that can become infected. Compartment I stands for *infected* or *infectious* agents, i.e. agents with the ability to infect susceptibles. The remaining individuals correspond to *detected and removed* or *immunised* agents, and belong to class R. This simple model assumes all infected agents are infectious (exposed individuals become instantly infectious), and implies that detected cases are instantly removed (there is no notion of cryptically infectious individuals becoming progressively detectable over time). Such assumptions can be easily relaxed by adding new compartments. For sake of simplicity, and following the literature, we also denote with S, I and R the *numbers* of susceptible, infected, and detected and removed agents respectively. Let N represent the total population size. By neglecting demographic processes (i.e. birth and death), which fall outside the scope of this work, we notice $N = S + I + R$.

We can now detail the *epidemiological transitions* occurring between the compartments, that are described through *rates*, i.e. quantities which are physically homogeneous to quantities per time unit. Susceptibles become infectious at rate Λ, also called the *force of infection*. In the mass-action (frequency-dependent) formalism [section 2.1. of Keeling & Rohani, 2008], the force of infection Λ is

1. as a good approximation [Ferguson *et al.*, 2001a].

assumed to be proportional to the *current proportion of infectious* in the population :

$$\begin{aligned} [\text{force of infection}] &= [\text{constant}] \cdot [\text{current proportion of infectious}] \\ \Lambda(I) &= \beta \cdot (I(t)/N) \ , \end{aligned} \tag{1.1}$$

where the constant β represents the *rate of infection transmission*, i.e. the rate at which an infectious agent infects susceptibles. Infectious agents are detected and removed at rate γ. On average, infected agents remain infectious over time period $1/\gamma$, a quantity referred to as the *infectious period*, which represents the average time period spent by an infected until detected and removed. Removed agents eventually re-enter the susceptible category at rate ν, either because of waning immunity within hosts or because agents reconstitute their stocks of susceptibles. Typical examples of outbreaks to which the SIRS can be applied include chlamydia, influenza, the porcine reproductive and respiratory syndrome and FMD in countries where FMD remains endemic.

Integrating the processes described above, notably the assumption on the force of infection $\Lambda(I)$ (1.1), the SIRS model is formalised with ordinary differential equations (ODE) and reads :

$$\begin{aligned} \frac{dS}{dt} &= \nu R \ - \ \Lambda(I)S \ , \\ \frac{dI}{dt} &= \Lambda(I)S \ - \ \gamma I \ , \\ \frac{dR}{dt} &= \gamma I \ - \ \nu R \ . \end{aligned} \tag{1.2}$$

The SIR model corresponds to the particular case when $\nu = 0$. Other noteworthy sub-models derived from the SIRS model include the SI model (when $1/\gamma \to \infty$, a common assumption for HIV on the short run ; see chapter 3) and the SIS model (when $\nu \to \infty$; a model we apply to describe the spread of *P. ramorum* in ornamentals ; see chapter 4).

1.1.2 Model exploration

We now briefly explore the behaviour of the SIRS compartmental model (1.2) and explain typical practical insights gained from its analysis.

We start by describing the *epidemic threshold* phenomenon. By construction, when most agents are initially susceptible ($S(0) \approx N$), an infected individual will infect β susceptibles per time unit (1.1). Since an infected remains infectious for an average time period $1/\gamma$, it will eventually infect β/γ hosts until being detected and removed. This ratio β/γ is defined as R_0, the *basic reproduction number*, undoubtedly one of the most important quantity in infectious disease epidemiology [Anderson & May, 1991]. R_0 represents the *average number of susceptible infected by a single infectious individual in an entirely susceptible population.* It naturally follows [2] that if $R_0 \leq 1$, the epidemic will never take-off

2. the following only stands for the ODE formalism, as stochastic infection processes with R_0 larger than 1 can get

(Fig 1.1A), while if $R_0 > 1$, the pathogen will invade the population (Fig 1.1C-D). We say that R_0 is a threshold parameter, as it controls the qualitative behaviour of the SIRS model, depending on its position with respect to 1.

Let us now describe *epidemic dynamics* when $R_0 > 1$. For $R_0 > 1$ kept constant, we represent two types of epidemic trajectories derived from the SIRS model : SIR epidemics where the pool of infected will eventually die out (Fig 1.1C), and SIRS epidemics where infected become endemic (Fig 1.1D). Here we take $R_0 = 4$, a value compatible with estimations from observed outbreaks of FMD and HIV [chapters 1 and 2 of Keeling & Rohani, 2008], both roughly corresponding to SIR epidemics. As far as the SIR model is concerned (Fig 1.1C), the only difference between simulated FMD and HIV outbreaks does not lie in their shapes, but rather on their relative timescales controlled by the time unit underlying rates (typically expressed in weeks for FMD and years for HIV).

Finally, the SIRS model and R_0 can also prove useful to inform control policies. Let us take the example of vaccination coverage. Imagine we would like to prevent future outbreaks, which proportion of agents should be immunised ? Let p denote the proportion of agents to immunise. Solving $\frac{dI}{dt} \leq 0$ in (1.2) at time 0, we notice an epidemic can never take-off if $S(0) = N(1-p) \leq N/R_0$. It follows $p \geq 1 - 1/R_0$. In other words, the SIRS model predicts it is sufficient to vaccinate a critical proportion of agents $p_c = 1 - 1/R_0$ to prevent future outbreaks (75% of herds in the case of FMD with $R_0 = 4$). The fact that $p_c < 1$ for most values of R_0 is called *herd immunity*, meaning that we do not actually need to vaccinate all susceptibles to prevent an outbreak.

Despite its extreme simplicity, the SIRS compartmental model can represent a wide range of epidemic dynamics, predict the future course of potential outbreaks and suggest relevant control measures [Anderson & May, 1992; Grenfell & Dobson, 1995]. In particular, the core ideas underlying the SIRS compartmental also apply to many exchange-driven infectious diseases in humans, animals and plants, and constitute the basis of all studies included in this work.

extinct due to random extinction events in the early stage of an outbreak [see section 6.3.3. of Keeling & Rohani, 2008]

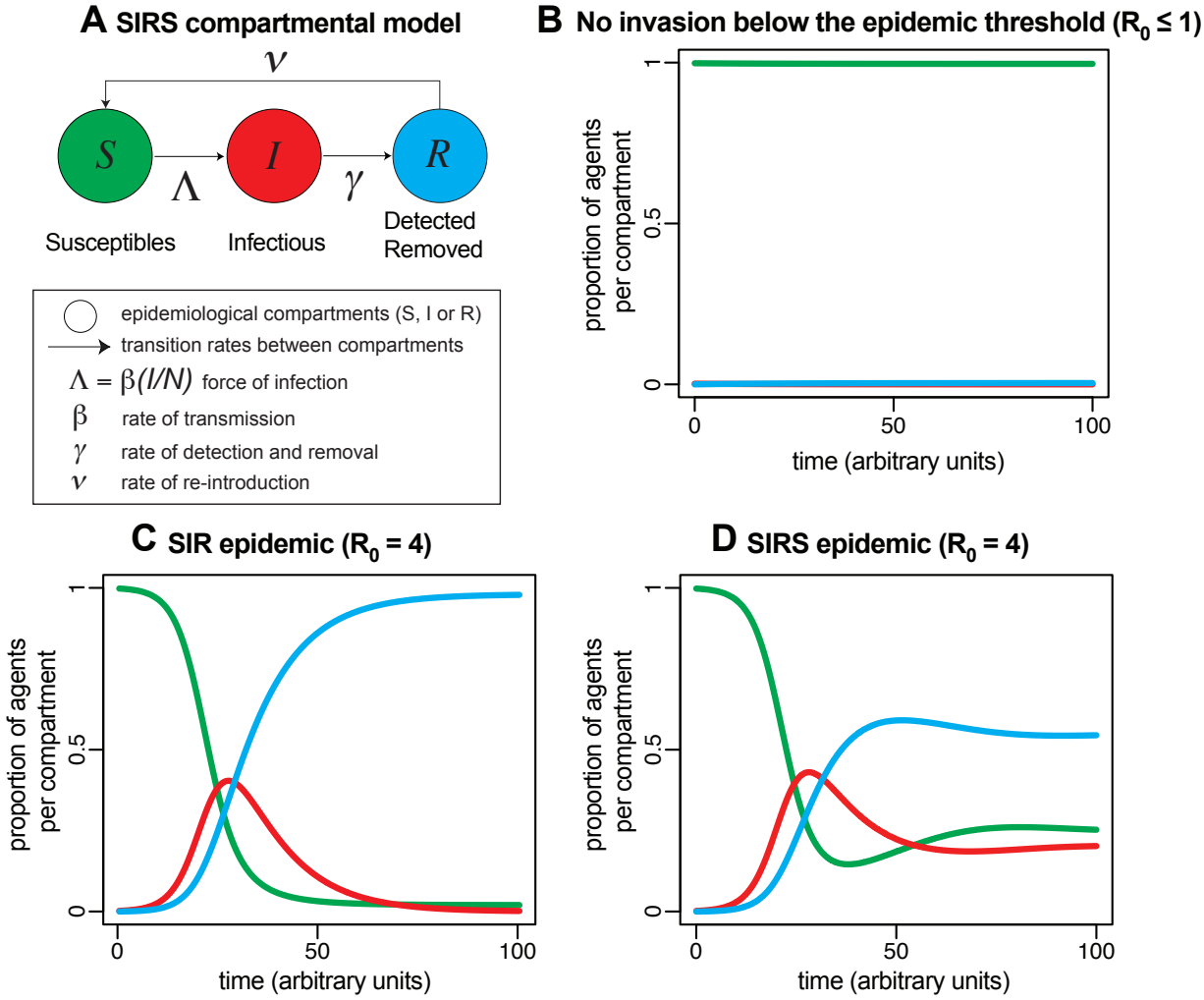


FIGURE 1.1 – **The SIRS compartmental model.** Model structure (**A**) and typical model outcomes (**B-D**). Agents are susceptible (S), infected (I) or detected and removed (R). Susceptible agents become infectious at rate $\Lambda = \beta \cdot (I/N)$. Infectious agents are detected and removed at rate γ. Removed agents recover at rate ν. The total population size $N = S + I + R$ is assumed to remain constant. $R_0 = \beta/\gamma$ denotes the basic reproduction number and must be strictly larger than 1 for an epidemic to occur. Epidemic trajectories (**B-D**) represent S/N, I/N and R/N as function of time. We consider three typical outcomes of the SIRS model : an absence of epidemic (**B**, with $\beta = 0.05 < \gamma = 0.1$), an SIR epidemic (**C**, with $\beta = 0.4 > \gamma = 0.1$ and $\nu = 0$) and an SIRS epidemic (**D**, with $\beta = 0.4 > \gamma = 0.1$ and $\nu = 0.04$). Rates are expressed in arbitrary units. For both cases, 20 agents are initially infected ($I(0) = 20$) in a total population of $N = 10000$.

1.2 From the compartmental to the network-based standpoint

Though straightforward, a compartmental framework does not come without drawbacks and over-simplifying assumptions. Here, we present existing approaches which attempt to uncover the contact and infection mechanisms underlying the compartmental SIRS model by mapping it to particular *network-based* epidemiological frameworks.

1.2.1 Opening the compartmental black box

The assumptions underlying the SIRS compartmental model can be made explicit by adopting a top-down approach. The general method employed in the literature [box 2.1 in Keeling & Rohani, 2008] consists in breaking down the rate of transmission β (1.1), which essentially acts a as black box, into the product of two mechanistic terms : a *contact term*, describing the epidemiological contact structure, and a *per contact transmission term*, specifying the infection process. We start by shedding light on the contact term by adopting a network-based standpoint. Then, we specify the per contact transmission term.

Specification of the contact term

Here, we focus on the identification of the contact term underlying the SIRS compartmental model. To mimic outbreaks, the SIRS compartmental model indeed specifies an implicit epidemiological contact structure, i.e. a structure of contacts among agents responsible for infection transmission. We can uncover this contact structure by modelling explicitly the N agents and their relationships using *networks*, which constitute intuitive tools to describe exchanges occurring among agents [Goyal, 2009]. A *network* or *graph* is a set of *nodes* or *vertices* interconnected by *links* or *edges*. Here we focus on *exchange networks*, where nodes and links represent agents and exchanges respectively. Interestingly, exchange networks explicitly describe the epidemiological contact structures underpinning infection events. Rephrasing the initial problem, we would like to find which network(s) most likely underpin the compartmental approach, i.e. identify network structures spanned by the compartmental framework. We present in turn and compare three approaches from the literature which provide a mechanistic, network-based, interpretation of the contact term : the complete-network-based interpretation, the random regular network interpretation, and the dynamical regular network interpretation.

The complete-network-based interpretation [Aparicio & Pascual, 2007]. Assuming all agents are mutually inter-connected, Aparicio & Pascual [2007] set the contact term to $N-1$. This interpretation of the contact term corresponds to the *complete network*, the unique network in which each of the N agent is connected to the $N-1$ other agents (Fig. 1.2A).

The random regular network interpretation [Bansal *et al.*, 2007]. Assuming all agents have a constant number of contacts, Bansal *et al.* [2007] rather set the contact term to a constant number of links per agent, that we denote k. By construction, $1 \leq k \leq N-1$, and the authors are typically interested in cases where $k \ll N-1$, as observed in empirical networks. Note that we recover the approach by Aparicio & Pascual [2007] for $k = N-1$.

This interpretation of the contact term corresponds to a random regular network, i.e. a network generated at random where each agent has exactly k links. In practice, a random regular network is constructed with the configuration model (CM) [Molloy & Reed, 1995], a key algorithm to generate networks randomly and under constraint from any distribution of links per agent, denoted P_k. The CM algorithm works as follow : *i*) each agent i is assigned a random number k_i of 'stubs', i.e. ends of edges emerging from the agent. k_i is drawn from a given degree distribution P_k under the constraint that $\sum_{i=1}^{N} k_i$ must be even ; *ii*) the network with pre-assigned degree distribution P_k is constructed by connecting pairs of stubs uniformly at random to yield actual links. In particular, random regular networks are generated with the CM algorithm with $P_k = 1$ for a chosen k [Molloy & Reed, 1995].

The dynamical regular network interpretation [Volz & Meyers, 2007]. Assuming contacts change over time, Volz & Meyers [2007] introduce a neighbour-exchange (NE) model to study the influence of the dynamical rewiring of links on SIR disease dynamics. The NE model described a rewiring process involving pairs of links in groups of four agents, where a link is swapped with another link at a given mixing rate. It follows the number of links per agent at any given time remains unchanged. As a limiting case, when the mixing rate becomes infinite, and when each agent has exactly one concurrent link, the NE model with SIR dynamics converges to the SIR compartmental model.

The contact term can hence be interpreted as an infinite mixing rate, and the resulting network is both dynamical and regular with $k = 1$.

Comparison of the three interpretations. All three interpretations of the contact term appear reasonable, and imply regular contact patterns where each agent has exactly the same number of contacts or/and the same contact rate. These particular structures of contacts are usually referred to as *homogeneous-mixing* or *random-mixing* in the epidemiological literature.

However, we find the complete-network-based interpretation more parsimonious for three main reasons :

– *Homogeneous contagion mechanism.* In the compartmental model formulation (1.2), we notice the force of infection $\Lambda(I) = \beta \cdot (I/N)$ (1.1) is applied to *all* susceptibles in the population. In other words, each infected individual is assumed to contribute to the infection of all susceptibles in the population, which intuitively corresponds to the complete-network-based (see section 1.3 for a proof) and the dynamical regular network (see [section 2b of Volz & Meyers, 2007] for a

proof) interpretations.

– *Strict identifiability.* The dynamical regular network and complete-network-based interpretations have the advantage of providing a unique — hence unambiguous — mapping from the compartmental to the network-based standpoint. For a given number of agents N, there is only one complete network and only one dynamical regular network (when the mixing rate goes to infinity), but a large number of random regular networks (one degree of freedom to specify : k).
– *Easier network generation and representation.* While generating a complete-network-based network is straightforward, generating a dynamical regular network is computationally demanding. In addition, a complete network can be easily depicted on a simple figure, which is not the case for dynamical networks.

From this point, and while the three interpretations are tenable, we will set the contact term to $N-1$, i.e. we adopt the complete-network-based interpretation of Aparicio & Pascual [2007]. In addition, using a complete network allows us to introduce interaction constraints in a progressive way (Fig. 2.1).

Specification of the per contact transmission term

Here, we specify the per contact transmission term, denoted ω, which underpin the SIRS compartmental model. We have just set the contact term to $N-1$, which implies agents constitute the nodes of a complete network. By construction, the rate of transmission is hence given by $\beta = (N-1)\cdot\omega$. Since $N-1$ is a number, ω is necessarily homogeneous to a rate for β to be a rate.

1.2.2 The SIRS complete-network-based model

The per contact transmission rate ω alone is not sufficient to generate SIRS epidemics on a complete networks. In fact, we further need to introduce an epidemic process accounting for explicit contacts among agents, i.e. an epidemic model which is actually compatible with a network standpoint. We refer to this epidemiological model as the *SIRS complete-network-based model* (Fig. 1.2A-B). This network-based framework is essentially very close to its compartmental counterpart, and we only highlight major differences between the two model formulations.

The SIRS complete-network-based model is a particular case of well-established *stochastic* epidemiological processes [section 6.3 of Keeling & Rohani, 2008]. Such stochastic processes explicitly describe, for any agent in a population, epidemiological states by *random variables* and transitions between states by *random events*. Stochastic processes are usually employed to model epidemics at the scale of individual agents whenever within-host infection dynamics are neglected [see Gog *et al.*, 2014, for further details on within-host infection dynamics].

We start by describing agents' states. Let i denote any of the N agents. Variables S_i, I_i and R_i represent the possible states of agent i at any point of time, standing respectively for the susceptible,

infectious and detected and removed states respectively. Each variable is binary (equal to 0, 'no' ; or 1 'yes') and informs us on the current epidemiological state of i. For instance, $I_i = 1$ means that agent i is infected and infectious. Since an agent can only occupy one state at a time, we have $S_i + I_i + R_i = 1$.

We then describe transitions between epidemiological states. In the general case, stochastic transitions between a series of states can be modelled with a Markov jump process [Bartlett, 1949]. Let us take the example, for an agent i, of a transition from a state X_i to a state Y_i occurring at rate r_{XY}. We assume Y_i is the only state that can be reached from state X_i. Mathematically, the time spent by i in state X_i is exponentially distributed with rate parameter r_{XY}. It follows the inverse of the transition rate ($1/r_{XY}$) is the average time spent by the individual i in state X_i before moving to state Y_i. In the SIRS complete-network-based model, each agent i can experience in turn three types of epidemiological transitions : an infection (transition from S_i to I_i), a detection-removal (transition from I_i to R_i) or a re-entry event (transition from R_i to S_i). Detection-removal and re-entry events strictly correspond to the compartmental framework : any infectious agent is detected and removed at rate γ, and any removed agent eventually re-enter the susceptible category at rate ν. However, infection events are more complex than in the SIRS compartmental framework, as they depend on the local system state in the neighborhood of susceptible agents. More precisely, a susceptible agent i becomes infected at rate λ_i^c, where λ_i^c is a *local force of infection*. We recall the rate of transmission is $\beta = (N-1)\omega$, where each agent is connected to the $(N-1)$ other agents and ω is the per-contact transmission rate. λ_i^c hence only depends on the number of infectious agents j connected to i. Assuming the local force of infection is proportional to the number of infectious agents j in contact with i, we notice :

$$\begin{aligned} [\text{local force of infection}] &= [\text{rate of infection per contact}] \cdot [\text{total number of infectious } j \text{ connected to } i] \\ \lambda_i^c &= \omega \cdot \sum_{j \neq i} I_j(t) \ . \end{aligned} \tag{1.3}$$

The resulting SIRS complete-network-based framework typically yields stochastic dynamics (100 grey curves in Fig. 1.2C, corresponding to 100 replicate runs of the model). In this particular case [3], the variability in the total number of infected is weak because we use a complete network, with a large R_0 (set to 4) and initially infect an important number of individuals ($\sum_{a=1}^{N} I_a(0) = 20$). Since this model is agent-based, we could also keep track of the state of each agent at any point of time (data not shown).

3. Examples of factors that could yield a larger variability in epidemic trajectories, typically with a mixture of extinctions and actual outbreaks, include a small population size, an heterogeneous contact structure in number of links per agent, a moderate value of R_0 (for instance 2), a small number of initially infected hosts, or a short infectious period [see e.g. chapter 6 in Keeling & Rohani, 2008].

1.3 Comparison and assumptions of the compartmental and network-based standpoints

Though different, we show the two models yield, following the law of large numbers, identical results, and share key assumptions, namely homogeneous-mixing and passive host behaviour.

1.3.1 Comparison of the two frameworks

We start by explaining how, and under which conditions, the two frameworks yield identical results. We then highlight irreducible differences between the two approaches.

Relating the two frameworks

To relate the two frameworks, we follow the ideas presented in the appendix of [Aparicio & Pascual, 2007].

Since the compartmental model relies on differential equations, we place ourselves in the limit of arbitrarily large populations ($N \to \infty$) to compare the two standpoints.

We start by re-formulating the SIRS complete-network-based model. In a complete network, any agent i has exactly $N-1$ neighbours. For large N, $N-1 \approx N$, which simplifies the expression of the force of infection (1.3) to :

$$\lambda_i^c \approx \lambda^c(I) = \omega \sum_{a=1}^{N} I_a = \omega I \ , \tag{1.4}$$

where λ^c denotes the approximate local force of infection and $I = \sum_{a=1}^{N} I_a$ is the total number of infectious at a given time. Clearly, $\lambda_i^c \to \lambda^c$ when N goes to infinity.

Now that we have an approximate expression of the local force of infection that is independent of the focal agent i, we can formally relate the compartmental and stochastic framework. Strikingly, we remark the approximate local force of infection is identical to the compartmental force of infection Λ (1.1) :

$$\lambda^c(I) = (N\omega) \cdot (I/N) = \beta \cdot (I/N) = \Lambda(I) \ . \tag{1.5}$$

From (1.5), and given the infectious period is $1/\gamma$, we deduce the basic reproduction number of the complete-network-based model : $R_0 = \beta/\gamma = N\omega/\gamma$. We conclude the two frameworks should be strictly equivalent when β and ω are chosen so that R_0 is kept constant, provided N is large enough. In practice, values of $N \geq 1000$ can already be considered large [Moslonka-Lefebvre *et al.*, 2009]. Numerical experiments show, on average, a perfect agreement between the two frameworks (black and red curves in Fig. 1.2C).

Irreducible differences between the two frameworks at hand

While yielding identical results when properly put into relation, the two frameworks exhibit irreducible differences in scales of complexity and nature of epidemic trajectories.

While the SIRS complete-network-based model describes infectious dynamics at the scale of individual agents, the SIRS compartmental framework operates at a much coarser scale.

In addition, the SIRS compartmental model based on ODE yields deterministic trajectories. In practice, if the initial number of infected is low, even epidemics with $R_0 > 1$ can get extinct due to random extinction effects, a phenomenon not captured by deterministic models [4] [see chapter 6 of Keeling & Rohani, 2008].

Such differences between the modelling approaches we present are important when deciding upon which framework should be preferentially used to address a given research question.

1.3.2 Key assumptions shared : homogeneous-mixing and passive host behaviour

Balancing from the compartmental to the complete-network-based standpoint allows us to gain interesting insights on the underlying epidemiological contact structure, i.e. the particular structure of contacts responsible for subsequent infections.

In the SIRS complete-network-based model, susceptible and infectious agents are all inter-connected as part of a complete network. In other words, susceptible and infectious agents are assumed well-mixed. In the general case, homogeneous-mixing does not necessarily imply a complete network, but all network-based interpretations of the SIRS compartmental model rely on some sort of homogeneous-mixing (see section 1.2.1).

Interestingly, we can further remark the only processes altering homogeneous-mixing are the detection-removal and re-introduction events, which respectively de-connect and re-connect agents within the complete network without further consequences. Once re-entering the network, an agent is automatically reconnected to all susceptibles and infected. Clearly, this implies agents exhibit *passive behaviour*, i.e. act as disease-conducing hosts that do not adapt their behaviour in response to joint exchange and infection dynamics.

As discussed at large in the next chapter, homogeneous-mixing and passive host behaviour constitute strong assumptions that we contribute to relax by explicitly accounting for interaction constraints and adaptive host behaviour.

4. Though not presented here, stochastic compartmental models do account for such extinction events.

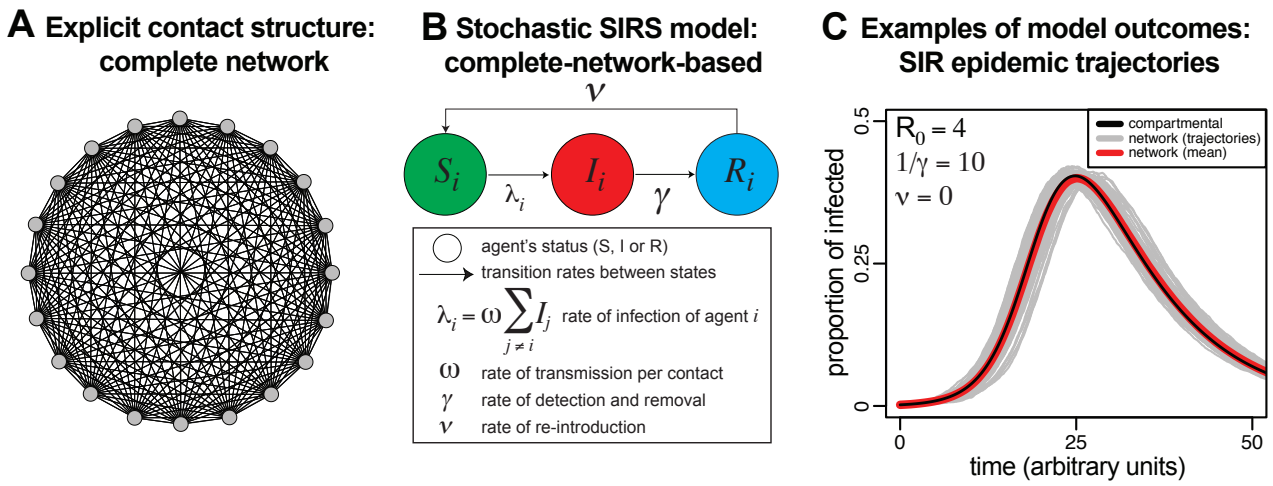


FIGURE 1.2 – **The SIRS complete-network-based model.** The SIRS complete-network-based model couples a complete network (**A**) to a stochastic network-based epidemic process (**B**), hence shedding light on the implicit assumptions underlying the SIRS compartmental model (**C**). Each agents is susceptible ($S_i = 1$), infected ($I_i = 1$) or detected and removed ($R_i = 1$). By construction, $S_i + I_i + R_i = 1$, and we neglected demographic processes, which implies a constant population size N. Susceptible agents become infected at rate λ_i^c (1.3). Infected agents are detected and removed at rate γ. Removed agents recover at rate ν. $R_0 = \beta/\gamma = N\omega/\gamma$ denotes the basic reproduction number, where β is the transmission rate of the compartmental model and ω the rate of transmission per contact of the complete-network-based model. Notice the correspondence $\beta = N\omega$ [Aparicio & Pascual, 2007]. We compare SIR dynamics as simulated with the compartmental model on the one hand (**C**, in black), and with the complete-network-based model on the other hand (**C**, in grey for individual trajectories and in red for the average out of 100 replicate runs). The stochastic process is implemented with Gillespie's direct algorithm [Keeling & Rohani, 2008]. For both cases, we initially infect 20 agents and set $\gamma = 0.1$, $\nu = 0$, and $N = 10000$. We take the values of β and ω for which $R_0 = 4$. Rates are expressed in arbitrary units.

Chapitre 2

Towards a mechanistic description of exchange-driven epidemics

Simple SIRS models are flexible and can provide interesting insights on exchange-driven epidemics. Whether compartmental or network-based, such simple models share two implicit mechanisms to account for infection dynamics, namely homogeneous-mixing and passive host-behaviour. Taken together, these assumptions correspond to a particular epidemiological contact structure : a complete network connecting all susceptible and infected agents at any point of time. The panmictic nature of contacts remains essentially invariant, including when agents are detected and removed or eventually re-enter the network (chapter 1). Common sense suggests homogeneous-mixing and passive host behaviour constitute implausible mechanisms to describe exchange-driven epidemics realistically.

Here, we review the literature and detail empirical evidence arguing against homogeneous-mixing and passive host behaviour. As an attempt to provide a mechanistic description of exchange-driven epidemics, we concentrate on two mechanisms influencing joint exchange and infection dynamics : *interaction constraints*, which tend to restrict homogeneous-mixing among agents (section 2.1) ; and *adaptive host behaviour*, which relaxes the assumption of passive host behaviour (2.2). We then explain how simple SIRS models, whether compartmental or network-based, can be generalised to incorporate interaction constraints and adaptive behaviour, and yield novel complementary insights (section 2.3). The in-depth study of interaction constraints and adaptive host behaviour with constrained interactions constitute the bulk of part II and part III respectively.

2.1 Accounting for interaction constraints

Interaction constraints correspond to barriers restricting exchanges among agents. But constraints also create structural diversity. By departing from homogeneous-mixing, i.e. a reference unconstrained scenario, interaction constraints span a wide range of constrained epidemiological contact structures, which can in turn drastically affect disease dynamics. Accounting for constraints in a given contact

structure will, unsurprisingly, tend to reduce the expected severity of outbreaks when compared against its unconstrained counterpart. However, it appears far more challenging to predict the influence of constrained contact structures on epidemics *when the overall level of constraint is kept constant*. For instance, imagine we were to distribute a constant limited pool of interactions to N agents. What would be the resulting contact structures and their influence on disease dynamics ? These two questions have generated much research in network theory and network epidemiology [Keeling & Eames, 2005].

Here, we first present a short and progressive typology of constraints generating contrasted epidemiological contact structures. For each constraint considered, we discuss research questions of epidemiological importance and knowledge gaps that we contribute to address in this work. Then, we show how constraints can be quantified through simple data analyses. We focus on two examples largely studied in this work : the weighting constraint affecting individuals having sex and the frictional constraint affecting agricultural agents exchanging livestock. Finally, we explain constraints are frequently intricate and related by multiple trade-off. We provide a rationale for focusing our epidemiological studies on particular constraints.

2.1.1 Constraints diversify epidemiological contact structures and raise questions

While homogenous-mixing spans a precise range of contact patterns, interaction constraints lead to an explosion in the diversity of potential epidemiological contact structures. Constrained contact structures are hence likely to exhibit a large range of epidemiological responses. Here, we provide a brief and progressive typology of interaction constraints. We plan not to be exhaustive, but show in which sense constraints can enrich contact structures. Though some constraints can also be incorporated in a compartmental formulation, we place ourselves in the case where constrained interactions are depicted on networks. The network-based standpoint has the advantage to make apparent the diversification implied by constraints.

We progressively present and combine four types of constraints : starting with the *sparseness constraint*, we add the *weighting constraint*, the *directional constraint* and finally the *frictional constraint* (Fig. 2.1). For each constraint, we highlight research questions of epidemiological interest and refer to the corresponding chapters where we contribute to shed light on current knowledge gaps.

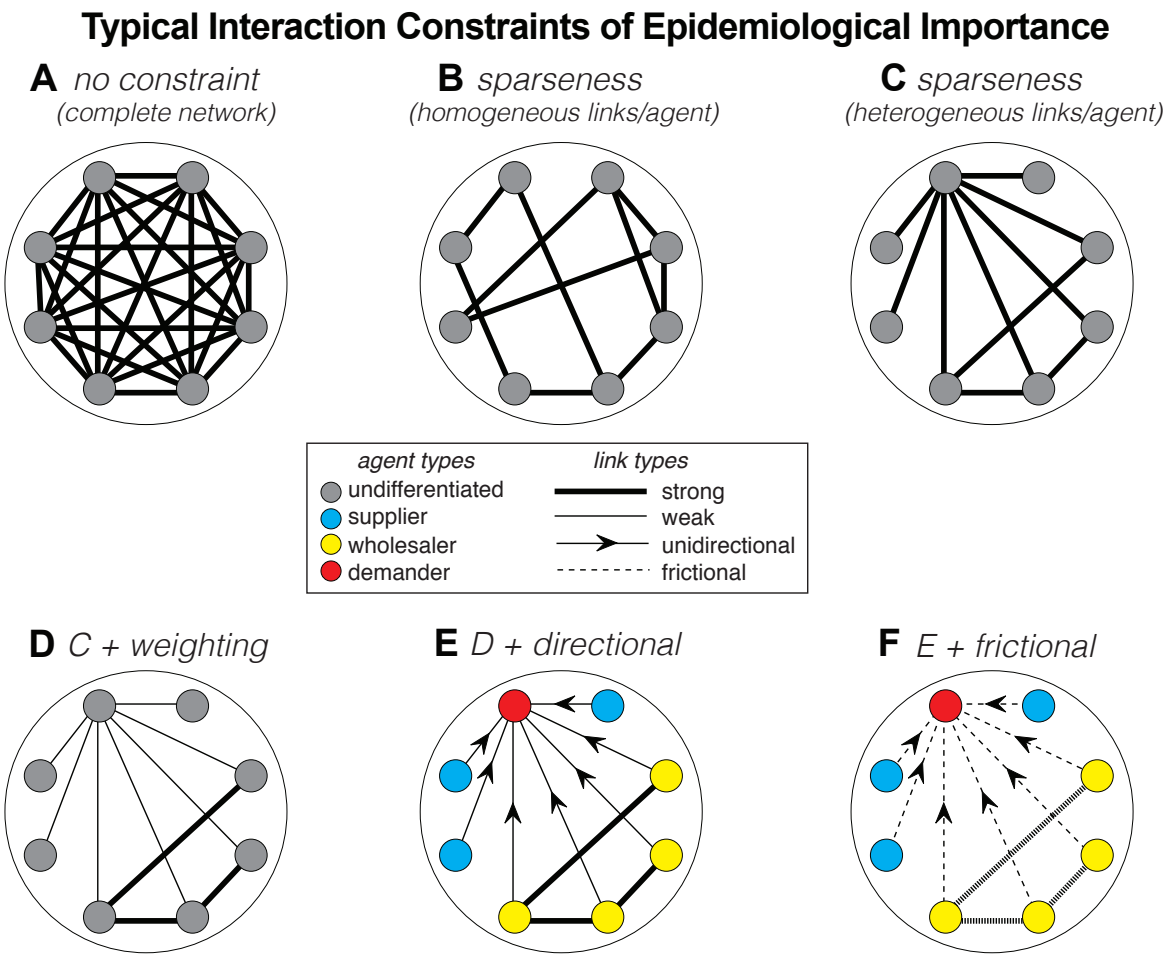


FIGURE 2.1 – **Typical interaction constraints influencing exchange-driven epidemics.** In contrast with the complete network structure underlying simple SIRS models (**A**), observed architectures of exchanges frequently exhibit multiple *interaction constraints* which contribute to explain subsequent transmission patterns. Typical constraints include the *sparseness constraint*, where only a small fraction of possible links exist, leading to distributions of links per agents either homogeneous (**B**) or heterogeneous (**C**) ; the *weighting constraint* with links exhibiting heterogeneous strengths (**D**) ; the *directional constraint*, where directions of links can reflect different types of agents such as suppliers, wholesalers demanders in markets (**E**) ; and the *frictional constraint*, implying links only exist at specific points in time (**F**).

The sparseness constraint

Most empirical contact structures depart from the complete network (Fig. 2.1A). For instance, for a given exchange type such as sexual contacts [Liljeros *et al.*, 2001] or livestock movements [Noremark *et al.*, 2011], most agents are connected to a very small subset of the total number of agents, implying infection can only spread to a limited number of agents compared with homogeneous-mixing. We call this particular barrier to exchanges the *sparseness constraint* (Fig. 2.1B-C).

Generally speaking, a network is sparse if only a small fraction of the maximal number of links exists. For a given network $\mathcal{G}$ with undirected interactions, let N and ℓ denote its number of nodes (agents) and links (exchange relationships) respectively. $\frac{N(N-1)}{2}$ is the maximal number of links within $\mathcal{G}$. Mathematically, $\mathcal{G}$ is sparse if condition $\ell \ll \frac{N(N-1)}{2}$ is satisfied. For instance, the random regular networks described by Bansal *et al.* [2007] are sparse for a constant number of links per node set to $k \ll (N-1)$ (see section 1.2.1). From a more local perspective, the sparseness constraint is not necessarily affecting all agents uniformly, i.e. all sparse networks are not necessarily regular. For a given agent i, we can quantify the intensity of the sparseness constraint by the number of partners of agent i, a quantity referred to as the *degree* in the network literature. Let $0 \leq k_i \leq N-1$ denote the degree of an agent i. As extreme cases, i may either be an entirely isolated orphan node ($k_i = 0$) or a hub connected to most agents in the network (k_i approaches the maximal degree, i.e. $N-1$). A regular network is a network for which $k_i = k$ for all i.

Due to potential variations in local connection patterns, sparse contact structures are extremely diverse. Keeping N and ℓ constant, we can typically contrast two main types of sparse networks : *networks with homogeneous degree* (Fig. 2.1B), where the spareness affects all agents uniformly ; and *networks with heterogeneous degree* (Fig. 2.1C), where agents are unequally affected by the spareness constraint.

Exploring the range of networks spanned for a given level of sparseness (i.e. for N and ℓ kept constant) and their influence on epidemic dynamics was at the core of network epidemiology [see e.g. Keeling & Eames, 2005, for a review]. In section 2.3, we explain how the sparseness constraint can be incorporated in network-based models and explain how it affects epidemiological dynamics. While the sparseness constraint has already been largely studied, the subtle epidemiological implications of the other constraints we consider remain largely unknown. This is a knowledge gap we contribute to address in this work.

The weighting constraint

So far, we implicitly assumed all contacts within a contact structure were identical. In fact, contacts may considerably vary in *strength*, i.e. can be characterised by different *weights*, implying a *weighting constraint* (Fig. 2.1D). For instance, the relative trade volume among pairs of farms exchanging cattle

can considerably vary [Vernon & Keeling, 2009]. So does the number of sex acts per partnership in sexual contacts [Blower & Boe, 1993]. Both examples of contacts can hence be described by *weighted links* belonging to *weighted networks*.

In what follows, a weight is simply a value assigned to a link, which may be indifferently static or dynamical. By summing the weights of links associated with an agent i, we obtain the weight or strength of this agent, which to some extent can be considered a generalisation of the degree [Barrat *et al.*, 2004b]. Provided distributions of weights of links are heterogeneous, i.e. consist of mixture of weak and strong ties, weighted networks can drastically depart from their unweighted counterparts [Moslonka-Lefebvre *et al.*, 2012a].

Altering the weight-per-agent distribution on a given network can have drastic consequences on epidemic dynamics [Moslonka-Lefebvre *et al.*, 2012a]. In chapter 3, we detail a complete framework to understand SIR epidemics on weighted networks.

The directional constraint

Many contact structures involve some kind of asymmetric relationships among pairs of agents, which in turn associates with an asymmetric risk of infection transmission. We hence call this particular restriction to exchanges the *directional constraint* (Fig. 2.1E). One frequent reason for this asymmetry in links is the differing nature of agents composing a network. For instance, the risk of HIV transmission differs between men and women during a heterosexual vaginal intercourse [Gomez-Gardenes *et al.*, 2008]. Another example is given by the directional shipments of contagious ornamentals from producers to wholesalers, and then from wholesalers to demanders [Pautasso *et al.*, 2010].

Here, we preferentially describe the asymmetric relationships resulting from the interaction of differing agents with *directed networks*, where asymmetric interactions between agents correspond to directed links and weights. The directional nature of connections implies two standpoints should be distinguished when characterizing agents : the in-wards and the out-wards standpoints, respectively focusing on in- and out-connections. For instance, we can define an in-degree and an in-weight for an agent by summing all links and weights pointing towards this agent respectively. In other approaches, the multiple natures of agents composing a directed network are explicitly modelled with multipartite networks, i.e. networks composed of more than one type of agents [Gomez-Gardenes *et al.*, 2008]. Of course, multipartite networks are not always directed, and agents with dissymmetrical interactions are not necessarily differing in nature. This is a simplification we make in this work.

In practice, we are again particularly interested in networks exhibiting a marked heterogeneity between the in-wards and out-wards standpoints, and made of multiple agent types. The directional constraint, and its likely influence on disease dynamics, is specifically studied in chapter 4.

The frictional constraint

What we call a connection or link does not always exist at any point in time. For instance, two farms bounded with a long-standing business partnership are not continuously exchanging livestock. Rather, the actual interaction between two farms operates as discrete jumps, which occur at a certain frequency that we call the transaction rate. A transaction results in a shipment of a batch of animals, i.e. a set of animals transported at the same time. The continuum ranging from fluid (quasi-permanent) business relationships to highly frictional (rare) transactions corresponds to increasing intensities of the *frictional constraint* (Fig. 2.1F). The notion of friction is widely used in labour economics to explain the existence of unemployment (see the model of Diamond, Mortensen and Pissarides [Pissarides, 2011]), but remains novel in epidemiology.

Formally speaking, friction is quantity defined at the level of links that tends to slow down the interaction rate between two agents, but without necessarily affecting their rate of exchange. For instance, if two agents exchange 100 goods per year, they can do so in 100 transactions of 1 good each, or in 1 transaction of 100 goods, with all possible intermediary cases in between. Assuming equilibrium in exchanges, friction can be estimated as the average batch size exchanged from an agent to another : the larger the batch, the slower the transaction rate for a given exchange rate (part III). From a mechanistic standpoint, frictions aggregate physical impediments to exchanges, such as the time and effort needed for searching business partners, cutting deals, and delivering goods. At the scale of a whole network, the overall level of friction can be calculated as the aggregated exchange rate divided by the aggregated transaction rate.

The influence of friction on disease dynamics is not intuitive, because increasing the overall level of friction results in two conflicting effects : one the one hand, the probability of infection transmission is expected to increase with the average batch size ; one the other hand, the transaction rate, which constitute the epidemiological rate of contact, will decrease, which will in turn tend to reduce the rate of transmission. Part III addresses this knowledge gap from both a compartmental and a network-based standpoints.

2.1.2 Quantifying constraints in data

There are multiple interactions constraints, and combining them results in enriched epidemiological contact structures. But this remains a conceptual point-of-view. How, in practice, can we decipher the most important constraints to focus on for a particular epidemiological system of interest ? A possible answer consists in studying exchange data. Here, we analyse such data for two systems largely studied in this work : agents having sexual relationships, which e.g. contribute to the transmission of HIV, and agricultural agents exchanging livestock, which can result in the propagation of livestock diseases such as FMD.

Data on sexual behaviour suggest a constant weighting constraint per agent

Sexual relationships constitute one of the central pathways along which sexually transmitted infections (STI) are propagated. Understanding the contact structure implied by sexual contacts is crucial to shed light on STI dynamics and propose appropriate control policies.

While the detailed and complete architecture of sexual-contact networks remains largely inaccessible (chapter 3), interesting insights on interaction constraints can be gained by analysing indirect data. For instance, the National Survey of Sexual Attitudes and Lifestyles [NATSAL Johnson *et al.*, 2001] contains valuable information on sexual behaviours in the UK, which notably reports how many sexual partners and sex acts were declared by surveyed participants over various time periods.

Our analyses of the NATSAL data (Fig. 2.2 and chapter 3) suggest the existence of a trade-off between the total number of partners (x) and the average number of sex acts per partnership (y) : a larger degree generally implies a lower weight per link. Regressions of log-linearised power-laws (fits of $y = ax^b$ where a an b are parameters to estimate) suggest this budget is constant for a given period of time and roughly equal to the product xy for all individuals. In other words, individuals appear to have a limited sexual budget, expressed in total sex acts per time period, which seems to correspond to a constant weighting constraint per agent. The idea of bounded sexual budget per individual is in line with previously reported findings [Blower & Boe, 1993].

Livestock-exchange data imply the existence of trade friction

The frictional constraint may be easier to grasp by adopting a comparative perspective. Here, we compare two datasets recording movements of animals occurring in France : the BDNI dataset for cattle in year 2009, and the BDPorc dataset for swine in year 2010. We neglect foreign movements and movements to slaughterhouses.

While the trade flow, i.e. total exchange volume per time unit, in the French cattle and swine markets appear relatively similar, we notice a stark difference in their total transaction rates, reflecting the fact that on average, far more swine are shipped per transaction than cattle (see chapter 6 for further details). Interestingly, we remark agents exchange batches of animals rather than individual animals. Let κ denote the typical average batch size in a given market, which is simply the total trade flow divided by the total transaction rate. Since it is impossible to exchange (living) animals in less than one unit, it follows κ is necessarily larger than 1. Due to a combination of production constraints and costly transportation of single animals, we further expect $\kappa > 1$.

κ corresponds to friction, and simply means agents tend to exchange animals as batches of a typical size κ. Now, imagine we could increase κ for a given system, by e.g. setting taxes on transports or imposing a minimal shipment size. What would be the epidemiological consequences and would it necessarily affect trade flow ? This is one of the central question we tackle in part III.

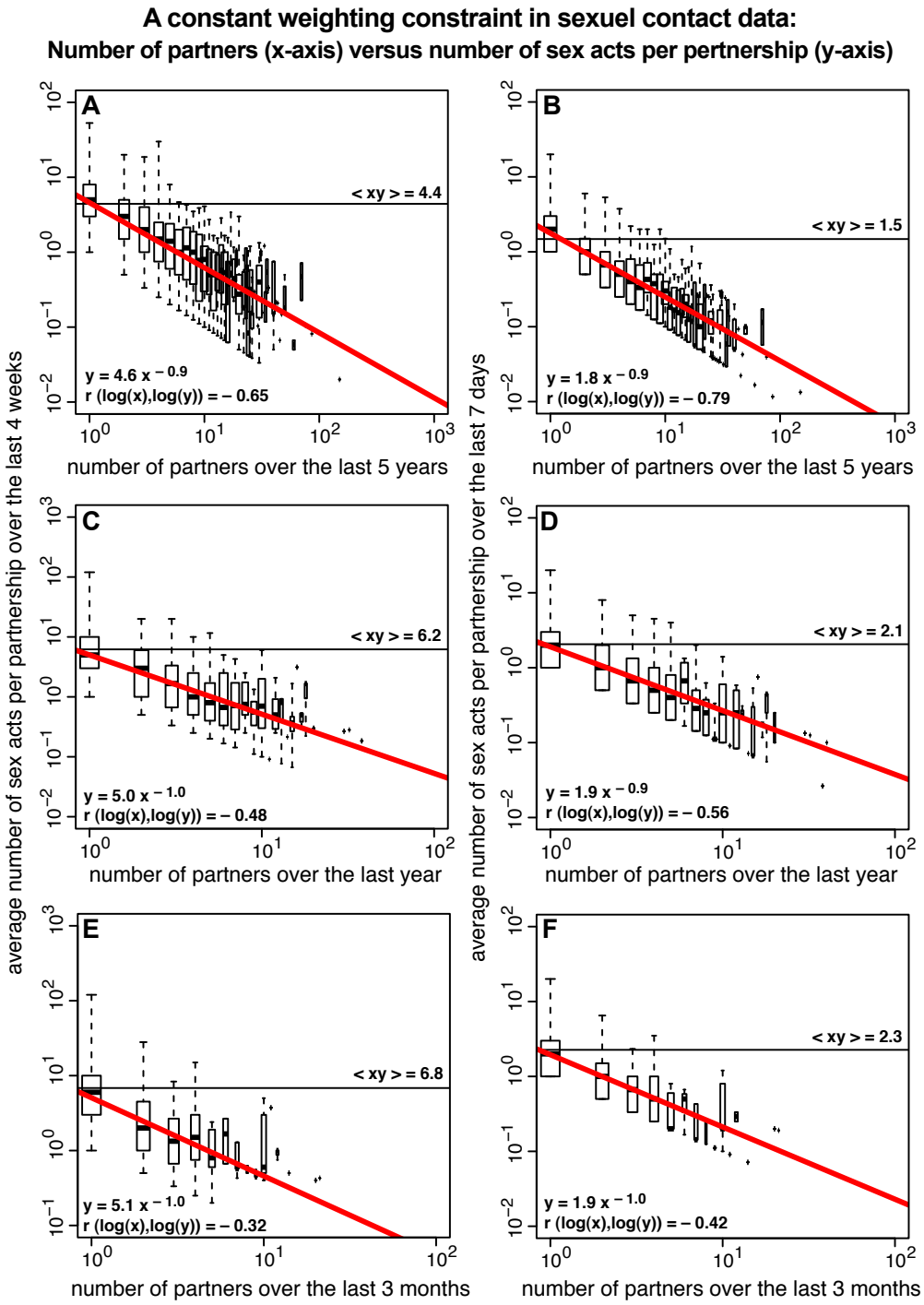


FIGURE 2.2 – **Sexual-contact data suggest the existence of a constant weighting constraint per agent.** Average number of sex acts per partnership over the last 4 weeks (**A,C,E**) and 7 days (**B,D,F**) (y) as function of the total number of sexual partners over the last 5 years (**A,B**), year (**C,D**) and 3 months (**E,F**) (x). Data are extracted from the NATSAL survey [Johnson *et al.*, 2001]. Red lines correspond to the best fits to data of the power-law model $y = ax^b$ using log-linearised regressions. r denotes the corresponding Pearson's linear correlation coefficient on the log-log scale. Black lines represent $\langle xy \rangle$, the average number of sex acts per person as calculated from data. The relative size of a boxplot is proportional to the log number of individual data it contains.

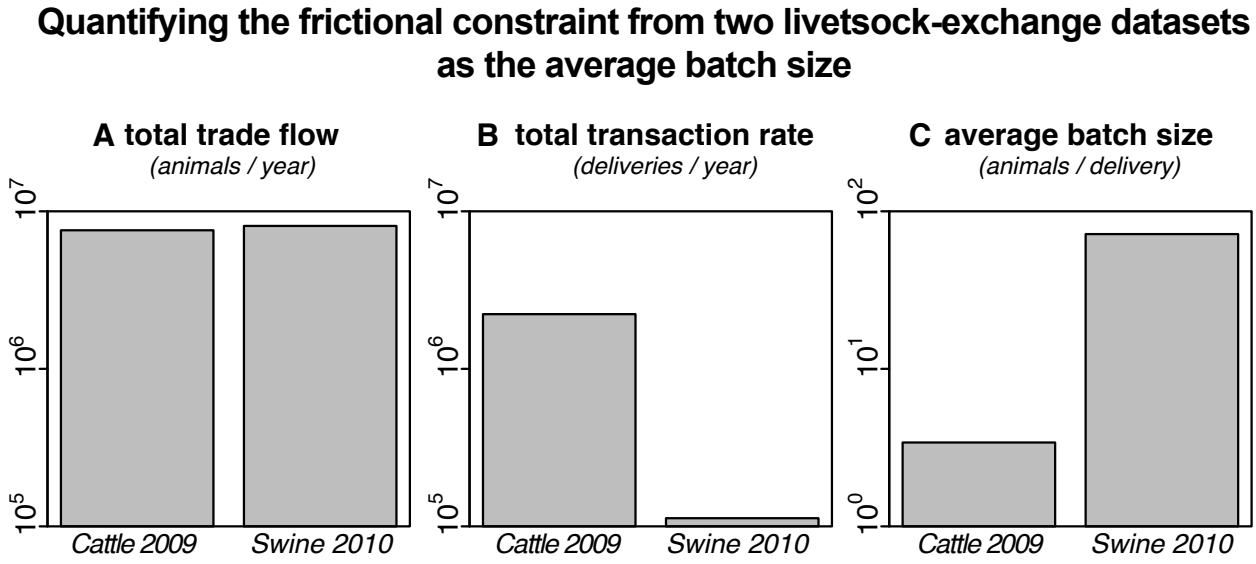


FIGURE 2.3 – **Livestock-exchange data imply the existence of friction.** Total trade flow (**A**), total transaction rate (**B**) and average batch size (**C**) calculated from two French livestock-exchange datasets : the BDNI dataset for cattle in year 2009, and the BDPorc dataset for swine in year 2010. For both cases we neglect foreign movements and movements to slaughterhouses. The total trade flow represents the total number of animals exchanged per time unit, while the total transaction rate represents the total number of deliveries per time unit. The average batch size, which we assimilate to fiction, is obtained by dividing the total trade flow by the total transaction rate, and represents the average number of animals shipped per transaction.

2.1.3 Constraints are intricate with multiple trade-offs : a bottom-up approach

Constraints are calculable from data and can be presented as elementary trade-offs. For instance, the weighting constraint in sexual-contact data translates as a trade-off between total sexual partners and average sex acts per partner ; and the frictional constraint in livestock-exchange data is described as a trade-off between the transaction rate and the average shipment size. Sometimes, the study of constraints is simplified because they do operate in isolation from one another. For instance, in contrast with sexual-contact networks, the trade-off between an agent's degree and average weight per link expressed as a trade flow, is not observed in the French livestock movements (data not shown). This difference between the two exchange systems probably stems from a stark difference between the types of agents composing them. In contrast with agricultural holdings which exhibit a strong heterogeneity in sizes, types and trade volumes (see e.g. chapter 5), individuals composing a sexual network appear more alike in their number of sex acts per time unit (Fig. 2.2).

However, in the general case, constraints are frequently intricate and related by multiple trade-offs. Typically, the sparseness constraint is almost always present in empirical contact structures. And additional constraints appear intricate. Let us take the case of sexual contacts. Not only are agents facing a constant weighting constraint over all their partnerships, but they are also likely to exhibit a frictional constraint per partnership. In other words, individuals would operate for each partnership of a given weight, e.g. expressed in total time spent having sex, a trade-off between the frequency

of interaction and the average time spent per intercourse. Though this particular assumption seems reasonable, we do not have access to data that could back up this assumption. Another example for which we do have data is livestock-exchange, which is characterized jointly by the frictional and the directional constraints.

Since constraints are related by multiple trade-offs, their epidemiological implications appear all but trivial and we study them using a bottom-up approach. As a first step, part II concentrates on the study of unique or pairs of constraints contributing to shape disease dynamics (part II). As a second step, part III implies a more general standpoint where all possible constraints are integrated in a general framework centered on agents's adaptive behaviour.

2.2 Accounting for adaptive host behaviour

Assuming intelligent agents are passive when facing crises such as epidemics appears unrealistic. Typical responses include the use of condoms, voluntary vaccinations or restrictions on trade exchanges with areas at risk of transmitting infections. Not only agents do protect themselves, but they also tend to adjust their levels of protective measures to the intensity of perceived threats [Funk *et al.*, 2010]. This particular behaviour is referred to as *adaptive risk-aversion*, because agents' decisions are motivated by fear of becoming infected (risk-aversion) and somehow depends on epidemic dynamics (adaptive behaviour). In the general case, *adaptive behaviour* means agents will adjust their behaviour in response to disruptions including disease outbreaks, and potentially their exposure to risk. Such adjustments may be caused by complex social or economic mechanisms [Klein *et al.*, 2007; Funk *et al.*, 2014].

Here, we first concentrate on presenting adaptive risk-aversion, by far the most documented example of adaptive behaviour in the epidemiological literature. Then, we show adaptive behaviour does not restrict to risk-aversion by taking the example of markets propagating epidemics. Many markets can sustain and disperse infectious diseases through trade pathways. For instance, trade of livestock can support the transmission of Foot and Mouth Disease and other infections [Keeling & Rohani, 2008]. From the converse standpoint, epidemics can disrupt disease-conducive trade pathways and influence agents' decisions to exchange, thus creating feedbacks loops between economic and epidemiological processes. Coupling a model of adaptive market behaviour with a model of infectious disease dynamics could hence lead to interesting insights on the mutual influence of trade and disease dynamics. We discuss the potential economic epidemiological implications of such a coupling based on a toy example. Building upon this toy model, part III presents a detailed analyses of a complex economic-epidemiological framework that we analyse progressively. We finally highlight interaction constraints can be subsumed into adaptive behaviour. Part III hence accounts both for adaptive behaviours and constrained interactions.

2.2.1 Risk-aversion : a standard example of adaptive behaviour

Agents tend not to remain passive when facing an epidemiological threat, and the corresponding adaptation is frequently proportionate to the perceived level threat. Typically, an increasing number of agents will tend to take protective measures such as vaccination to reduce their risk of acquiring infection in response to larger epidemic prevalence Funk *et al.* [2010].

Studies that aim at identifying effective strategies for the control of infectious diseases usually do not take into account the likely possibility that there may be an interaction or feedback loop between epidemic dynamics and host behaviour. There are some recent exceptions though, which have concentrated on the epidemiological influence of one particular adaptive response : adaptive risk-aversion behaviour [Funk *et al.*, 2009, 2010; Durham & Casman, 2012; Nicolaides *et al.*, 2013].

Risk-aversion behaviour is a form of disease prevention where asymptomatic agents decrease their exposure to infection by acting upon the two components underlying infection transmission (subsection 1.2.1) : the contact term (e.g. by staying home) and/or the per contact transmission term (e.g. by wearing protective masks). Since agents respond to a perceived threat, risk-aversion implies agents have some information about a given disease outbreak and act on their own initiative rather than relying on community-wide measures by regulatory bodies. If this behaviour is determined by the ongoing perception of a variable risk, then it is said to be 'adaptive' risk beahviour.

In practice, adaptive risk-aversion can be incorporated in epidemiological models as a function increasing with disease prevalence or outbreak awareness [Funk *et al.*, 2010], or evaluated via complex optimisation of an agent's economic weighing between the benefits of interacting with other agents and the cost of infection that may be incurred through such contacts [Fenichel *et al.*, 2011; Morin *et al.*, 2013]. All models have in common to decrease the rate of transmission in response to an increase in the perceived threat. It follows epidemiological models neglecting risk-aversion behaviour tend to overestimate the probability of occurrence and severity (e.g. infectious peak and cumulative cases) of epidemics in comparison with their counterparts accounting for risk aversion [Funk *et al.*, 2009, 2010; Fenichel *et al.*, 2011; Morin *et al.*, 2013].

2.2.2 Adaptive behaviour does not reduce to risk-aversion : the case of market behaviour

To the best of our knowledge, epidemiological modelling studies have focused on adaptive behaviour that is altered solely in response to awareness of outbreaks. However, we intuitively expect adaptive behaviour in response to outbreaks does not boil down to risk aversion. Here, we focus on a more general class of adaptive behaviour, namely *market behaviour*. We first argue market behaviour encompasses but does not restrict to adaptive risk aversion. Then we introduce a toy model coupling economic-epidemiological processes as a basis to investigate joint exchange and infection dynamics and their implications for disease control, which may be difficult to anticipate.

Behavioural responses encountered in markets

As a first approximation, a market corresponds to a complex system where economic agents freely exchange products, money and services through particular trade pathways. Trade pathways shape a contact structure among agents which has the potential to propagate infectious diseases through the exchange of contaminated products. Examples of markets propagating infectious processes include trade of livestock [Kiss *et al.*, 2006; Rautureau *et al.*, 2011; Lentz *et al.*, 2011] ; prostitution [Rocha *et al.*, 2011] ; and airline transportation [Colizza *et al.*, 2006]. Through similar mechanisms, epidemics occur through exchange of information on the Internet [Lloyd & May, 2001] and exchange of debt in financial markets [May & Arinaminpathy, 2010; Haldane & May, 2011].

Interestingly, markets involve an epidemic-conducing contact structure primarily conditioned by economic decisions very different from those underlying disease risk-aversion. Not only are markets influenced by agents's individual behaviours such as decisions to sell and buy Mas-Colell *et al.* [1995], but also by collective coordination processes such as actual exchanges between supplying and demanding agents [Fair & Jaffee, 1972] and formation of prices at an even larger scale [Mas-Colell *et al.*, 1995]. The integration of these rich behavioural processes leads to a particular balance between supply and demand, which contributes to drive trade dynamics and agents' behaviour in return, thus creating a purely economic feedback loop between agents' behaviour and market dynamics [McCauley, 2009].

When an epidemic occurs in a market, it has of course a direct influence on agents through e.g. removal and re-entry events. It can also alter agents' behaviours through adaptive risk aversion. However, the influence of disease dynamics on the complex mechanisms underpinning the feedback loop between agents' behaviour and market dynamics is not so trivial. For instance, actions and behavioural adjustments following an outbreak may either help to restore or further disturb the balance between supply and demand. Sanatory regulation and risk aversion aimed at reducing infectious contacts can reduce supply and demand. Conversely, the responses exhibited by agents can have either positive or negative impacts on disease dynamics. For example, agents that try to establish alternative but potentially-infectious trade relationships could outweigh the effect of regulation and risk aversion efforts, i.e. the effort of individual agents to adjust their own supply and demand to changing price could worsen disease outbreaks. Furthermore, the actual establishment of trade relationships underpinning the epidemiological contacts is conditioned by friction (subsection 2.1.1). By constraining the development of potentially-infectious trade contacts, friction may have a suppressive effect on epidemics. Phenomena such as friction and adjustment in supply and demand illustrate that human behaviour in response to trade-driven epidemics does not simplify to regulation and risk aversion.

A toy adaptive process coupling trade and epidemic dynamics

Clearly, the complexity in behavioural processes underpinning the mutual influence of trade and epidemics prompt for an integrative framework coupling economic and infection mechanisms. Here we present a toy adaptive process coupling trade and epidemic dynamics, where agents adapt their current and future decisions to exchange in response to ongoing or past outbreaks.

We consider a stochastic SIRS infectious process propagated on a toy market composed of four adaptive agents (Fig 2.4) [1]. Starting with a unique case (time t_0), trade connections will promote infection of other susceptible agents (t_1). Infected agents will be detected and removed (t_2 and t_4) and eventually allowed to re-enter the market after sanitation (t_3 and t_6). Removal of agents will generate strong incentives for the remaining agents to re-allocate their trade efforts towards asymptomatic

1. The toy process presented here illustrates potential outcomes of our heterogeneous market and market-epidemiological models (hFTM and hME models) with adaptive partitioning (see sections 7.1.3 and 7.2)

agents (t_5). If remaining agents are susceptible or already infected, this re-adjustment process will contribute to eradicate the outbreak, as shown here. But if a non-negligible subset of remaining agents are cryptically infected (e.g. if either agent c or d was infected at time t_5), re-adjustment in relative trade flow may actually contribute to disease progression. In addition, we notice adaptation of agents in response to epidemics (t_5) exhibits hysteresis because current decisions of agents to interact depends on the past states of the system. Hysteresis stems from agents' memories on past events and implies that the network is unlikely to return to its original state when the outbreak eventually dies-out (t_∞ versus t_0).

This simple example enables us to reinforce the intuition that adaptive market behaviour can have counter-intuitive effects on disease dynamics, which differ from adaptive risk aversion. To ease comprehension, the toy process is presented as a dynamical network, but adaptive behaviour can also be investigated with other classes of models, such as the compartmental framework. In this work, we resort to a combination of network-based and compartmental approaches to study the joint- influence of trade and epidemic dynamics (part III).

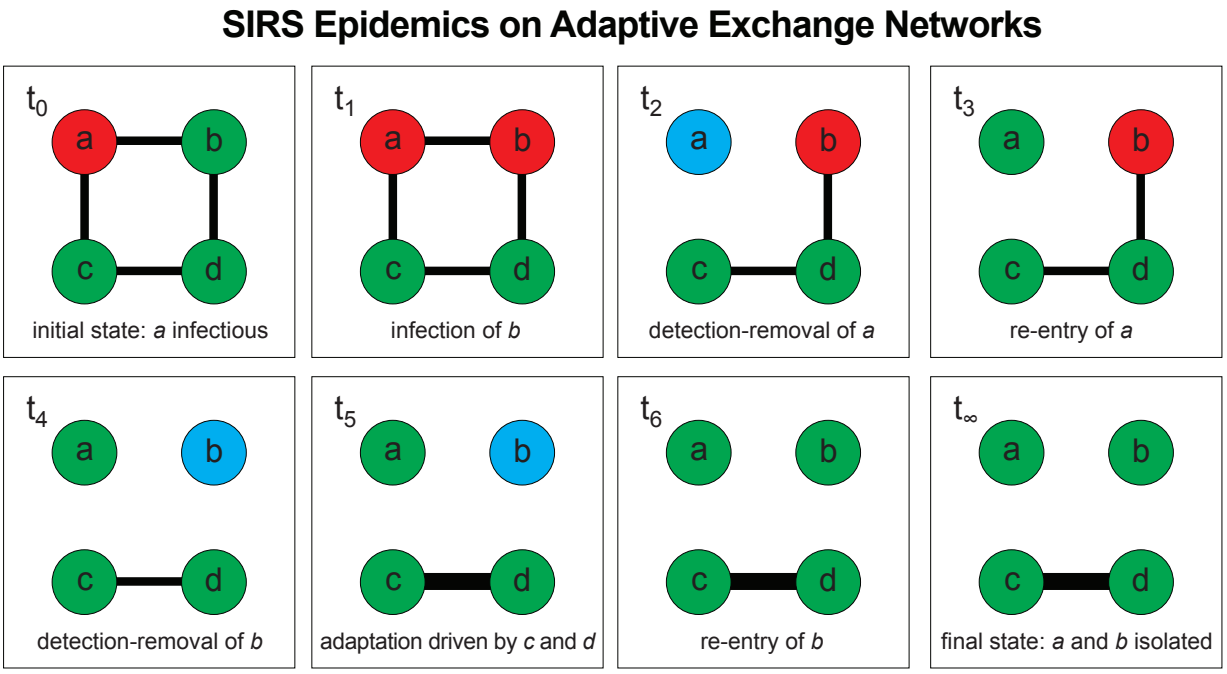


FIGURE 2.4 – **Adaptive host behaviour influence joint exchange and epidemic dynamics.** We consider an infectious process propagated on an adaptive exchange network made of four agents a, b, c and d. The eight panels (t_0 – t_∞) depicts the full time course of an SIRS outbreak (S status in green, I in red, R in blue), starting with the initial infection of a at time t_0, up to the system final state t_6 – t_∞.

2.2.3 Interaction constraints can be subsumed into adaptive host behaviour

Our intuition, that we translate into integrative models of trade and disease dynamics (part III), is that an agent's adaptive decision to exchange necessarily integrates all possible constraints it faces (section 2.1). From this standpoint, intricate trade-off in constraints arises because any agent's willingness to exchange must be bounded by some kind of limited exchange budget, which corresponds to an agent' physical limitations. It follows any agent has to constantly arbitrate between various conflicting constraints to reach a given exchange objective. For instance, an agent may decide to spend its available budget by searching for new partners (i.e. depart from the sparseness constraint). But once this budget is consumed, the agent will not be able to increase its overall interaction weight per link (i.e. decrease the burden of the weighting constraint), or increases its relative frequency per interaction (i.e. relax the frictional constraint).

2.3 Approaches used to relax assumptions of simple SIRS models

Simple SIRS models are grounded on core assumptions governing exchanges and subsequent epidemic dynamics, most importantly homogeneous-mixing and passive host behaviour. Such mechanisms are clearly at odds with empirical evidence documenting exchange structures [Keeling & Eames, 2005] and exchange dynamics [Funk *et al.*, 2010]. This advocates for the inclusion of interactions constraints (section 2.1) and adaptive host behaviour (section 2.2) into epidemiological models

To relax the above assumptions and allow for more general exchange mechanisms such as adaptive behaviour with constrained interactions, we present two complementary approaches widely used in the epidemiological literature : one approach generalises the network-based standpoint, while the other builds upon the compartmental framework. In other words, both the compartmental standpoint and its network-based counterpart constitute flexible frameworks that can be extended to overcome the limitations of simple SIRS models and improve the description of underlying exchange-driven dynamics. While more precise than a compartmental model, a network-based framework requires detailed data that are not always available or only partially known. Besides, network-based models are usually complex and not amenable to detailed mathematical derivations and quick computational analyses. It follows the compartmental approach is complementary to the network-based standpoint. Both approaches are valuable, and the best standpoint depends on research objectives and available data. For instance, in this work, we exclusively adopt the network-based standpoint to study interaction constraints with passive host behaviour and their implications on epidemics (part II). However, when we do take into account adaptive host behaviour, which incorporates constraints plus other complex intricate processes, we resort to both approaches (part III).

2.3.1 Generalising the network-based standpoint

Networks constitute natural tools to model interaction constraints. Here we first explain how complex networks and subsequent infection processes can be formally described. Then, we apply this generalised approach to incorporate and explore the epidemiological influence of the sparseness constraint, which has been largely studied in the epidemiological literature (subsection 2.1.1)

The general approach

We start by describing epidemiological contact structures as weighted directed networks. In the general case, a network can be represented by an $N \times N$ matrix $\mathbb{W}$, whose element $w_{ij} \leq 0$ represents the weight at which agent i interact with agent j, following the order of the directed link (i, j). As particular cases, a network is *unweighted* if w_{ij} only takes binary values, where 1 denotes a link and 0 an absence of link ; and a network is *undirected* if $w_{ij} = w_{ji}$ for all links. The complete network corresponds to the case where $w_{ij} = 1$ for any couple of agents (i, j), except on the diagonal $(i = j)$,

which is set to zero. By allowing the modeller to specify *any* value or function for w_{ij}, i.e. by departing from the complete network and homogeneous-mixing, the network-based standpoint can overcome the limitations of simple SIRS models [Keeling & Eames, 2005].

Now we specify an infection process adapted to weighted directed networks. In the formulation of the SIRS complete-network-based model (Fig. 1.2), we notice the local force of infection λ_i^c for an agent i (1.3) explicitly depends on its number of infected neighbours. In the general case, the local force of infection, that we denote λ_i, becomes :

$$\begin{aligned} [\text{local force of infection}] &= [\text{rate of infection per contact}] \cdot [\text{sum of the directed weights } (j \rightarrow i) \text{ over all agents } j] \\ \lambda_i^c &= \omega \sum_{I_j=1} w_{ji} \ . \end{aligned} \tag{2.1}$$

The rest of the model remains identical to its complete-network-based counterpart.

Example of application : epidemics on sparse networks

Understanding how interaction constraints influence epidemics is one of the central question investigated in our work (part II). Here, we show, using the well-studied spareness constraint (subsection 2.1.1) as an example, how constraints can be integrated in a network-based framework and how they can influence epidemics.

Generating sparse networks. We place ourselves in a case where the overall level of sparseness is kept constant, i.e. we set the number of agents N and average number of links per agents $\langle k \rangle$. We consider two contrasted models for generating sparse networks : the model of Erdős–Rényi (ER)[Albert & Barabasi, 2002] and the model of Barabási–Albert (BA) [Barabasi & Albert, 1999]. In a nutshell, the ER model draws new links uniformly at random and independently among N pre-existing agents ; while the BA model connects new agents and corresponding links proportionally to the degree of existing agents. In both cases, the network generation process is stopped when a network with N agents and $N \langle k \rangle /2$ links is obtained. The ER and BA models yield sparse networks with Poisson (i.e. homogeneous) and power-law (i.e. heterogeneous) degree distributions respectively (Fig. 2.5A-B). In practice, empirical networks exhibit a large variance in degree and are hence usually best described by BA rather than ER networks [Barabasi & Albert, 1999].

Influence of differing sparse networks on the invasion threshold. ER and BA networks appear very different, which may in turn influence epidemics. Here, following the proofs detailed in [Aparicio & Pascual, 2007; Durrett, 2007], we generalise the original definition of the epidemiological threshold R_0 to any sparse network, and show the types of modifications brought by local heterogeneity in the sparseness constraint.

Whenever sparse networks exhibit heterogeneity in degree, the usual definition of R_0 should be slightly modified to account for subtle correlative effects between infectious agents arising in the early stage of an outbreak. The usual approach consists in defining R_0 as the average number of susceptibles infected by a *typically* (rather than an average) infected agent over its entire infectious period. In practice, a typically infected agent is defined as a secondary case, i.e. an agent infected by the very first infectious case. R_0 hence becomes the average number susceptible agents infected by a secondary case. We now derive R_0. Let τ and ζ denote the transmissibility and the average excess degree of a secondary case respectively. The transmissibility τ is the probability that an infection spread along a link before recovery, and is hence given by $\tau = \omega/(\omega+\gamma)$. The average excess degree of a secondary case ζ is the average of the number of agents that can get infected by a secondary case, i.e. their average degree-minus-one. Removing one to the degree is necesarry because a secondary case was necessarily infected by an initial case through a common link. By construction, R_0 is given by the product $\tau\zeta$. We now derive ζ to re-express R_0. We assume the initial case is selected uniformly at random among the N agents. It follows the initial case has degree k with probability P_k, where P_k is the degree distribution. Now what is the probability, denoted Q_k, that a secondary case has degree k ? Given the index case is an *agent* selected uniformly at random, Q_k is also the probability to pick an agent with degree k by following a *link* selected uniformly at random. Q_k is hence proportional to kP_k, and we have : $Q_k = kP_k/(\sum_v vP_v) = kP_k/\langle k\rangle$, where $\langle\cdot\rangle$ denotes the average operator and $\langle k\rangle$ is the average degree. ζ is hence given by : $\zeta = \sum_{k=1}^{\infty}(k-1)Q_k = \langle k^2\rangle/\langle k\rangle - 1$, where $\langle k^2\rangle$ is the second moment of the degree distribution [2]. Finally, re-expressing the basic reproduction number as function of original model parameters leads to :

$$R_0 = \frac{\omega}{\omega+\gamma}\left(\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right) , \tag{2.2}$$

We notice the first and second moments of P_k fully specify the influence of network structure on subsequent epidemic threshold properties. In contrast with degree-homogenous networks such as ER networks where $\langle k^2\rangle \approx \langle k\rangle^2$, degree-heterogeneous networks such as BA networks are characterised by $\langle k^2\rangle \gg \langle k\rangle^2$. Other being equal, the risk of epidemic invasion is hence larger in networks with heterogeneous degree distributions. Note that key relation (2.2) only stands for static networks that are both unweighted and undirected [3].

Influence on epidemic dynamics. We now simulate multiple SIR trajectories for one ER and on BA network with $N = 10000$ and $\langle k\rangle = 4$ (Fig. 2.5C-D). We compare two epidemic scenarios. On the one hand, we set $\omega = 0.2$ for both networks (**C**, in grey and black for ER, and cyan and red for BA). On the other hand, we take for each network the value of ω for which $R_0 = 4$ (**D**, identical color code). The

2. Note that we implicitly assume that the first moment $\langle k\rangle$ and the second moment $\langle k^2\rangle$ exist, which is always true for networks with a finite N

3. see chapters 3 and 4 for attempts to generalise (2.2) to weighted networks or directed networks respectively.

scenario where ω is kept constant suggests the BA network appears more at risk than the ER network, with a quicker disease propagation and a higher infectious peak. This finding is not surprising given $R_0 = 12$ for the BA network, while $R_0 = 4$ for the ER network. In the second scenario, we hence adjust the value of ω for the BA network so that $R_0 = 4$ for both networks. The results are reversed : the ER network appears now far more at risk than the BA network. In contrast with simple SIR models which always exhibit the same type of epidemic response, BA and ER networks exhibit drastically differing behaviour when facing an epidemiological process characterised by a given R_0 and infectious period $1/\gamma$.

Consequences for disease control. As simple as it seems, the sparseness constraint can have drastic consequences on subsequent infection processes, especially when sparse networks are heterogeneous in degree, i.e. networks largely departing from homogeneous-mixing. In such cases, network epidemiology can help in determining which agents should be preferentially targeted for control [Eames & Keeling, 2003]. In practice, targeting hubs preferentially is frequently the most effective control strategies to destroy the epidemiological contact structure of sparse networks with heterogeneous degree, and hence block potential transmission events [Albert *et al.*, 2000].

2.3.2 Specifying the sub-components of the compartmental framework

The SIRS compartmental model can be extended as such, retaining the compartmental formulation without resorting to networks. The approach consists in two main steps. First, the rate of transmission β (1.1), which essentially acts a as black box, is broken down as a product of two terms : a *contact term*, describing the epidemiological contact structure, and a *per contact transmission term*, specifying the infection process [Keeling & Rohani, 2008]. For β to be physically homogeneous to a rate, either the contact term or alternatively the per contact transmission term is required to be a rate. Second, the SIRS compartmental model can be enriched by explicitly specifying the mechanistic processes underlying the contact term and/or the per contact transmission term [e.g. Fenichel *et al.*, 2011].

2.3.3 Choosing the most parsimonious approach

At first sight, it seems difficult to know whether we should adopt preferentially the compartmental or the network-based standpoint. This is actually a general problem for modellers, because there exist a plethora of epidemiological frameworks that extend far beyond the particular compartmental and network-based standpoints we present [Keeling & Rohani, 2008], for instance spatial models with dispersion kernels [Keeling *et al.*, 2001].

Here we provide the general criteria used to favour one approach against the other. The general idea consists in comparing models based on the *objectives* they can address, relatively to their *levels of complexity*. We do not plan to be exhaustive here, only clarify the scope and limitations of any

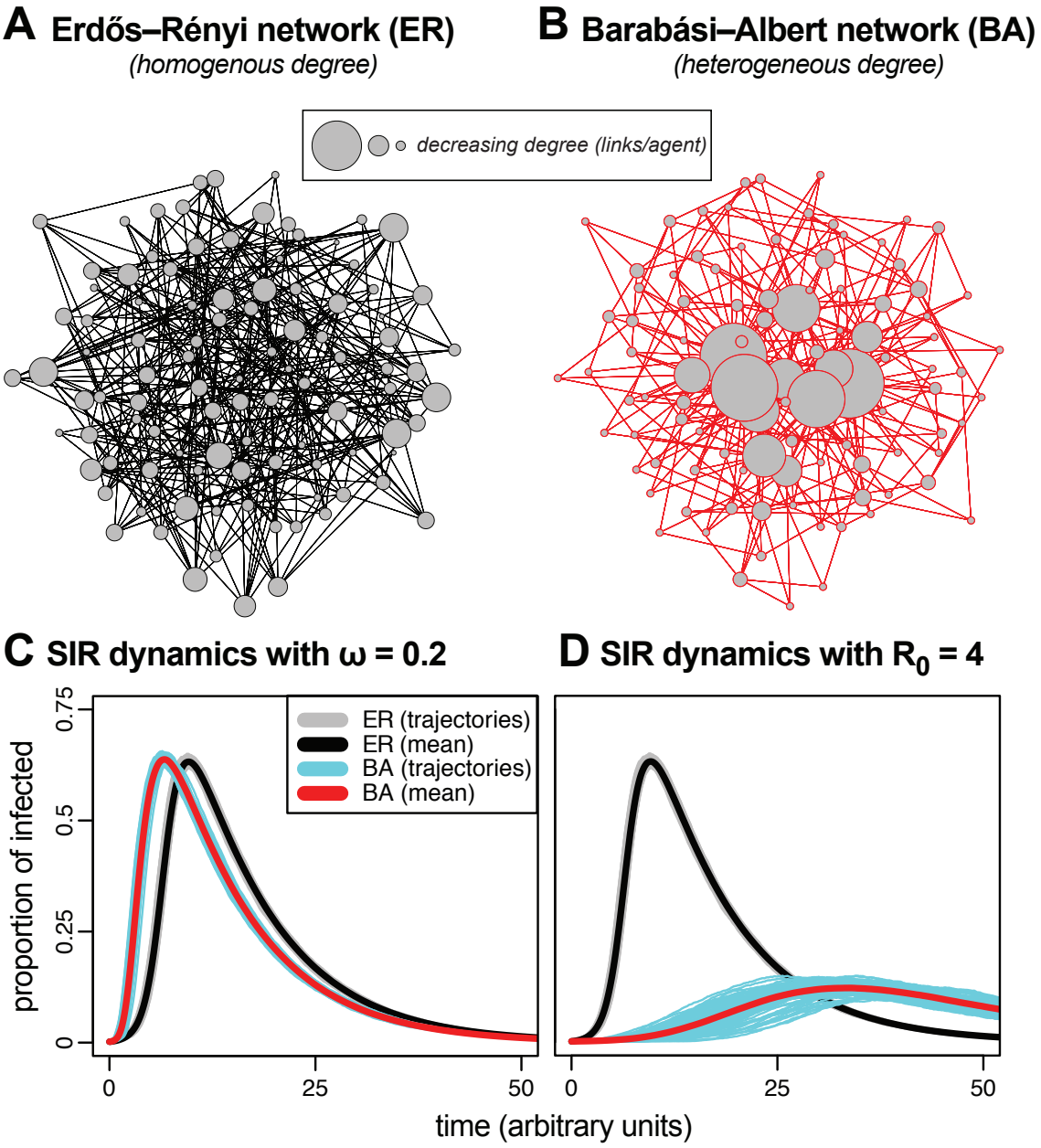


FIGURE 2.5 – **Epidemics on sparse networks with contrasted degree distributions.** Influence of networks with contrasted degree distributions (**A**-**B**) on subsequent SIR epidemics (**C**-**D**). The epidemic process we consider, including its notations and implementation, is essentially identical to Fig. 1.2, with the exception that susceptible agents now become infected at rate λ_i (2.1). The basic reproduction number becomes $R_0 = \omega/(\omega+\gamma)(\langle k^2\rangle/\langle k\rangle - 1)$, where $\langle k\rangle$ and $\langle k^2\rangle$ respectively denote the first and second moments of the degree distribution. Two types of standard sparse networks are contrasted : the Erdős–Rényi network (ER, **A**) and the and Barabási–Albert network (BA, **B**), with resulting homogeneous and heterogeneous degree distributions respectively. For graphical convenience, we display networks with $N = 100$ agents and $\langle k\rangle = 4$. Then we simulate stochastic SIR dynamics on a single ER network and a single BA network, with $N = 10000$ and $\langle k\rangle = 4$. For both cases, we initially infect 20 agents and set $\gamma = 0.1$ and $\nu = 0$. We compare two epidemic scenarios. On the one hand, we set $\omega = 0.2$ for both networks (**C**, in grey and black for ER, and cyan and red for BA). On the other hand, we take for each network the value of ω for which $R_0 = 4$ (**D**, identical color code). Rates are expressed in arbitrary units.

modelling approach. First, we need to make sure the modelling framework we adopt has the potential to address our particular research objectives. A model is always designed to answer a precise question.

Some models may uncover the most important mechanisms leading to disease transmission, predict the evolution of an ongoing outbreak or evaluate the particular agents that are most at risk. Others may inform us on how best collecting data in the field to provide valuable insights on disease dynamics. Second, we will tend to favour models with appropriate levels of complexity to address a given research question. For instance, scales encompassed by models may include intra-host dynamics Mideo *et al.* [2008] and nested layers of inter-host dynamics, ranging from interacting individuals to entire countries and even continents Moslonka-Lefebvre *et al.* [2011]. Some models consider epidemics as deterministic processes, while others describe them as stochastic Keeling & Rohani [2008]. Choosing the most parsimonious level of complexity while still successfully reaching a given objective, is related with Richard Levins's modelling trade-off Orzack & Sober [1993]. For Levins, any model is expected to exhibit a trade off between generality (in the sense of flexibility, rather than mathematical generality), realism and precision.

We can now apply Levins' trade-off to our particular problem. When compared against a SIRS network-based framework, a SIRS compartmental model is more general (can be easily tuned to explain an average epidemic behaviour) but less realistic (the underlying contact structure implies homogeneous-mixing and subsequent epidemic events are deterministic) and more imprecise (we cannot access individual infectious status, only quantify the fractions of hosts belonging to particular compartments, i.e. subpopulations). The network-based model is intrinsically specific and data-costly as it requires the specification of an explicit contact structure, for which we do not necessarily have data. It follows the choice of a given standpoint depends on the problem we wish to tackle and the data we have access to.

In practice, modellers can sometimes overcome Levins' trade-off and lack of detailed data by e.g. approximating epidemiological processes occurring on complex models by simpler models [Pellis *et al.*, 2014]. Simpler does not necessarily mean over-simplistic, and an approximated model can sometimes retain the most important properties of its original model, while enabling the use of data available at coarser scales (we illustrate this point in chapter 3).

Deuxième partie

A path from epidemic models with interaction constraints, but passive agents...

Foreword to part II

Exploring the influence of differing network structures on overlying epidemic dynamics constitutes a long-standing challenge of network epidemiology [Strogatz, 2001]. Epidemiological contact structures spanned by exchanges appear extremely rich due to a combination of intricate interaction constraints faced by agents (section 2.1), which suggests interaction constraints are likely to exert a leading influence on epidemics. Interaction constraints arise because agents' exchange capacities are bounded in the number of partners they can have (sparseness constraint), rate and direction of exchanges (weighting and directional constraints) or in their rate of encounter with other agents (frictional constraint). While the sparseness constraint has been exhaustively studied, the epidemiological influence of the three other constraints and their interactions remains poorly understood [Meyers *et al.*, 2006; Pellis *et al.*, 2014].

In this part, we thoroughly investigate how interaction constraints shape epidemics for an overall exchange level kept constant. We place ourselves in a case where exchange networks are not affected by agents' behaviours, a simplifying assumption enabling us to gain detailed mathematical insights on the structural factors driving infection dynamics. Identifying the key structural factors determining infection transmission is central to design relevant disease control policies. We start by exploring the epidemiological implications of the weighting constraint, which reveals crucial to understand STI dynamics (chapter 3). Then we study the influence of the directional constraint on epidemics, with a focus on trade-driven epidemics in ornamental plants (chapter 4). Finally, we show combining the weighting and directionality constraints can prove insightful to assess and prevent joint economic-epidemiological risk in livestock-exchange markets (chapter 5). Note that the frictional constraint is rather investigated when accounting for agents' adaptive behaviour (part III), because its adequate study requires an explicit modelling of exchange dynamics. As a general rule, we detail novel approaches and insights gained from projects entirely initiated as parts of this doctoral work. In contrast, we only briefly present in the main text research projects that were completed, but not started, during this PhD project. All research chapters included have in common to shed light on exchange-driven epidemics, and constitute ports of call of the same intellectual odyssey.

Chapitre 3

Weighting matters for epidemics. Application to STI dynamics on networks.

This chapter presents the epidemiological implications of the weighting constraint and is based on two research articles published in *Journal of Theoretical Biology* [Moslonka-Lefebvre *et al.*, 2012a] and *PLoS Computational Biology* [Kamp *et al.*, 2013]. Note that we only detail [Kamp *et al.*, 2013] in the following, since this particular study was entirely initiated as part of this doctoral work. We also emphasise the key mathematical concepts and derivations of this chapter were contributed by Christel Kamp. Our main contributions lies in the epidemiological motivation of the study, survey data analysis (see also Fig. 2.2), comparisons of contrasted approaches to generate actual weighted networks from joint probability distributions in degree and weight, and numerical explorations of subsequent disease dynamics to confirm analytical predictions.

In [Moslonka-Lefebvre *et al.*, 2012a] (see appendix A), we show weighting for sex acts is crucial to understand the spread of STI on networks. Human sexual networks exhibit a heterogeneous structure where few individuals have many partners and many individuals have few partners. Network theory predicts that the spread of sexually transmitted infections (STI) on such networks should exhibit striking properties (e.g. rapid spread or vanishing epidemic threshold). However, these properties cannot be found in epidemiological data. Current network models typically assume a constant STI transmission risk per partnership, which is unrealistic because it implies that sexual activity is proportional to the number of partners and that individuals have the same activity with each partner. We develop a framework that allows us to weight any sexual network based on biological assumptions. Our results indicate that STI spreading on the resulting weighted networks do not have heterogeneous-related properties, which is consistent with data and earlier studies. Generally speaking, our study suggests altering the weight-per-agent distribution on a given network without modifying its cumulative weight can have major consequences for epidemic dynamics.

In [Kamp *et al.*, 2013], we confirm analytically the numerical results reported in [Moslonka-Lefebvre

et al., 2012a] and propose novel mathematical tools to approximate SIR epidemics on any weighted network. In a nutshell, we follow a configuration-type approach, which implies a weighted network is fully characterised by its joint probability distribution P_{kl} in degree k (e.g. number of sexual partners par individual) and weight l (e.g. number of sex acts per individual per time unit). We can then capture the full time course of SIR epidemics by solving a system of partial differential equations of probability generating functions derived from P_{kl}. This new approach is hence applicable to survey data which give us access to P_{kl} and have the advantage of being easily accessible in contrast with sexual-contact networks obtained by contact-tracing studies (section 2.3). In addition, the framework allows us to calculate analytical estimates for the basic reproduction number (R_0) and the rate of early epidemic growth (r_0). In, particular, we find R_0 is approximated by :

$$R_0 \approx \frac{\pi \langle l/k \rangle}{\pi \langle l/k \rangle + \gamma} \left(\frac{\langle kl \rangle}{\langle l \rangle} - 1 \right) , \tag{3.1}$$

where, π is the probability of transmission per interaction event, $\langle l/k \rangle$ is the average number of interaction events per contact per time unit, γ the recovery rate, $\langle l \rangle$ the average number of interaction events per agent per time unit and $\langle kl \rangle$ is the second moment of the joint probability distribution P_{kl}. By remarking the per-contact transmission rate ω (see subsection 1.2.1) is given by $\omega = \pi \langle l/k \rangle$, we can rewrite R_0 as :

$$R_0 \approx \frac{\omega}{\omega + \gamma} \left(\frac{\langle kl \rangle}{\langle l \rangle} - 1 \right) . \tag{3.2}$$

If weights per agent are constant ($l \propto 1$ in (3.2)), as observed in first approximation for sexual-contact networks (Fig. 2.2), R_0 reads :

$$R_0 \approx \frac{\omega}{\omega + \gamma} \left(\langle k \rangle - 1 \right) \leq \frac{\omega}{\omega + \gamma} \left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right) , \tag{3.3}$$

where the right-hand side is the value of R_0 when the network is unweighted ($l \propto k$ in (3.2), which is identical to (2.2)). Note the inequality between the left and right hand sides holds since $\langle k \rangle^2 \leq \langle k^2 \rangle$, with equality only when the variance of k, given by $\langle k^2 \rangle - \langle k \rangle^2$, is negligible, i.e. when networks are strictly regular. Relation (3.3) confirms analytically the results reported in [Moslonka-Lefebvre *et al.*, 2012a], and demonstrates the importance of considering concurrently the sparseness and the weighting constraints.

Taken together, [Moslonka-Lefebvre *et al.*, 2012a; Kamp *et al.*, 2013] suggest targeting prevention measures on the most connected individuals may not be the best strategy to prevent the spread of STI. Rather, general prevention campaigns (e.g. TV spots to promote the use of condoms) may prove more cost-effective in such systems.

Epidemic spread on weighted networks.

Christel Kamp[1,*], Mathieu Moslonka-Lefebvre[2,3] and Samuel Alizon[4]

[1] Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Paul-Ehrlich-Straße 51-59, 63225 Langen, Germany.

[2] INRA, UR 0341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France.

[3] AgroParisTech, F-75005 Paris, France.

[4] Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2), 911 avenue Agropolis, 34394 Montpellier Cedex 5, France.

[*] Corresponding author, email : christel.kamp@pei.de

Abstract

The contact structure between hosts shapes disease spread. Most network-based models used in epidemiology tend to ignore heterogeneity in the weighting of contacts between two individuals. However, this assumption is known to be at odds with the data for many networks (e.g. sexual contact networks) and to have a critical influence on epidemics' behavior. One of the reasons why models usually ignore heterogeneity in transmission is that we currently lack tools to analyze weighted networks, such that most studies rely on numerical simulations. Here, we present a novel framework to estimate key epidemiological variables, such as the rate of early epidemic expansion (r_0) and the basic reproductive ratio (R_0), from joint probability distributions of number of partners (contacts) and number of interaction events through which contacts are weighted. These distributions are much easier to infer than the exact shape of the network, which makes the approach widely applicable. The framework also allows for a derivation of the full time course of epidemic prevalence and contact behaviour, which we validate with numerical simulations on networks. Overall, incorporating more realistic contact networks into epidemiological models can improve our understanding of the emergence and spread of infectious diseases.

Keywords : weighted networks, epidemics, SIR model, R_0, outbreaks

Authors summary

Understanding how infectious diseases spread has public health and ecological implications. The contact structure between hosts strongly affects this spread. However, most studies assume that all type of contacts are identical, when in reality some individuals interact more strongly than others. This is particularly striking for sexual-contact networks, where the number of sex acts is not identical for all partnerships. There are two limitations to current frameworks that can explain the lack of studies on weighted networks. First, analytical results are difficult to obtain, which requires numerical simulations. Second, inferring weighted networks from survey data is extremely difficult. Here, we present a novel framework that allows to alleviate these two limitations. Building on configuration type network epidemic approaches, we manage to capture disease spread on weighted networks from the distribution of the number of contacts and distribution of the number of interaction events (e.g. sex acts). This allows us to derive analytical estimates for the epidemic threshold and the rate of spread of the disease. It also allows us to readily incorporate survey data, as illustrated in this study with data from the National Survey of Sexual Attitudes and Lifestyles (NATSAL) carried out in the UK.

3.1 Introduction

Contact structure between hosts is known to have a key influence on disease spread [Anderson & May, 1991]. A striking result is for instance that the more heterogeneous the contact network is, i.e. the higher the variance in the number of contacts per individual, the more rapid the initial disease spread.

One way to capture contact structure is to use a network [Newman, 2002]. Such contact networks are typically described by a square binary adjacency matrix, where each term on the ith line and jth column can take the value 0 or 1 to indicate respectively the absence or the presence of a contact between individuals i and j. Contact networks are widely used because they possess several convenient properties, one of which being that the dominant eigenvalue of the adjacency matrix is an indicator of the initial propagation speed of an infectious disease spreading on this network [Li *et al.*, 2012; Moslonka-Lefebvre *et al.*, 2012b].

The main limitation of contact networks is that their exact shape is often difficult to infer. This is why there is a continuous effort to predict disease spread from network summary statistics that are easier to estimate, such as the distribution of the number of contacts (degrees). For instance, the number of secondary infections generated by a typical infected host in a fully susceptible population, i.e. the basic reproductive number R_0 [Anderson & May, 1991], scales with the ratio of the second moment $\langle k^2 \rangle$ and first moment (mean) $\langle k \rangle$ of the distribution in the number of contacts k. This result holds both for static networks (denoted R_0^{stat})[Durrett, 2007] as well as for fully mixed, dynamic networks

(denoted R_0^{mix}) [May & Anderson, 1987; Moreno *et al.*, 2002] with

$$R_0^{\text{mix}} = \frac{\omega}{\gamma}\frac{\langle k^2\rangle}{\langle k\rangle} = \frac{\omega}{\gamma}\left(\frac{\sigma_k^2}{\langle k\rangle} + \langle k\rangle\right) \tag{3.4a}$$

$$R_0^{\text{stat}} = \frac{\omega}{\omega+\gamma}\left(\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right) = \frac{\omega}{\omega+\gamma}\left(\frac{\sigma_k^2}{\langle k\rangle} + \langle k\rangle - 1\right) \tag{3.4b}$$

where $\sigma_k^2 = \langle k^2\rangle - \langle k\rangle^2$ is the variance of the distribution of the number of contacts. The static case corresponds to networks in which the identity of contacts is fixed (as approximatively seen in sexual contact networks) and the fully mixed dynamic case corresponds to a situation in which individuals update their contacts dynamically in a fully mixed fashion within the population (as approximatively seen in airborne infections).

R_0^{stat} and R_0^{mix} represent the lower and upper bounds of the basic reproductive ratio [Volz & Meyers, 2009] for SIR epidemics on random networks if infectious individuals transmit the infection at the per-contact transmission rate ω and recover from the infection at a rate γ. These findings have implications on the epidemic threshold which is inversely proportional to R_0. Both static and dynamic heterogeneous networks with a large or even diverging variance in the distribution of the number of contacts have a very small or even vanishing epidemic threshold.

One of the typical key assumptions epidemiological models on networks make to obtain such elegant expressions for R_0 is that the transmission rate is the same between all pairs of individuals. This is materialized by the fact that all the edges of the contact matrix have a weight of 0 or 1. This is known to be a simplifying assumption [Barrat *et al.*, 2004a]. A well-studied example related to infectious diseases is that of sexual contact networks, where the number of sex acts per unit of time is not constant in all partnerships [Blower & Boe, 1993; Nordvik & Liljeros, 2006; Britton *et al.*, 2007]. More generally, the number of interaction events (which correspond to potential transmission events) may vary among contact pairs and is likely to decrease with the number of contacts an individual has (see also Fig. B.3). Simplifying the reality is commendable but the problem is that tampering with the weighting of the network has been shown to lead to the loss of important epidemiological properties of heterogeneous unweighted networks, such as the low value of the epidemiological threshold or the negative correlation between the epidemiological threshold value and network size [Moslonka-Lefebvre *et al.*, 2012a]. To summarize, although contact networks appear to be 'scale free' in structure, they might not exhibit the properties one might expect from this structure.

An increasing number of studies in epidemiology point to the importance of considering weighted networks. Some examples include the spread of sexually-transmitted infections [Moslonka-Lefebvre *et al.*, 2012a], disease transmission in sheep flocks [Schley *et al.*, 2012], respiratory diseases of humans [Stehle *et al.*, 2011] or general infectious diseases of human spreading on a social contact network [Eames *et al.*, 2009] or on airline connection networks [Colizza *et al.*, 2006]. Several more conceptual studies have also been published in the theoretical physics literature (e.g. [Newman, 2002; Joo & Lebowitz, 2004;

Wang *et al.*, 2007; Britton *et al.*, 2011]). Most of these studies have in common that they use weighted networks and resort to (heavy) numerical simulations. Indeed, contrary to unweighted networks, we lack analytical frameworks to analyze epidemic spread on weighted networks.

Here, we present an original framework, which builds on configuration type network epidemic approaches [Volz, 2008; Kamp, 2010] that offers an alternative to simulating epidemics on full networks. It allows to model the dynamics of a disease spreading on a weighted network and to estimate key epidemiological variables from the network's properties. The framework provides us with explicit expressions for the rate of early epidemic expansion (r_0) and the basic reproductive ratio (R_0) of the infection without requiring strong simplifying assumptions regarding epidemiological processes or the distribution of weights on the contact network. It also allows for a derivation of the full time course of epidemic prevalence and contact behaviour of susceptible, infected and recovered individuals (in terms of the probability generating functions – PGFs – of the degree distributions). As sketched in Fig. 3.1, the parametrisation is done in a general fashion using the joint probability distribution P_{kl} of an individual to have k contacts among which (s)he randomly distributes l interaction events. We validate our analytical results using numerical simulations on networks.

Importantly, since this framework relies on summary statistics of the network and does not require knowledge of the exact shape of the network, it can be parametrized using large scale survey data. The network information we lose by using these summary statistics requires that we make the assumption that there is no assortativity between individuals in our framework. However, we show that even with these assumption we can approximate epidemic dynamics better than with non weighted networks.

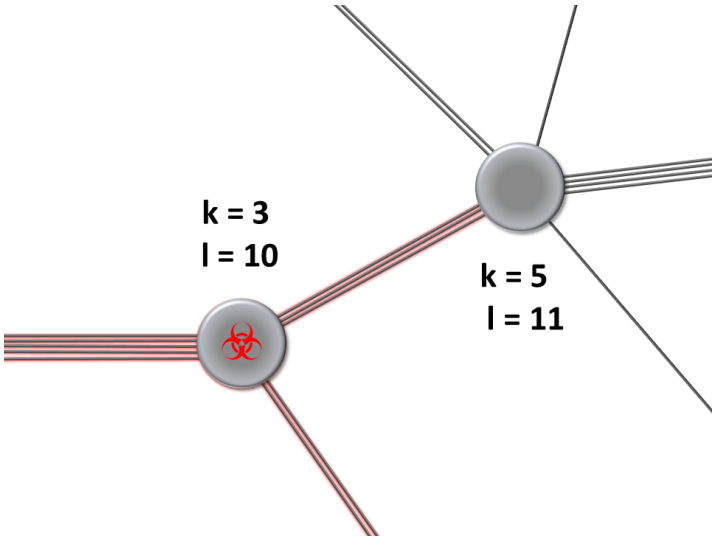


FIGURE 3.1 – **Weighting between contacts.** Both the number of contacts that an individual maintains and the weight that (s)he assigns to each contact are relevant for the spread of an infectious agent. Here, each individual has l interaction events that (s)he can distribute among his/her k contacts. On the scale of the transmission network, these are modelled by the joint probability distribution P_{kl} to find an individual with k contacts and l interaction events per time interval.

3.2 Materials and Methods

3.2.1 The configuration model

Individual have k contacts and l interaction events per time interval, which are distributed among their k contacts (l is sometimes also referred to as the strength of the node [Barrat *et al.*, 2004a]).

The model is broadly applicable as it can be parameterised through any joint probability distribution of the number of partners (k) and number of interaction events (l). Such a joint probability distribution P_{kl} can be written as the product $P_{kl} = P_k \; P_{l|k}$, $P_{l|k}$ being the probability distribution of the number of interaction events per time l given that the individual has k contacts. If $P_{l|k} = \delta_{lk}$, where δ_{lk} is 1 if $l = k$ and 0 otherwise, we are then back to a 'classical' network case, with an exact linear dependency between the number of contacts and the number of interaction events (for a detailed discussion, see section B.4). Our framework can capture more general situations by explicitly choosing $P_{l|k}$.

Our analytical approximation assumes that an individual distributes his/her l interaction events multinomially among his/her k contacts, which are static and are randomly assigned as in configuration models [Volz, 2008; Kamp, 2010]. This individual is infected at a rate proportional to his/her average number of interaction events with i infected contacts among his/her total k contacts (so $0 \leq i \leq k$). This averaging implies the choice of a time scale for the number of interaction events such that $\langle l \rangle > \langle k \rangle$ (the network shape is assumed to be constant over the time period considered so the number of contacts is not affected by the time scale). The analytical approach only relies on the nodes' statistics and does not consider the constraints for half-contacts to match half-contacts with similar weight, whereas in reality, the weight on a link between two nodes should be the same for the two nodes. This could lead to an unrealistic network segregation for some artificial networks where $\langle l \rangle > \langle k \rangle$. This network segregation can affect epidemic dynamics in these networks in a way that is not seen in the analytical approach *per se* but can be considered through corrections in the analytical approach as we discuss later on.

3.2.2 The epidemiological model

A susceptible individual becomes infected at a rate proportional to the number of interaction events per time interval (l), the transmission probability per interaction event of the pathogen (π) and the probability for each of his/her contacts to points to an infected individual (p_{SI}). The population dynamics of susceptible, infected and recovered individuals with k contacts and l interaction events per time interval (denoted S_{kl}, I_{kl} and R_{kl} respectively) are thus captured by the following set of

differential equations :

$$\dot{S}_{kl} = -\pi\, p_{SI}\, l\, S_{kl} \tag{3.5a}$$

$$\dot{I}_{kl} = \pi\, p_{SI}\, l\, S_{kl} - \gamma\, I_{kl} \tag{3.5b}$$

$$\dot{R}_{kl} = \gamma\, I_{kl}, \tag{3.5c}$$

where γ is the rate at which hosts recover and become immune to subsequent infection (see Table 3.1). This setting can capture lifelong infections if $\gamma = 0$.

The dynamics of the total number of susceptible, infected and recovered individuals can be obtained by summing over k and l in equation system 3.5. This leads to :

$$\dot{S} = -\pi\, p_{SI}\, \langle l \rangle_S\, S \tag{3.6a}$$

$$\dot{I} = \pi\, p_{SI}\, \langle l \rangle_S\, S - \gamma\, I \tag{3.6b}$$

$$\dot{R} = \gamma\, I. \tag{3.6c}$$

where $\langle l \rangle_S = \sum_{k,l} l\, S_{kl}/S$ is the average number of interaction events per time a susceptible individual has.

To close the equation system 3.5–3.6, we need expressions for the temporal dynamics of the p_{AB}, i.e. the probabilities for a status A individual's contact to be with an individual in state B. These can be derived through a careful assessment of the links/contacts among susceptible and infected individuals over the course of an epidemic. This means following the dynamics of the joint probability distribution to find k contacts and l interaction events per time among susceptible individuals $P_{Skl} = S_{kl}/S$ through its PGF, denoted $G_S(x, y, t)$. The temporal dynamics of $G_S(x, y, t)$ can be calculated by observing that the dynamics of the corresponding joint probability distribution of contacts and interaction events of susceptible individuals are governed by the equation $\dot{P}_{Skl} = \frac{\dot{S}_{kl}}{S} - \frac{\dot{S}}{S} P_{Skl}$. Hence, we close equation system 3.5–3.6 with the following equations :

$$\dot{p}_{SI} = p_{SI} \left(\pi\, p_{SS}\, \frac{\langle kl \rangle_S}{\langle k \rangle_S} - \pi\, (1 - p_{SI} + p_{SS})\, \frac{\langle l \rangle_S}{\langle k \rangle_S} - \gamma \right) \tag{3.7a}$$

$$\dot{p}_{SS} = -\pi\, p_{SI}\, p_{SS}\, \frac{\langle kl \rangle_S - 2\langle l \rangle_S}{\langle k \rangle_S} \tag{3.7b}$$

$$\dot{G}_S(x,y,t) = \pi\, p_{SI}\, (\langle l \rangle_S\, G_S(x,y,t) - y\, G_S^{(0,1)}(x,y,t)), \tag{3.7c}$$

where $G_S^{(0,1)}(x,y,t)$ is the partial derivative of $G_S(x,y,t)$ with respect to y. Note that $\langle l \rangle_S = G_S^{(0,1)}(1,1,t)$ and $\langle k \rangle_S = G_S^{(1,0)}(1,1,t)$. For furthers details about the terms in this equation system, see Table 3.1.

The set of differential equations describing the epidemic process is derived by careful bookkeeping of the links along which an infectious agent spreads (a detailed derivation is provided in Section B.1).

TABLE 3.1 – **Notations used in chapter 3.**

$\dot{f}(x,t) = \frac{\partial}{\partial t} f(x,t)$	partial derivative of function f with respect to t
$f^{(a,b)}(x,y,t) = \frac{\partial^a}{\partial x^a}\frac{\partial^b}{\partial y^b} f(x,y,t)$	partial derivative of function f a times with respect to x and b times with respect to y
A_{kl}	number of individuals in group A with k contacts and l (potential)transmission events (per time interval)
$A = \sum_{k,l} A_{kl}$	number of individuals in group A
$N_{kl} = \sum_{A} A_{kl}$	number of individuals with k contacts and l transmission events (per time interval)
$N = \sum_{k,l} N_{kl}$	total number of individuals
$P_{Akl} = \frac{A_{kl}}{A}$	probability for an individual in group A to have k contacts and l transmission events per time interval
$G_A(x,y,t) = \sum_{k,l} P_{Akl}(t) x^k y^l$	probability generating function (PGF) of $P_{Akl}(t)$
$\langle k \rangle_A = G_A^{(1,0)}(1,1,t)$	average number of contacts of A individuals
$\langle l \rangle_A = G_A^{(0,1)}(1,1,t)$	average number of transmission events per time interval of A individuals
$\langle kl \rangle_A = G_A^{(1,1)}(1,1,t)$	average number of contacts times transmission events per time interval of A individuals
$P_{kl} = \frac{N_{kl}}{N}$	probability for an individual to have k contacts and l (potential) transmission events per time interval
$G(x,y,t) = \sum_{k,l} P_{kl}(t) x^k y^l = \sum_A \frac{A}{N} g_A(x,t)$	probability generating function (PGF) of $P_{kl}(t)$
$\langle k \rangle = G^{(1,0)}(1,1,t)$	average number of contacts of individuals
$M_A = \sum_{k,l} k A_{kl} = A G_A^{(1,0)}(1,1,t)$	number of links coming from A individuals
$M = \sum_A M_A$	number of links
M_{AB}	number of links coming from A individuals and pointing to B individuals
$p_{AB} = \frac{M_{AB}}{M_A}$	probability for a link starting from an A individual to point to an B individual

A,B correspond to epidemic stages, i.e. S, I, R for susceptible, infected, recovered

3.2.3 Analytical results and their validation

Distributions used

To validate the analytical approach, we generated four types of networks corresponding to the combinations of homogeneous and heterogeneous behaviour in the number of contacts k and the number of interaction events l as well as the corresponding networks with a linear dependency between k and l. We studied the spread of an infectious agent on these artificial networks using the analytical approach by plugging the corresponding joint probability generating functions into equations 3.5 and 3.7.

More precisely, we consider combinations of Poisson and power law distributions for the number of contacts k and interaction events per time l. As we neglect isolated hosts, we use the Poisson distribution $P_n = \frac{1-\delta_{n0}}{1-e^{-\langle n \rangle}} \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}$ with support for $n > 0$ and probability generating function $g(x) =$

$\frac{e^{\langle n\rangle(x-1)} - e^{-\langle n\rangle}}{1 - e^{-\langle n\rangle}}$ and power law distributions with exponent λ and cut-off κ, $P_n = \frac{n^{-\lambda} e^{-\frac{n}{\kappa}}}{\mathrm{Li}_\lambda\left(e^{-\frac{1}{\kappa}}\right)}$ and $g(x) = \frac{\mathrm{Li}_\lambda\left(xe^{-\frac{1}{\kappa}}\right)}{\mathrm{Li}_\lambda\left(e^{-\frac{1}{\kappa}}\right)}$ (normalisation through the Polylogarithm Li_λ) for homogeneous and heterogeneous behaviour, respectively. If the joint probability distribution P_{kl} is given by the product $P_{kl} = P_k P_l$ of independent distributions with PGFs $g_1(x)$ and $g_2(y)$, their joint PGF $G(x, y)$ is also given by the product $G(x, y) = g_1(x)g_2(y)$. In the linear case with $P_{kl} = P_k \delta_{kl}$, the PGF $G(x, y)$ is given by $G(x, y) = g_1(xy)$.

Generating networks from PGF

Networks were generated by assigning each host a number k of 'half-contacts' (stubs) and l interaction events per time interval drawn from the distribution P_{kl}. Each host then shared his/her interaction events equi-probably at random (i.e. multinomially) among his/her k contacts. P_{kl} was chosen to satisfy $\langle l\rangle = 2\langle k\rangle$ corresponding to a timescale in which a host has on average 2 interaction events per contact.

The matching of half-contacts (stubs) was done at random but by respecting their assigned number of interaction events per contact. The problem is that not all probability distribution P_{kl} can be realised topologically within a network. Indeed, weighting of interactions between contacts can impose strong constraints on network topology [Serrano *et al.*, 2006]. The fact that contacts can only occur between stubs with the same weight can lead to network segregation and assortative effects that are not seen in the analytical approach *per se* due to its node-centric view (see section B.5). It is even possible to devise joint probability distributions P_{kl} that cannot be represented through a network topology (although this is not a problem for P_{kl} empirically derived from realised networks). In simulated networks, this necessity for exact matching segregates the network into subnetworks, which can show assortativity either with respect to the degree of connected nodes or with respect to their edge weights.

Correcting for assortativity

If the number of interaction events per time and per contact is large (i.e. if $\langle l\rangle/\langle k\rangle \gg 1$), hosts nearly equi-distribute their interaction events among all their contacts due to convergence under the law of large numbers. Decreasing $\langle l\rangle$ yields a more realistic network, as this introduces some variability in the number of interaction events among a host's contacts, which offers more flexibility in the assignment of interaction events on short time scales and reduces assortative effects.

Weight assortativity arises in networks that have a (nearly) constant number of contacts k per individual and where the number of interaction events l is distributed in a heterogeneous way, thus leading to an early expansion among the most highly active individuals. This can accelerate the initial expansion of an epidemic but, at the same time, it also constrains disease spread compared to what one

could expect from the analytical approach. Alternatively, the network can also segregate with respect to the number of contacts an individual holds (i.e. *contact or degree assortativity*). An extreme case can be observed when a (nearly) constant number of interaction events has to be distributed among a heterogeneous number of contacts. This leads to a (near) isolation of individuals with single contacts from the epidemic process (see section B.5).

We can introduce some tolerance in 'negotiating' the number of interaction events per contact. Another way to deal with weight and degree correlations between neighbouring individuals is to drop the assumption that weights are multinomially distributed among an individual's contact. Indeed, heterogeneous weight distributions among an individual's contacts can reduce correlations among neighbouring individuals' degrees and weights [Serrano *et al.*, 2006]. However, this cannot be done without changing P_{kl} to an empirical distribution $\bar{P}_{kl}$, while at the same time deviations might arise from the analytical approach as the assumption of multinomial distribution of weights is violated.

3.2.4 Simulations on weighted networks

Networks for simulation are obtained by first generating 10,000 nodes with k half-contacts (stubs) and l interaction events as drawn from the probability distribution P_{kl}. The l interaction events a node has are then distributed multinomially among its k stubs. Stubs are randomly matched together, with matches being rejected if the weights of the stubs differ by more than one interaction event. In addition matches are rejected if they differ by more than 10 % of the smaller weight involved to avoid biases in nodes with few links If stubs with non-identical weights are matched, the contact is assigned the mean weight randomly rounded to the next integer. This results in the empirical distribution $\bar{P}_{kl}$ as mentioned in the previous subsection.

We use the Gillespie direct algorithm [Keeling & Rohani, 2008] to run stochastic SIR epidemics on continuous time. For each susceptible node i, transmission occurs at rate $\sum_{j \in I} \pi W_{j,i}$, where π is the probability of transmission per sex act, W is the weighted adjacency matrix listing the number of interaction events per time unit between all pairs of nodes, and I is the set of infected nodes connected to node i. Infected nodes recover at rate γ. For each case analyzed, 20 nodes were initially infected uniformly at random in a population of 10,000 and 100 replicate simulations were carried out over each of 20 replicate networks.

3.2.5 Derivation of r_0 and R_0

The derivation of the early exponential growth rate r_0 is based on the observation that the rate of epidemic expansion as described by equations (B.2a-B.2c) is proportional to the number of links/contacts between susceptible and infected individuals M_{SI}, i.e. $p_{SI}\, S\, \langle l \rangle_S = \frac{M_{SI}}{M_S}\, S\, \langle l \rangle_S = M_{SI}\, \frac{\langle l \rangle_S}{\langle k \rangle_S}$, with $M_S = S\, G_S^{(1,0)}(1,1,t) = S \langle k \rangle_S$ being the number of contacts of susceptible individuals.

As there is no explicit expression for M_{SI} we have to rely on approximate equations that are valid in the early stages of an epidemic. In the mean field approximation, M_{SI} is approximated by the number of infected individuals (I) and the average number of contacts per individual found originally in the total population ($\langle k\rangle = G^{(1,0)}(1,1,0)$). As soon as the epidemic is set, i.e. once we are beyond the mean field approximation relying on a randomly picked node, M_{SI} is given by the product of I and a slightly more sophisticated estimate of the number of contacts of infected hosts than in the mean field approximation. More precisely, each infected node contributes to M_{SI} by the average excess degree of a recently infected node chosen proportional to its number of interaction events $\frac{G^{(1,1)}(1,1,t)}{G^{(0,1)}(1,1,t)}-1 = \frac{\langle kl\rangle}{\langle k\rangle}-1$ (for details, see section B.1). This means that the contact the infection has spread from is discounted and that all 'new' contacts are assumed to still be susceptible in the early phase of an epidemic. In order to correctly estimate M_{SI} in the full chain of early infections, it is necessary to discount not only the contact from an individual got infected but also the contact along which the epidemic spreads further. This results in $M_{SI} \approx I\left(\frac{G^{(1,1)}(1,1,t)}{G^{(0,1)}(1,1,t)} - 2\right) = I\left(\frac{\langle kl\rangle}{\langle k\rangle} - 2\right)$.

Altogether we have,

$$\begin{aligned}\dot{I} &= \pi p_{SI}\langle l\rangle_S S - \gamma I \\ &= \pi M_{SI}\frac{\langle l\rangle_S}{\langle k\rangle_S} - \gamma I\end{aligned}$$

We also have $M_{SI}\,\frac{\langle l\rangle_S}{\langle k\rangle_S} \approx M_{SI}\frac{\langle l\rangle}{\langle k\rangle}$ where the approximation holds in the early phase of the epidemic.

If we assume that $M_{SI} \approx I\langle k\rangle$, then

$$\dot{I} = (\pi\langle l\rangle - \gamma)I \tag{3.8}$$

If we assume that $M_{SI} \approx I\left(\frac{\langle kl\rangle}{\langle k\rangle} - 2\right)$, then

$$\dot{I} = \left(\left(\frac{\langle kl\rangle - 2\langle l\rangle}{\langle k\rangle}\right)\pi - \gamma\right)I. \tag{3.9}$$

The exponential growth rate of the infected population in equation 3.9 corresponds to r_0. Approximation (3.8) corresponds to a 'mean field approximation' representing the neighbourhood of a randomly picked node, i.e. not a node picked according to its number of interaction events per time interval. Approximation (3.9) considers that an infected individual has been picked with a probability proportionally to its number of interaction events per time interval. The doubling time t_D can be derived from the early exponential growth rate r_0 as $t_D = \ln(2)/r_0$.

The basic reproductive ratio R_0 is the average number of secondary infections that a typical infected host produces in a fully susceptible population. As for SIR models on classical random networks, it is derived by first evaluating the distribution of excess contacts of a typically infected host, i.e. the

probability for a node chosen according to its number of interaction events per time (l) to have k excess contacts. This probability is $Q_{kl} = \frac{lP_{(k+1)l}}{G^{(0,1)}(1,1)}$. R_0 is calculated in section B.3 as the number of infections that spread along these excess contacts before recovery of the typically infected host.

3.3 Results

3.3.1 Validation of the analytical model with simulations on networks

In order to test our analytical model, we consider epidemiological dynamics taking place on artificial networks on which we release the constraints found in 'classical'/unweighted networks by assuming that the number of interaction events l an individual has does not necessarily increase linearly with his/her number of partners k. To create these networks, we used combinations of Poisson and power law distributions for the number of contacts k and interaction events per time l. This allowed us to introduce arbitrary combinations of homogeneous or heterogeneous behaviour in the way contacts are made and in the number of interaction events established, that may be either independent or dependent (as in the linear case).

To validate the model, we compared the epidemic prevalence (I) from repeated simulation runs with the results derived from the analytical approach using the probability generating functions corresponding to P_{kl} and $\bar{P}_{kl}$, $G(x,y)$ and $\bar{G}(x,y)$. (Note that $P_{kl}(0) = P_{Skl}(0)$, $\bar{P}_{kl}(0) = \bar{P}_{Skl}(0)$, $G(x,y,0) = G_S(x,y,0)$ and $\bar{G}(x,y,0) = \bar{G}_S(x,y,0)$.) The epidemiological dynamics are summarised in Fig. 3.2. In addition of the analytical approach for P_{kl} ($G(x,y)$), $\bar{P}_{kl}$ ($\bar{G}(x,y)$), we also show an approximation (applied to P_{kl}), in which we exclude individuals with one contact. The latter is relevant for networks with heterogeneous number of contacts and (nearly) constant number of interaction events per individual (contact or degree assortativity). Quantitative measures to assess the level of discrepancy between simulations and approximations are provided in section B.6 and table B.1. The derivation of analytical expressions for the error (e.g. a 95% confidence interval) is likely to be an extremely difficult task as these expressions should consider the complex implications on epidemic dynamics that are caused by deviations in the network topology from those of random networks. The simplest way might be to use numerical simulations to estimate the magnitude of the error for given networks.

Overall, the analytical approach matches the simulation results well if the constraints imposed by P_{kl} on the specific networks topology are properly taken into account. The assortativity correction (in orange) is most relevant for a heterogeneous distribution of contacts (k follows a power law) and a homogeneous distribution of interaction events (l is Poisson distributed). This is because the strongest constraints on network topology are expected to occur with these distributions (Fig. 3.2, top panel, right side). On the contrary, for more homogeneous networks (Fig. 3.2, left panel), the assortativity correction is less needed. In the linear case, the correction is not relevant (Fig. 3.2, bottom panel).

A particularly interesting observation is that epidemics spread slower on networks with heterogeneous contacts if the network is weighted than if the number of interaction events scales linearly with the number of contacts for each individual (i.e. 'classical' linear networks). This is particularly true in networks where the average weight per contact is inversely proportional to the number of contacts an individual has. In our simulations, this effect is the clearest when P_k follows a power law distribution and P_l a Poisson distribution (Fig. 3.2 top panel, right side).

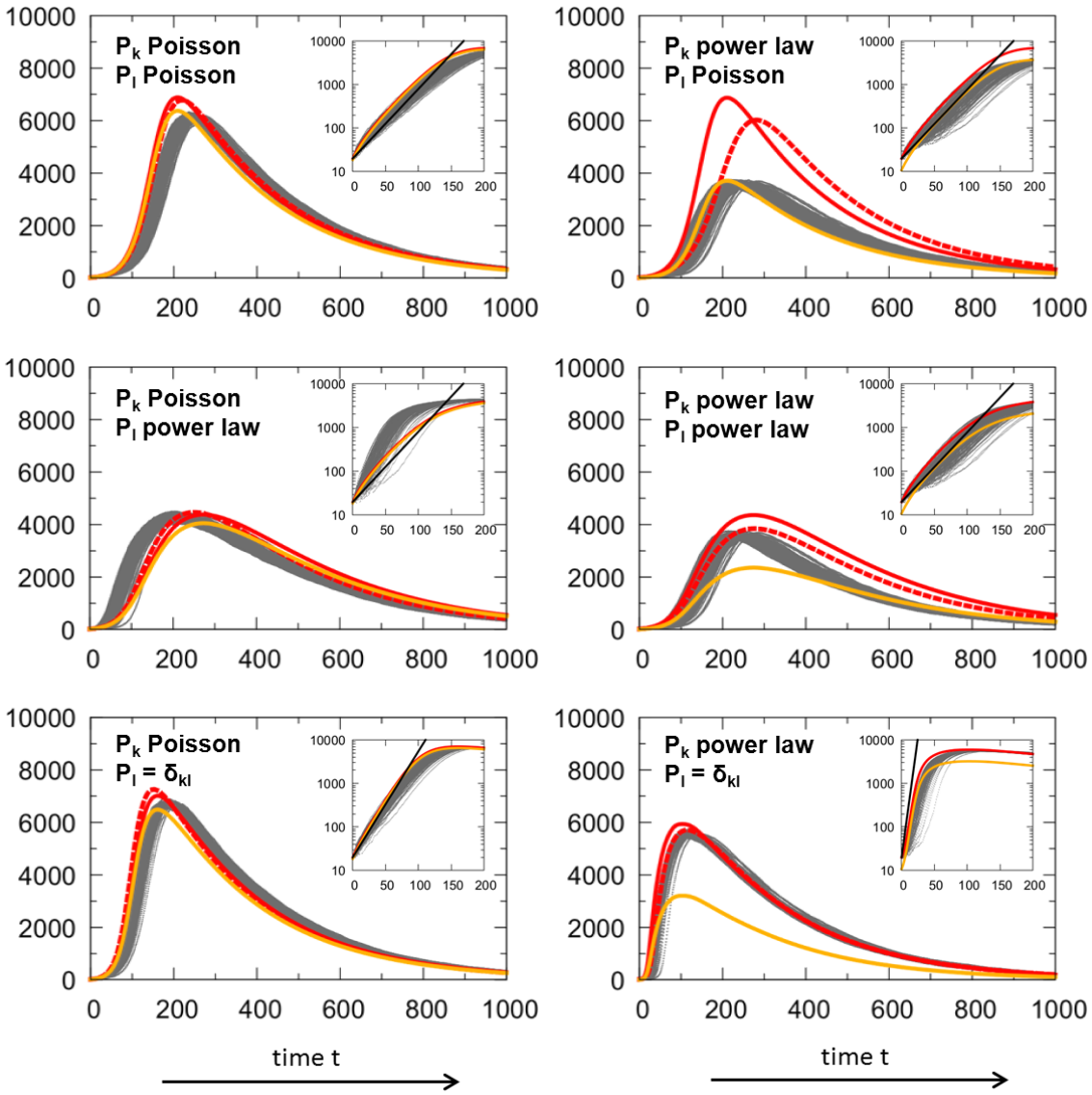


FIGURE 3.2 – **Dynamics of the number of infected hosts (I) on different types of networks.** The distributions in the number of contacts (k) and interaction events per time (l) are either homogeneous (Poisson) or heterogeneous (power law). For the number of interaction events, we also show the linear case in which l is strictly proportional to the number of contacts k, i.e. $P_l = \delta_{kl}$. k and l are drawn from joint distributions P_{kl} with $\langle l \rangle = 2\langle k \rangle$ (except for the analytical P_{kl} model's linear case where $l = k$ being compensated by a double transmission rate). The figures show the epidemic prevalence I as the outcome of the simulation runs (grey, dotted lines), of the the numerical solution of the analytical model with P_{kl} (red, solid line) and $\bar{P}_{kl}$ (red, dashed line). In addition, we show the epidemic prevalence when excluding individuals with only one contact ($k = 1$) which is relevant for epidemics on networks with heterogeneous numbers of contacts including many individuals with $k = 1$ in combination with a (nearly) constant number of interaction events, as realised through a Poisson distribution (orange line, cf. specifically power law, Poisson). Parameters chosen correspond to $\langle k \rangle = 4$ (Poisson case : $\langle k \rangle = 4$, $\langle l \rangle = 8$, power law case : $\lambda_k = 1.4$, $\lambda_l = 0.89$, $\kappa_k = \kappa_l = 22$). Epidemiological parameters are $\pi = 0.01$ (0.02 for the analytical P_{kl} model's linear case), $\gamma = 0.004$ in arbitrary units and $I(0) = 20$. The insets show the same data for the early epidemic expansion in logarithmic scale showing early exponential growth according to $I(0)e^{r_0 t}$ (black line) with r_0 from table 3.2.

3.3.2 Capturing epidemic characteristics (expressions for r_0 and R_0)

There are different ways to assess the initial propagation of an infectious agent in an otherwise fully susceptible population. One possibility is to estimate the initial exponential growth rate in the number of infected individuals (r_0). Another possibility consists in estimating the number of secondary cases created by a newly infected host in a fully susceptible population, which is classically referred to as the basic reproductive ratio R_0 [Anderson & May, 1991].

Subtle effects arise depending on whether we choose the neighbourhood of a random individual as a reference or the neighbourhood of a 'typically' infected individual. The first case corresponds to what is usually referred to as a 'mean field approximation' and captures well the very first infection events. The second case (using a 'typical' infected individual) is more appropriate to capture the next stages of early epidemic expansion because it accounts for the fact that spatial structure has been sensed or set by the epidemic process. We thus use it for the derivation of R_0.

The expressions for r_0 and R_0 for the SIR model are shown in Table 3.2 and are derived in details in the Materials and Methods and in appendix B, respectively. Note that these do not involve any approximation beyond those implied in the model's assumptions, i.e. they are exact within the model framework. The derivation for R_0 when $\gamma > 0$ is the only case that requires some further approximations to obtain an explicit formula (section B.3).

In the expression for r_0 and R_0, all the occurrences of the transmission probability (π) are weighted by the number of interaction events per contact. This makes sense because this corresponds to a transmission rate. Note that there is a slight difference between r_0 and R_0 because in the former we have the ratio of the means ($\langle l \rangle / \langle k \rangle$), whereas in the latter we have mean of the ratios ($\langle l/k \rangle$). This is due to the fact that the averaging is done at a different step in the calculations.

More interestingly, the expressions for r_0 and R_0 both scale with the second moment $\langle kl \rangle$ of the joint probability distribution P_{kl}. This implies that the number of contacts (k) and the interaction events (l) an individual maintains equally affect epidemic spread. At the same time the correlation between these quantities is relevant to model rapid epidemic spread : for epidemic control, targeting individuals with most contacts or interaction events can prove to be much less efficient than targeting those who maximise both. The formulae in Table 3.2 are generalisations of formula 3.4b, which corresponds to the linear case where the number of interaction events scales with the number of contacts a person maintains, i.e. $\langle kl \rangle \sim \langle k^2 \rangle$. Note that earlier approaches on weighted networks correspond to the fully mixed situation in formula 3.4a in which k is interpreted as the number of interaction events [May & Anderson, 1987].

Figure 3.2 shows epidemic expansion in different types of simulated networks (grey, dotted lines) in comparison with analytical approximations (red and orange lines). The insets in logarithmic scale show that the exponential growth rate $r_0 = \left(\frac{\langle kl \rangle}{\langle l \rangle} - 2 \right) \frac{\langle l \rangle}{\langle k \rangle} \pi - \gamma$ calculated for the early epidemic

expansion approximates well the simulation and analytical results. The early exponential growth rate r_0 can slightly underestimate the dynamics if the contact network is homogeneous (P_k is Poisson distributed), while the distribution of interaction events is heterogeneous (P_l follows a power law) . This underestimation is due to the fact that weight assortativity generates a subnetwork of individuals with many interaction events, which speeds up epidemic expansion in its early phase (Fig. 3.2, middle panel, left side).

TABLE 3.2 – r_0 **vs.** R_0**.** π is the transmission probability, γ the recovery rate, $\langle l \rangle$ the average number of interaction events an individual has, $\langle k \rangle$ the average number of contacts per individual, $\langle kl \rangle$ is the second moment of the joint probability distribution P_{kl} and $\langle l/k \rangle$ is the average number interaction events per contact. Note that the equation for R_0 is exact for $\gamma = 0$.

	Early epidemic growth rate r_0	Basic reproductive ratio R_0
Epidemic expansion from randomly picked index case, (mean field approximation) $M_{SI} \approx I\langle k \rangle$	$\dot{I} \approx (\pi \langle l \rangle - \gamma) I$ $r_0^{MF} = \pi \langle l \rangle - \gamma$	$R_0^{MF} = \frac{\pi \langle l \rangle}{\gamma}$
Early epidemic expansion, structure set by epidemic, $M_{SI} \approx I\left(\frac{\langle kl \rangle}{\langle l \rangle} - 2\right)$	$\dot{I} \approx \left(\left(\frac{\langle kl \rangle}{\langle l \rangle} - 2\right)\frac{\langle l \rangle}{\langle k \rangle}\pi - \gamma\right) I$ $r_0 = \left(\frac{\langle kl \rangle}{\langle l \rangle} - 2\right)\frac{\langle l \rangle}{\langle k \rangle}\pi - \gamma$	$R_0 \approx \frac{\langle \frac{l}{k} \rangle \pi}{\langle \frac{l}{k} \rangle \pi + \gamma}\left(\frac{\langle kl \rangle}{\langle l \rangle} - 1\right)$

3.3.3 Application of the model to epidemiological data

The knowledge of transmission networks along which an infectious agent can spread within a host population is of great importance to public health. These networks might be hard to assess for airborne infections because they are very dynamic [Edmunds *et al.*, 2006; Wallinga *et al.*, 2006] but easier to infer for sexually transmitted infections (STI) because they are more static. Such sexual contact networks have been surveyed in many studies covering homosexual as well as heterosexual populations and different societal contexts [Schneeberger *et al.*, 2004; Johnson *et al.*, 2001; Hamilton *et al.*, 2008] to understand and prevent the spread of STI. The National Survey of Sexual Attitudes and Lifestyles (NATSAL [Johnson *et al.*, 2001]) provides detailed data on the situation in the United Kingdom, including distributions in the number of sexual partners (k) and sex acts (interaction events, l) a person has within certain time frames.

As shown in Fig. 3.3A, both the number of partners (contacts, k) and sex acts (interaction events, l) an individual has are heterogeneously distributed. However, their joint distribution P_{kl} does not show a linear behaviour, implying that the number of sex acts l does not scale linearly with the number

of partners k an individual has. This is also supported by Pearson's correlation coefficient which is 0.15, i.e. positive but not indicating a strong linear relationship between k and l (see also Fig. B.3). When they combine a linear relationship between number of partners and number of sex acts with the observed broad distributions of sexual contacts and sex acts, several models predict extremely rapid early epidemic expansion and an epidemic threshold that is potentially vanishing in the limit of infinite network size [Durrett, 2007; May & Anderson, 1987; Moreno *et al.*, 2002; Volz & Meyers, 2009; Boguna *et al.*, 2003; Bansal *et al.*, 2007], as can be seen from equation 3.4.

In our 'Validation' section we have shown that, regarding the number of interaction events, a deviation from the linear behaviour decreases epidemic expansion and peak prevalence, especially for transmission networks that are characterized by a heterogeneous distribution in the number of contacts per individual k and regardless of whether the distribution of interaction events/sex acts is homogeneous (Poisson) or heterogeneous (power law). This is also reflected more quantitatively in the expression for R_0, which is dominated by the second moment $\langle kl \rangle$ of P_{kl}. In other words, (as explained above when interpreting our approximations) the more the number of partners (k) correlates with their number of sex acts (l) the faster the early epidemic expansion.

These differences between weighted and unweighted networks are visible in Fig. 3.3, which shows the epidemic expansion of a susceptible-infected (SI) epidemic with transmission probability $\pi = 0.01$ per sex act in two scenarios. We chose to model SI dynamics by setting $\gamma = 0$ in order to simulate HIV spreading on a short time scale, but the model could be evaluated with analogous results for SIR dynamics (Fig. B.1). In the first scenario (black curve, Fig. 3.3B), the number of sex acts grows linearly with the number of contacts respecting the average number of sex acts ($\langle l \rangle = 6.05$ during 4 weeks). In the second scenario we use a weighted network, where the joint probability distribution P_{kl} is obtained from the NATSAL data (red curve, Fig. 3.3B) and can be complemented by a correction for assortative effects (orange curve, Fig. 3.3B).

The exponential growth rates of the epidemics are $r_0 = 0.021$ per year for the linear, unweighted, network and $r_0 = 0.0034$ per year for the weighted network. This confirms that the 'linear' scenario supports faster epidemic expansion. The correction for assortative effects underestimates the epidemic prevalence in the network because in the NATSAL network heterogeneity in the number of contacts and interaction events does not lead to a strong network segregation, i.e. individuals with a single or few contacts are not isolated (see Fig. 3.2 and the case in which both P_k and P_l follow power law distributions). Although the survey data shown in Fig. 3.3A only provide us with a rough picture of the real transmission network and although relying on the number of partners during 5 years overestimates the number of concurrent partners, the data are sufficient to confirm a remarkable reduction in the speed of epidemic expansion when shifting from a classical unweighted transmission network towards a more realistic weighted transmission network. This finding is in particular consistent with an earlier simulation study [Moslonka-Lefebvre *et al.*, 2012a].

The framework can also be used to follow the dynamics in the epidemic subgroups and to identify risk groups along the course of the epidemics (sectionB.2). Fig. 3.3C and D show that the average number of contacts $\langle k \rangle$ and sex acts $\langle l \rangle$ per individual increases early on in the infected population and decreases in the healthy population. This reflects the over-representation of some individuals among those being infected as sketched by the insets : during the epidemic (here shown at t = 10 years) the probability of being infected grows with the number of contacts k an individual has, and even more so with the number of sex acts l (s)he has.

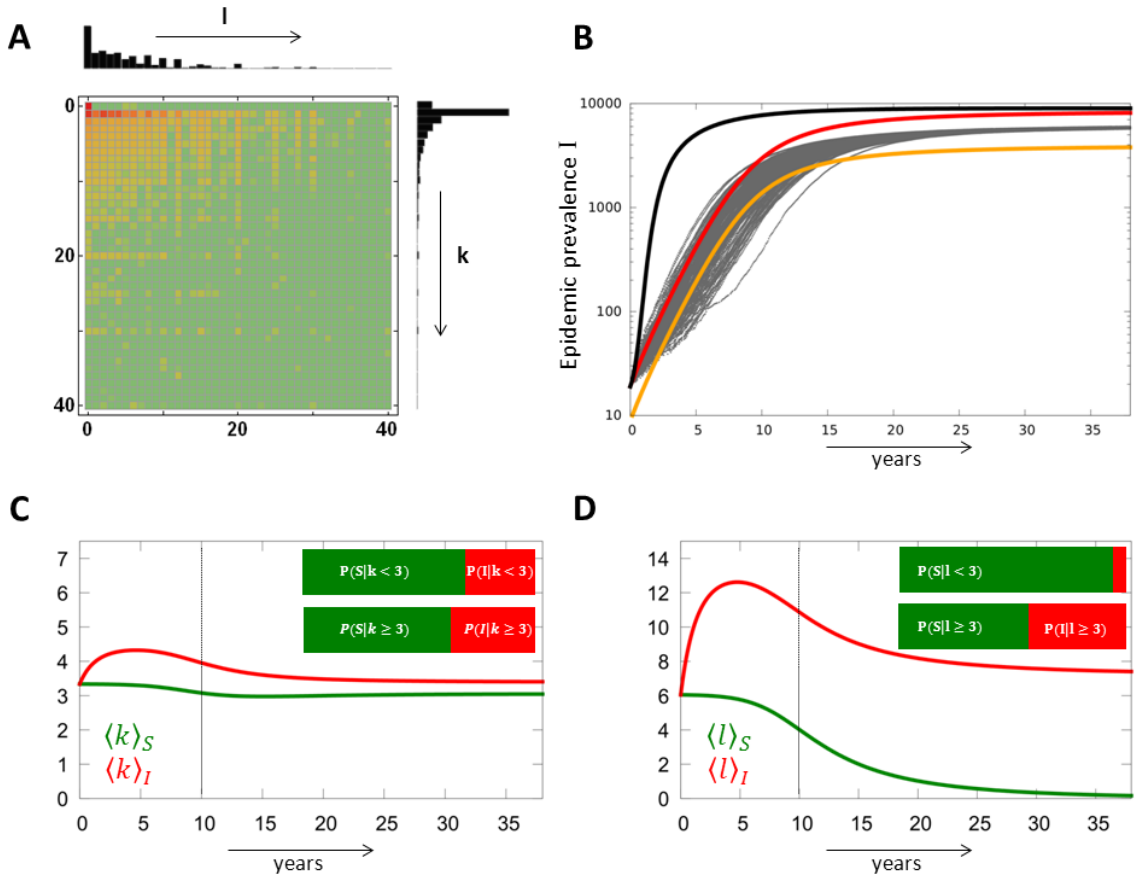


FIGURE 3.3 – **Disease spread on a network inferred from data.** A) Characteristics of the heterosexual contact network inferred from the NATSAL contact tracing study [Johnson *et al.*, 2001]. The network shows a heterogeneous joint probability distribution P_{kl}, which is the probability for an individual to have k contacts during the last 5 years and l sex acts during the last 4 weeks (higher values of P_{kl} are in red and lower values are in green). This heterogeneity is also seen for the marginal distributions P_k (on the right) and P_l (on the top). B) Dynamics of an SI epidemic spreading on an unweighted (black line) or a weighted sexual contact network. The results of simulations on the weighted network are in grey, the approximations of our model are in red or in orange for the case with assortativity correction. The network has been reduced to nodes with $k > 0$ and transmission probability per sex act is π = 0.01. C) Dynamics of the average number of contacts $\langle k \rangle$ of susceptible (in green) and infected individuals (in red) over the course of an epidemic spreading on the weighted network. The inset shows the probability of a host to be susceptible or infected at t = 10 years conditioned to the number of contacts during the last 5 years (more or less than 3 contacts). D) Same as panel C but for the number of sex acts $\langle l \rangle$. In Panels B, C and D, the weighting is done using the P_{kl} shown in panel A. Individuals with more contacts tend to be disproportionally infected (panel C). Individuals with more sex acts tend to be even more infected (panel D).

3.4 Discussion

Network theory has broadened our understanding of the spread of infectious agents — or other entities such as information, money, travellers or goods — in complex settings. In their simplest form, network models do not consider that contacts may show variability in their transmission capacity. However, the probability of disease transmission along a contact strongly depends on the intensity of the contact, transportation links vary in their throughput and information may not be shared equally among all possible channels. Although earlier studies have shown that this weighting in terms of interactions between contacts has non-negligible impact on the spreading dynamics, the modelling of epidemics on weighted networks largely focuses on simulation studies [Eames *et al.*, 2009; Chu *et al.*, 2009; Moslonka-Lefebvre *et al.*, 2012a], regular networks [Rattana *et al.*, 2013], mean field approximations [Chu *et al.*, 2011; Yang & Zhou, 2012] or discrete time dynamics [Britton *et al.*, 2011; Deijfen, 2011]. Therefore, explicit expressions for epidemic characteristics such as the basic reproductive ratio R_0 are available only in special cases.

It is possible to simplify the epidemiology by using a Reed-Frost model. For this, one needs to assume that infections take place in discrete time steps, with non-overlapping generations and that each infected individual recovers with certainty one time step after infection. These simplifications allow to assess outbreak probabilities using branching processes [Britton *et al.*, 2011]. In this formalism, as shown in [Deijfen, 2011], R_0, denoted R_0^{RF}, can be derived as the dominant eigenvalue of the mean offspring matrix $(m_{d,k\geq 2})$, where $m_{d,k}$ represents the expected number of individuals with k contacts that an individual with d contacts infects considering potentially degree-dependent network weights. Importantly, it is only because the Britton *et al.* make strong simplifying assumptions in their model, such as the independence between network weights and nodes' degrees, that they can derive an explicit form of R_0. In contrast to our findings on NATSAL data, the Reed-Frost approach systematically predicts negative exponential growth rates of the epidemics for both scenarios (the network average of linear case is $\langle R_0^{\mathrm{RF}} \rangle = 0.2$ and that of survey data case is $\langle R_0^{\mathrm{RF}} \rangle = 0.1$). The discrepancy between our model and that of [Britton *et al.*, 2011] stems from the implicit assumption of the Reed-Frost model with discrete time steps they use, which is that recovery occurs immediately after infection and therefore that $\gamma \gg \pi$.

We extend earlier results by developing a framework based on partial differential equations that allows to model continuous time SIR epidemic dynamics for general weighted networks defined through the joint probability distribution for an individual to have k contacts and l interaction events. From this we are able to derive the full epidemic dynamics in terms of the number of susceptible, infected and recovered individuals over time as well as explicit expressions for the basic reproductive ratio R_0 and the exponent of early epidemic expansion r_0. The application of the method to epidemics on artificial and empirically-motivated networks matches well with simulation results on these same

networks. Moreover, it also stresses the impact of assortative effects introduced by contact weighting on epidemic dynamics ; an aspect that will need closer attention in future research.

One limitation to our approach is due to potential errors in the inference of the network. There are known biases in the self-reporting of number of partners (with different trends between men and women [Smith, 1992]) and self-reported number of sex acts are likely to exhibit similar biases. One extension of this study would be to see how such noise in the network inference could affect epidemic spread. Our intuition is that the consequences should be less important than for non weighted networks because heterogeneity in the weights is already likely to dampen striking network properties in terms of disease spread [Moslonka-Lefebvre *et al.*, 2012a].

As many earlier methods, ours analyses model disease spread on networks from a node centric summary statistics, by considering the number of contacts and transmission events per time. Therefore, it inherently neglects correlation between nodes. In other words there is no consideration of assortativity between individuals based on their number of contacts or transmission events per time. At the same time, individuals share their activity randomly among all their contacts (weights are homogeneously, or multinomially,distributed among the edges that leave a node), which can enforce correlations among nodes in certain networks. Also clustering is observed in many contact networks [Danon *et al.*, 2012] and this issue should be addressed in an extended version of our model.

Most analytical and numerical models predict disease spread on network using only one summary statistics, the distribution of the number of partners. We show that additional insights can be gained, while maintaining some analytical results, by including another summary statistics, such as the distribution of the number of sex acts knowing the number of partners. These data are easier to collect than full information of the contact network (especially for a weighted network), which makes our framework widely applicable. We demonstrate this applicability here using data from the NATSAL study conducted in the UK. We note that for some artificial distributions, our results begin to diverge from simulations on real networks. However, the framework has proven to be applicable for empirical distributions and analysis of more empirical data will allows us to further test the robustness of the method using more realistic assumptions.

Acknowledgments

Many thanks to Hervé Monod and Elisabeta Vergu for helpful discussions and insights, Oliver Rübenkönig for consulting with the Mathematica code, Andreas Bunten for his support in the final stages of the manuscript, and to three anonymous reviewers for numerous suggestions. MML would like to thank the French Ministries in charge of Agriculture and Environment for financial support. SA is funded by an ATIP-Avenir grant from CNRS and INSERM and by the IRD and the CNRS.

Chapitre 4

Exchange directionality impacts transmission. The case of *P. ramorum* spread through trade of ornmentals.

This chapter highlights the epidemiological importance of the directional constraint and summarises a research article published in *Mathematical Biosciences* [Moslonka-Lefebvre *et al.*, 2012b]. The detailed results are available in appendix C, but are not included in the main text since the work was not initiated as part of this PhD project.

Trade flows are well described by directed networks. For example, the trade of ornamental plants is better approximated by directed networks, given that connections from e.g. producers to wholesalers and final customers are more likely than in the reverse direction [Pautasso *et al.*, 2010]. Most studies have focused on undirected networks where nodes are connected symmetrically. Directed networks, given their a priori asymmetric adjacency matrices, have been used relatively rarely to model epidemics [Newman *et al.*, 2001; Meyers *et al.*, 2006; Park & Kim, 2006]. An epidemiological model specifically designed for directed networks would enable a more realistic representation of epidemics spreading on markets. In [Moslonka-Lefebvre *et al.*, 2012b], we introduce and explore a model of epidemic spread on unweighted directed networks, which we term the *Susceptible-Infected-Susceptible along a continuum* (SIS_c) framework. The SIS_c model describes the infection status of each agent of a directed network along a continuum ranging between full susceptibility to full infection. This is realistic for epidemics in the plant trade, as plant nurseries and retail centres can be completely devoid of a pathogen, fully infected, or everything in between [Pautasso & Jeger, 2008]. Keeping the notations of part I, we present the SIS_c model in continuous time, where infectious individuals transmit the infection at the per-contact transmission rate ω and remove infectious stock at rate γ (see section 2.2.2. of appendix C). Let $Y_i(t)$ denotes the total number of infected products within agent i at time t. The SIS_c process models the

evolution of Y_i over time and reads :

$$\frac{dY_i}{dt} = \omega \sum_{j \neq i} w_{ji} Y_j - \gamma Y_i \ , \tag{4.1}$$

where, $w_{ji} = 1$ if j is connected to i, and $w_{ji} = 0$ otherwise. Note that $w_{ji} \neq w_{ij}$ in the general case (see also subsection 2.3.1). The SIS$_c$ model has many applications such as information, disease and rumour spread among individuals or organisations. Here, we focus on disease spread in horticultural trade networks, as observed with *P. ramorum* [Jeger *et al.*, 2007; Moslonka-Lefebvre *et al.*, 2011]. In such a system, an SIS process is particularly relevant because immunization is not possible : if plant growers carry on trading susceptible species, then they are still at risk of becoming infected even after eradication of a certain pathogen and the adoption of good practices.

We start by assessing the overall epidemic risk posed by a directed network. As for a weighted network (chapter 3), a directed network can be characterised by its joint probability distribution $P_{k^{in}k^{out}}$, where k_i^{in} and k_i^{out} represent the in- and out-number of partners (in- and out-degrees) of agent i respectively. The basic reproduction number R_0 [1] associated with the SIS$_c$ frameworks can be estimated from classical mean-field assumptions which neglect degree–degree correlations between pairs of agents [Restrepo *et al.*, 2007] (see section 3.1 of appendix C for further details) :

$$R_0 \approx \frac{\omega}{\gamma} \frac{\left\langle k^{in} k^{out} \right\rangle}{\langle k \rangle} \ , \tag{4.2}$$

where $\langle k \rangle$ is the average degree (the averages of the in- and the out-degree have to be the same), and $\left\langle k^{in} k^{out} \right\rangle$ is the second moment of the joint probability distribution $P_{k^{in}k^{out}}$. Equation (4.2) shows correlations between in and out-degrees ($\left\langle k^{in} k^{out} \right\rangle$) have a positive effect on R_0. As an extreme case, if the in- and out-degree of each agent i are identical ($k_i^{in} = k_i^{out} = k_i$ in (4.2)), R_0 reads :

$$R_0 \approx \frac{\omega}{\gamma} \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} \ , \tag{4.3}$$

which is identical to relation (3.4a), i.e. the formula of R_0 for fully mixed undirected networks.
In contrast, in the average case of uncorrelated directed networks (when $\left\langle k^{in} k^{out} \right\rangle = \left\langle k^{in} \right\rangle \left\langle k^{out} \right\rangle = \langle k \rangle^2$ in (4.2)), we notice :

$$R_0 \approx \frac{\omega}{\gamma} \langle k \rangle \leq \frac{\omega}{\gamma} \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} \ , \tag{4.4}$$

which implies a lower risk of invasion for uncorrelated [2] directed networks in comparison with positively correlated directed networks. Keeping the total number of directed links constant, we hence expect

1. For cases where agents are either strictly susceptible or strictly infected, we can reasonably postulate, by analogy with (3.4a)–(3.4b), that R_0 is rather given by $R_0 \approx \frac{\omega}{\omega + \gamma} \left(\frac{\left\langle k^{in} k^{out} \right\rangle}{\langle k \rangle} - 1 \right)$.

2. The risk would be even lower for networks with negative correlations between the in- and out-degrees at the agent-level

directed networks to be associated with a lower risk of epidemic invasion than undirected networks (where by definition directed links are symmetric for each pair of node, which yields maximal correlation between the in- and out-degree at the agent-level). These theoretical intuitions are in perfect agreement with the numerical results reported in [Moslonka-Lefebvre *et al.*, 2009].

For a given directed network, the SIS_c model also provides interesting insights to rank relative epidemic risks at the agent-level. Following Pautasso *et al.* [2010], we classify agents as producers, wholesalers or retailers depending on the relative difference of their in- and out- degrees. More generally, the hierarchical position of an agent i in a directed network, which positions i's in the supply chain, is quantified by a hierarchical index Δ_i defined as :

$$\Delta_i = \frac{k_i^{in} - k_i^{out}}{k_i^{in} + k_i^{out}} \ , \tag{4.5}$$

where by construction $-1 \leq \Delta_i \leq -1$. Producers can be visualized as mainly exporters (negative Δ_i), retailers as mainly importers (positive Δ_i), and wholesalers as agents with no preponderance of either export or import (values of Δ_i centered on zero), thus linking the other two categories. In other words, we assume the relative directions of links reflect a range of agents' specialisations in the supply chain (Fig. 2.1E). Note that we translated definition (24) in appendix C to ease comparison with the notion of flow polarity defined in chapter 5. Relation 4.2 suggests an effective control of disease spread should reduce high correlations between links in and links out by targeting wholesalers preferentially, an assumption which we confirmed numerically in [Pautasso *et al.*, 2010] and in section 3.2 of appendix C.

Chapitre 5

Combining weight and directionality to understand exchange-driven epidemics. Market-epidemiological analyses of livestock movements as a case study.

This chapter presents an article entitled *Market analyses of livestock trade networks and prevention of joint economic and epidemiological risks*, currently in resubmission – allowed by the editorial board – to *Journal of the Royal Society Interface*. In a nutshell, we describe livestock movements as weighted directed networks, because the total number of animals shipped per trade relationship over a given period of time is variable between pairs of holdings [e.g. Vernon, 2011], and animals are essentially transported from selling to buying holdings [e.g. Lal Dutta *et al.*, 2014]. Based on a new economic categorisation of agents which accounts both for holdings' weights (weighting constraint, see chapter 3) and their overall direction of exchange (directional constraint, see chapters 4), we identify cost-effective strategies for preventing epidemics with limited adverse effects on disease-conducing trade.

Trade of livestock is crucial for the agri-food sector, but also constitutes a path of infectious disease transmission [Vernon, 2011]. The architecture of animal movements is extremely rich and can be captured by networks of potentially-infectious contacts [Vernon & Keeling, 2009]. By allowing the calculation of structural determinants of epidemic risk, such network models prove invaluable to identify animal holdings that are most likely to drive epidemics [Bajardi *et al.*, 2012]. Examples of risk determinants include the total number of commercial partners of an agent [the total degree, see e.g. Rautureau *et al.*, 2011] and the number of premises that can be reached by following temporally-compatible links from a given holding [the outgoing infection chain, see e.g. Noremark *et al.*, 2011]. In the general case, any measure of network centrality for a given holding can be used as an indicator of the corresponding risk of contagion and / or transmission. To prevent and mitigate epidemics, it is naturally recommended

to surveil preferentially those holdings that are most central to livestock markets. However, targeting core market players can be costly, rendering such strategies inapplicable. Major holdings involve numerous animals to screen, incurring additional costs to the regulator, and their disturbance, e.g. through excessive prevention, can cause economic downturns [Paton *et al.*, 2009]. The evaluation of combined economic and epidemiological risk determinants remains a key knowledge gap.

Here, we aim at identifying efficient strategies for preventing epidemics with minimal disruptions to the market and limited costs through calculation of proxies for economic and epidemic burdens. We hence introduce a novel market-based categorisation to cluster the holdings based on economic rather than purely structural network summaries. In our framework, each holding i corresponds to an economic agent characterised by two quantities, namely *flow polarity*, denoted fp_i, and *flow share*, denoted fs_i, which we define as :

$$\begin{aligned} \mathrm{fp}_i &= \frac{\phi_i^{\mathrm{in}} - \phi_i^{\mathrm{out}}}{\phi_i^{\mathrm{in}} + \phi_i^{\mathrm{out}}} \ , \\ \mathrm{fs}_i &= \frac{\phi_i^{\mathrm{in}} + \phi_i^{\mathrm{out}}}{\sum_j (\phi_j^{\mathrm{in}} + \phi_j^{\mathrm{out}})} \ , \end{aligned} \tag{5.1}$$

where ϕ_i^{in} and ϕ_i^{out} respectively represent the trade volume bought and sold by agent i over a certain period of time. Flow polarity quantifies agent i's *position in the supply chain* (Fig. 5.1A,C), and constitutes a such an extension of the hierarchical index (4.5) to directed networks with heterogeneous weights. Flow share is a standard measure of market leadership in marketing studies [Kotler & Armstrong, 2013], and can be interpreted as the relative weight of an agent in a given network (Fig. 5.1B,D). We define *market categories* based on joint flow polarity and flow share and study the networks resulting from these categories (Fig. 5.3).

By analysing two datasets recording French cattle and swine movements, we show market categories correspond to contrasted groups of economic agents (Fig. 5.1C,D), are relatively time-invariant and prove useful to assess joint economic-epidemiological risk. While risk trends by category are similar for both markets, our analyses suggest trade of cattle is riskier than swine, which is in agreement with a former study [Rautureau *et al.*, 2012]. We finally compare contrasted strategies for preventing epidemics. Both static and dynamical preventive strategies are explored, whether based on real-time or past data. The strategies we describe are implemented by a public authority preferentially protecting agents belonging to particular market categories at risk. However, the perceptions of what is the best strategy depends on who evaluates it as the best, because protection is costly and may generate distortions to the market. We hence confront the regulator's and the market's standpoints when comparing strategies. From the sole point of view of the public authority, and if we assume intervention costs are proportional to the number of economic agents to protect, protecting preferentially those agents with the largest flow shares appears the best strategy, which is a standard recommendation in the

literature for similar systems [Albert *et al.*, 2000]. In sharp contrast, when the point of view of the market is also taken into account or when costs of prevention are rather proportional to the total trade volume as expected for vaccination campaigns, it may be more interesting to protect preferentially economically-minor agents, especially in cattle.

In essence, our study suggests that multiple standpoints should be adopted when evaluating targeted preventive strategies, because exchanges, subsequent infection events and the regulation of epidemics have differential consequences on and involve contrasted decisions from diverse agents. This finding has general implications for epidemiological and ecological studies aiming at prioritising interventions to maintain healthy and diverse (eco-)systems.

Market analyses of livestock trade networks to inform the prevention of joint economic and epidemiological risk

Mathieu Moslonka-Lefebvre[1,2,3,#], Christopher A. Gilligan[2], Hervé Monod[1], Catherine Belloc[4,5], Pauline Ezanno[4,5], João A. N. Filipe[2,*] and Elisabeta Vergu[1,*]

[1] INRA, UR 0341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France.

[2] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, United Kingdom.

[3] AgroParisTech, F-75005 Paris, France.

[4] INRA, UMR1300 Biologie, Epidémiologie et Analyse de Risques en santé animale, CS 40706, F-44307 Nantes, France.

[5] LUNAM Université, Oniris, Ecole nationale vétérinaire, agroalimentaire et de l'alimentation Nantes-Atlantique, UMR BioEpAR, F-44307 Nantes, France.

[#] Corresponding author, email : mathieu@moslonkalefebvre.com

[*] These authors contributed equally to this work

Abstract

Conventional epidemiological studies of infections spreading through trade-networks, e.g. via livestock movements, generally show that central large-size holdings ("hubs") should be preferentially surveyed and controlled in order to reduce epidemic spread. However, epidemiological strategies alone may not be economically optimal when costs of control are factored in together with risks of market disruption by targeting core holdings in a supply chain. Using extensive data for animal movements and the supply chains for cattle and swine in France, we introduce a method to identify effective strategies for preventing outbreaks with limited budgets while minimimising the risk of market disruptions. Our method involves the categorisation of holdings based on position along the supply chain and degree of market share. Our analyses suggest that trade is riskier in propagating epidemics through cattle networks, which are dominated by exchanges involving wholesalers, than for swine. We assess the effectiveness of contrasting interventions from the perspectives of regulators and the market, using percolation analysis. We show preferentially targeting minor, non-central agents can outperform targeting of hubs when the costs to stakeholders and the risks of market disturbance are considered. Our study highlights the importance of assessing joint economic-epidemiological risk in networks underlying pathogen propagation and trade.

Keywords : animal trade networks ; disease prevention ; economic epidemiology ; livestock ; multiple-criteria decision analyses ; risk-based surveillance.

5.1 Introduction

Trade is crucial for the economy, but can also drive infectious disease transmission, sustaining epidemics locally and promoting potentially long-distance introductions [e.g. Vernon, 2011]. Examples of markets that can contribute to epidemic outbreaks include trade of livestock such as cattle [Lal Dutta *et al.*, 2014], swine [Noremark *et al.*, 2011], and sheep [Kiss *et al.*, 2006] ; prostitution [Rocha *et al.*, 2011] ; and airline transportation [Colizza *et al.*, 2006]. In the epidemiological literature, the contact structure underpinning pathogen spread through trading contacts is usually described using network models ([e.g. Vernon & Keeling, 2009]). In such models, holdings (e.g. farms) are represented by nodes that are interconnected by links that represent exchanges among holdings (e.g. by movement of animals). In the past decade, network-based models have become increasingly popular to achieve a threefold objective : *i*) to describe the contact structure spanned by such markets [e.g. Vernon, 2011], *ii*) to assess the epidemic risk factors at the scale of individual holdings [e.g. Bajardi *et al.*, 2012] and *iii*) to design effective disease control strategies [e.g. Pellis *et al.*, 2014]. In particular, network analyses have proved useful in identifying super-spreading holdings, usually referred to as "hubs", that should be preferentially subjected to trade restrictions in order to prevent and mitigate epidemics. The disruption of such core market players, however, can cause economic shocks, a downside, which, to the best of our knowledge, has not been considered in network-based, data-driven epidemiological studies. Here, we consider the key, but yet unaddressed question of evaluating the trade-off between efficiencies of trade pathways as commercial trade routes and, simultaneously, the vulnerabilities as routes for the transmission of economically damaging pathogens. We investigate this trade-off for livestock exchange markets in France, for which we have access to extensive data.

The construction of network models of livestock markets requires the use of records of animal movements. In the EU, for instance, exhaustive tracing of livestock-movement is available in many nationally maintained datasets, originally for cattle since the 1990s (Council Directive 92/102/EEC of 27 November 1992 on the identification and registration of animals) and more recently for swine (Council Directive 2008/71/EC of 15 July 2008 on the identification and registration of pigs). In other countries, data with such levels of detail may not be available because, for example, the data may have been aggregated to comply with privacy laws, or routine data collection may not be implemented at the farm level. The architecture of animal movements is often extremely rich and can be described by more or less simple network models depending on the objective of the epidemiological study Vernon & Keeling [2009]. In some cases, livestock-exchange network models can account for more than one type of node (e.g. farms versus purely commercial holdings) [e.g. Rautureau *et al.*, 2011], direction of exchange (animals are essentially shipped from selling to buying holdings) [e.g. Dube *et al.*, 2008], weight of shipments (when the number of animals shipped in one go varies) [e.g. Vernon, 2011], and dynamical aspects (the shipment of animals occurs at certain points in time) [e.g. Vernon & Keeling, 2009].

It follows that empirical livestock exchange networks exhibit key features shared by many complex networks, namely they are *multipartite*, *directed*, *weighted* and *dynamic*. Recent studies with network analyses of livestock markets provide significant information on demographic aspects and vulnerability to pathogen transmission in cattle and pig markets in several countries, especially where detailed data are routinely collected, including cattle, for example in the UK, Sweden, France and Italy [Vernon, 2011; Noremark *et al.*, 2011; Rautureau *et al.*, 2011; Bajardi *et al.*, 2012] for cattle and Sweden, France and Germany [Noremark *et al.*, 2011; Rautureau *et al.*, 2012; Buttner *et al.*, 2013a], for pigs.

Various determinants of the risk of transmission of infection, and hence disease, can be calculated depending on the features included in a given network model. These determinants of risk include the total number of commercial partners per agent [the total degree, see e.g. Rautureau *et al.*, 2011] and the number of premises that can be reached through successive temporally-compatible links [Noremark *et al.*, 2011]. In the general case, any measure of network centrality for a given holding can be used as an indicator of the corresponding risk of contagion. By identifying potentially highly-contagious nodes, network analyses can inform the prevention and control of infectious disease transmission. It is general wisdom that such infectious "hubs" should be targeted preferentially by the regulator (the public authorities enforcing health policy, which sometimes precludes such hubs from exchanging) in order to prevent and mitigate infectious disease outbreaks in exchange network systems Albert *et al.* [2000]; Barthelemy *et al.* [2004]; Madar *et al.* [2004].

While the implementation of trade restrictions on key large-size holdings could be effective in mitigating epidemic outbreaks, it is often prohibitively costly to regulators and has potentially severe economic impact on markets. Specifically, the disruption of core market players through intensive preventive measures can cause economic shock and render such strategies inappropriate Paton *et al.* [2009]. The promising alternative of combining evaluation of economic and epidemiological risks, however, remains a key gap in the literature Ceddia *et al.* [2013]

Because the agents that are central to the market are also likely to act as sources for epidemic spread, we expect that in general there will be a strong association between economic and epidemiological risks, some of which may be negative (inverse) associations. In this study, we aim at identifying efficient strategies for preventing epidemics with minimal disruption to markets and limited cost to the stakeholders such as regulators and business owners. We introduce a market-based categorisation for aggregating the holdings based on economic as well as structural network summaries, namely *position along the supply chain* and *market share or leadership*. To study the economic-epidemiological implications of our categorisation, we analyse two datasets recording cattle and swine livestock movement in France during 2005-2009 and 2010, respectively (Section 5.2). We show that the market categories that we propose describe livestock exchange intuitively and provide insights on the underlying trading patterns (Section 5.3.1). Using these categories, which are easy to implement because they are empirically defined, we evaluate the joint economic-epidemiological risk of epidemic outbreak and associated

regulatory measures (Section 5.3.2). We consider both the regulator's and the market's standpoints and evaluate the effectiveness of different preventive strategies that target agents in selected market categories (Section 5.3.3). Both static and dynamical preventive strategies are explored, whether based on real-time or past data. We conclude by summarising our most important results and by highlighting some perspectives for future work (Section 5.4).

The principal contribution of our study stems from adopting data-driven epidemiological and economic standpoints in order to evaluate control strategies against pathogen spread in livestock trade markets. Our study also contributes to the literature on network epidemiology by identifying how the choice of optimal outbreak control strategies may depend on the system considered (here cattle and swine livestock exchanges). Specifically, the optimal strategy identified does not necessarily rely on preferentially targeting hubs, despite the latter being often regarded as an evidently best approach.

5.2 Materials and Methods

5.2.1 Trade networks and livestock-exchange

Describing livestock-exchange from a market-centric perspective requires a preliminary exposition of core concepts, at the crossroad of economics and network theory. After a brief introduction to markets, we present the livestock-exchange data that we analyse as trade networks.

Understanding markets from a network-centric perspective

As a first approximation, a market can be formally described as a network composed of economic agents (e.g. individuals, businesses or sovereign States) in interaction [Rosenbaum, 2000; Goyal, 2009]. From a network perspective, an agent corresponds to a node or vertex interacting with other nodes through links or edges. From an economic perspective, an agent is an entity that pursues its own interests through some kind of economic optimisation. Agents have generally divergent interests resolved through exchanges and price definition [Callon, 1998].

Based on Moslonka-Lefebvre *et al.* [2013], we define some core concepts to describe markets and their influence on epidemics. A market is made of supplying agents, i.e. *suppliers*, and demanding agents, i.e. *demanders*. Agents that are both supplying and demanding correspond to *wholesalers*. Agents interact during *transactions* by exchanging goods that can lead to disease transmission, where a transaction is a delivery from a supplier to a demander. A disease-conducing contact corresponds to *trade flow*, which represents the number of products (e.g. animals) traded from a supplier to a demander per unit time. Provided that trade is the only route for transmission, trade flow can be interpreted as the epidemiological contact rate (number of transactions per time unit) weighted by contact intensity (number of products exchanged per transaction).

Since the exchange of animals occurs from suppliers to demanders but often not in the reverse

direction, and since different numbers of animals are shipped per transaction, we say that exchanges are *directed* and *weighted*. It follows that trade flow has a direction, e.g. we can dissociate *in-* and *out-*trade flow. We can also calculate the *total*-trade flow, i.e. the sum of the in- and out-trade flows. Hence, markets are described by directed and weighted networks. Moreover, since exchanges occur at precise points in time, they also form *dynamical* networks. Although more realistic, dynamical networks are harder to analyse than static networks (obtained, for instance, by aggregating interactions over time). Here, we consider both types of networks.

French livestock-exchange data described as trade networks

We analyse and compare trade networks derived from two datasets recording livestock-exchange in France : the BDNI for cattle (managed by the French ministry in charge of Agriculture) over years 2005-2009 and BDPorc (managed by the French professional union BDPorc) in 2010 for swine. Each dataset details movements of animals occurring in France among all economic agents involved in the supply chain, from strictly breeding farms to slaughterhouses with various categories of structural wholesalers in between (e.g. breeding-fattening farms, strictly fattening farms, dealers). Data on imports and exports are also available. Traceability is imposed by the regulator at different scales : on individual animals in the case of cattle, and on batches (sets of animals shipped from a seller to a demander during a transaction) in the case of swine. Hence, we extract individual-level transactions directly from the cattle dataset. Moreover, we reconstruct individual-level transactions for pigs based on a simple matching processe [see ESM Section B.2 of Moslonka-Lefebvre *et al.*, 2013, for details].

Our core network-based analyses are carried out at the microeconomic business-scale, i.e. agricultural holdings, as far as national livestock-exchange is concerned. The datasets allow us to distinguish three groups of holdings : *farms*, i.e. agents aiming at producing livestock ; *trading agents* such as assembling centers, i.e. agents aiming at exchanging livestock ; and the *rest of the world*, a single entity aggregating all agents located outside of France, which we use to assess the importance of international restrictions on trade, as e.g. might occurin the case of a major outbreak. Following the epidemiological literature [Noremark *et al.*, 2011], we neglect slaughterhouses and movements to slaughterhouses from our analyses since including these movements would underestimate the risk of transmission associated with farms. However, we do include foreign movements to and from France as they can contribute to disease introduction, further dispersion on large geographical scales and major economic disruptions. Though we explored several temporal descriptions of networks (see appendix D), all analyses presented in the main text are carried out on static networks aggregating transactions at the yearly scale for the sake of simplicity.

5.2.2 Market-centric categorisation of economic agents for representing the trade networks

We introduce a generic categorisation of economic agents applicable to a variety of markets including livestock exchange. We sort agents according to two types of market summaries : *position along the supply chain* and *market leadership*, and then use these categories to define *market categories*. Let $\mathcal{T} = [t_1, t_2]$ represent the period of time over which we aggregate the transactions. We then use these aggregates to calculate the following summaries.

Position along the supply chain : flow polarity

Position along the supply chain is given by the overall direction of trade flow, that we quantify by a summary referred to as *flow polarity* and denoted fp_a, and that is given for any agent a by the difference of its in- and out-trade flow divided by its total-trade flow over a particular time period $\mathcal{T}$:

$$[\text{flow polarity of agent } a](\mathcal{T}) = \mathrm{fp}_a(\mathcal{T}) = \frac{[\text{in-trade flow to } a](\mathcal{T}) \; - \; [\text{out-trade flow from } a](\mathcal{T})}{[\text{in-trade flow to } a](\mathcal{T}) \; + \; [\text{out-trade flow from } a](\mathcal{T})}, \quad (5.2)$$

where $[\text{in-trade flow to } a](\mathcal{T}) = \phi_a^{in}(\mathcal{T}) = \sum_i \phi_{ia}(\mathcal{T})$ and $[\text{out-trade flow from } a](\mathcal{T}) = \phi_a^{out}(\mathcal{T}) = \sum_j \phi_{aj}(\mathcal{T})$, with $\phi_{ij}(\mathcal{T})$ the trade flow from i to j over $\mathcal{T}$ and where sums are over all nodes exchanging with a over the same period. By construction, $-1 \leq \mathrm{fp}_a \leq 1$ (5.2) and fp_a can take any value between these two extremes. In order to build discrete classes of agents based on flow polarity, we introduce an empirical threshold $\epsilon > 0$ which can take either predetermined values or be set equal to percentiles of a given distribution. Hence, agents a for which $\mathrm{fp}_a < -\epsilon$, $\mathrm{fp}_a \in [-\epsilon, \epsilon]$ and $\mathrm{fp}_a > \epsilon$ correspond to *suppliers*, *wholesalers* and *demanders* respectively. Flow polarity is an extension of the concept of *hierarchy* [Pautasso *et al.*, 2010; Moslonka-Lefebvre *et al.*, 2012b] to weighted and dynamical networks.

Market leadership : flow share

Following marketing studies Kotler & Armstrong [2013], we use flow share, i.e. the relative trade flow, to quantify *market leadership*. For any agent a, flow share, denoted fs_a, is defined as its total-trade flow over time period $\mathcal{T}$ divided by the sum of total-trade flow for all agents over $\mathcal{T}$:

$$[\text{flow share of agent } a](\mathcal{T}) = \mathrm{fs}_a(\mathcal{T}) = \frac{[\text{in-trade flow to } a](\mathcal{T}) \; + \; [\text{out-trade flow from } a](\mathcal{T})}{[\text{total in-trade flow}] \; + \; [\text{total out-trade flow}](\mathcal{T})}, \quad (5.3)$$

where $[\text{total in-trade flow}] = \Phi^{in}(\mathcal{T}) = \sum_a \phi_a^{in}(\mathcal{T})$ and $[\text{total out-trade flow}] = \Phi^{out}(\mathcal{T}) = \sum_a \phi_a^{out}(\mathcal{T})$, with sums over all active agents over $\mathcal{T}$. Flow conservation implies that $\Phi^{in}(\mathcal{T}) = \Phi^{out}(\mathcal{T})$. By definition, $0 \leq \mathrm{fs}_a \leq 1$ (5.3) and $\sum_a \mathrm{fs}_a = 1$, where the sum is over all active agents in the market during $\mathcal{T}$. Similarly to flow polarity, we introduce two empirical thresholds $\delta_1, \delta_2 > 0$ which can take either predetermined values or be set equal to percentiles of a given distribution. Agents a for which $\mathrm{fs}_a < \delta_1$, $\mathrm{fs}_a \in [\delta_1, \delta_2]$

and $\mathrm{fp}_a > \delta_2$ are denoted *nichers*, *followers* and *leaders* respectively.

Definition of market categories using flow polarity and flow share summaries

Market categories are defined based on a two-dimensional indicator $(\mathrm{fp}_a, \mathrm{fs}_a)$, i.e. by the combination of position along the supply chain and market leadership : *suppliers-nichers (SN), suppliers-followers (SF), suppliers-leaders (SD), wholesalers-nichers (WN), wholesalers-followers (WF), wholesalers-leaders (WL), demanders-nichers (DN), demanders-followers (DF), and demanders-leaders (DL)*. In addition to the categorization in 3×3 classes (with respect to the above definitions for fp_a and fs_a), finer grids can be adopted for more detailed analysis.

5.2.3 Elaboration, choice, and evaluation of targeted control strategies

Based on the categorisation of agents that we have introduced, we consider preventive strategies that involve preferential targeting of agents belonging to certain market categories. We evaluate generic forms of interventions for outbreak control. Specific practical examples of these interventions include preferential surveillance and vaccination of the agents that are deemed most at risk. We proceed in three steps : firstly, we elaborate a general class of strategies preferentially targeting agents belonging to specific market categories ; secondly, we identify meaningful targeting strategies by assessing which agents are most at risk according to network-based summaries of economic and epidemiological risks ; thirdly, we evaluate indirectly, from an economic-epidemiological perspective, how selected strategies influence systemic risk.

Preferentially targeting of agents in specific market categories

Let N denote the number of agents involved in at least one trade event during the time interval $\mathcal{T}$. The fraction $F_n = n/N$ of n agents to be targeted is chosen according to strategy $\mathcal{S}$ (scenario denoted $(F_n; \mathcal{S})$), which ranks each agent a from 1 to N according to the decreasing values of a rank function, denoted $R(\mathrm{fp}_a, \mathrm{fs}_a)$, and based on the market categories as defined by fp and fs. We define $R(\mathrm{fp}_a, \mathrm{fs}_a)$ as the product of functions $R_{\mathrm{fp}}(\mathrm{fp}_a)$ and $R_{\mathrm{fs}}(\mathrm{fs}_a)$:

$$\begin{aligned} R(\mathrm{fp}_a, \mathrm{fs}_a) &= R_{\mathrm{fp}}(\mathrm{fp}_a)\, R_{\mathrm{fs}}(\mathrm{fs}_a)\ , \text{with} \\ R_{\mathrm{fp}}(\mathrm{fp}_a) &= [2 - (\mathrm{fp}_a + 1)]^{z^{suppliers}}\, [\mathrm{fp}_a + 1]^{z^{demanders}}\ , \\ R_{\mathrm{fs}}(\mathrm{fs}_a) &= [\max \mathrm{fs} - \mathrm{fs}_a]^{z^{nichers}}\, [\mathrm{fs}_a]^{z^{leaders}}\ , \end{aligned} \tag{5.4}$$

where $z^{suppliers}$, $z^{demanders}$, $z^{nichers}$ and $z^{leaders} \geq 0$ are fixed parameters representing the preferences of the regulator for targeting specific market categories. As an example, wholesaling leaders are surveyed preferentially when we set : $z^{suppliers} = z^{demanders} > 0$, $z^{nichers} = 0$ and $z^{leaders} > 0$. In the case when two or more agents take the same value of R, we choose their relative orders uniformly at random.

Note that the variability resulting from this ranked ordering is weak, as confirmed by assessments of each strategy over 100 random replicate targeting for both datasets.

Identification of specific targeted control strategies

To choose meaningful preventive strategies, i.e. to set *a priori* appropriate values for $z^{suppliers}$, $z^{demanders}$, $z^{nichers}$ and $z^{leaders}$ in (5.4), we calculate various risk indicators per agent per market category. We specify for each indicator whether it quantifies an *economic risk*, an *epidemiological risk* or an *economic-epidemiological* risk. An *economic risk* is the risk of market disruptions caused by the failure of an agent. An *epidemiological risk* is the risk for an healthy agent to get contaminated and/or the risk of an infected agent to further transmit an infection to other agents. An *economic-epidemiological risk* is a combined risk. In identifying optimal interventions, we are specially interested in strategies that minimise both economic risk and epidemiological risk.

First, we consider three risk indicators at the agent level : flow polarity, flow share and trade flow (defined in sections 5.2.1 and 5.2.2). We consider flow polarity is a measure of joint economic-epidemiological risk, flow share constitutes a measure of economic risk and trade flow quantifies epidemiological risk (although the relationships between these measures are subject to economic and epidemiological reinterpretation) [Kellermann, 2011; Battiston *et al.*, 2012; Moslonka-Lefebvre *et al.*, 2013]. We also calculate, for each market category, the values of two additional indicators of epidemiological risk that are well-documented in the network epidemiology literature : the *proportion of agents belonging to the largest strongly connected component (LSCC)* and the *average betweenness centrality.* The LSCC is the largest subnetwork of agents for which a directed path exists from any other agent in the subnetwork. The betweeness centrality of a node is the fraction of shortest paths that passes through this node (see section D.1.1 for details).

Evaluation of the targeted control strategies using multi-criteria decision analyses (MCDA)

To evaluate targeting strategies, we carry out multi-criteria decision analyses (MCDA) from an economic-epidemiological perspective. The MCDA aim at finding optimal strategies for reaching one or multiple objectives with minimal efforts, potentially considering multiple objectives and types of effort simultaneously Mourits *et al.* [2010]. The capacity of a strategy for mitigating a disease is measured through *prevention-effectiveness* criteria, while the effort needed for reaching a given effectiveness is measured through *prevention-cost* criteria, where all criteria are to be defined. An optimal strategy is one that maximises prevention-effectivenes with minimal efforts. The strategies considered in this study are implemented by the regulator who decides to target certain agents preferentially. In practice, the cost of implementing control measures may fall on the regulator, the business owners, or other stakeholders of the market. However, the perception of what is an optimal strategy is subjective. We hence consider two complementary points of view when comparing control strategies : the regulator's

and the market's.

***Prevention-effectiveness* : derivation from the LSCC.** We define prevention-effectiveness as the benefit to the market stakeholders of implementing a preventive strategy, for instance the losses averted by avoiding an epidemic. The proportion of agents belonging to the LSCC is a standard epidemiological proxy to assess both the probability of disease invasion and the epidemic final size Meyers *et al.* [2006]; Rautureau *et al.* [2012]. Following node-percolation experiments Albert *et al.* [2000], we measure the prevention-effectiveness of a given strategy $\mathcal{S}$ when targeting a proportion F_n of agents by the *relative decrease in the LSCC size* of the network aggregated over the time period $\mathcal{T}$:

$$\begin{aligned}[\text{prevention-effectiveness}](F_n;\mathcal{S})(\mathcal{T}) &= [\text{relative decrease in the LSCC size}](F_n;\mathcal{S})(\mathcal{T}) \ , \\ &= 1 - \frac{[\text{number of agents remaining in the LSCC}](F_n;\mathcal{S})(\mathcal{T})}{[\text{number of agents in the LSCC without prevention}](\mathcal{T})} \end{aligned} \quad (5.5)$$

To assess the robustness of our findings based on the LSCC, we also use, in appendix D, the average infection chain as an alternative proxy for prevention-effectiveness Dube *et al.* [2008]; Noremark *et al.* [2011]; Lal Dutta *et al.* [2014]. In contrast with the LSCC, the average infection chain has the advantage of taking into account the sequence of dates when exchanges occur over time period $\mathcal{T}$ (and hence of considering only temporally-compatible paths between agents). However, due to considerable computation time required, calculation of the average infection chain is only performed for rather small networks. We also explore the influence of postponing the implementation of a given strategy $\mathcal{S}$ (see section D.1.2 for more details).

***Prevention-costs* : relative costs to the regulator and market distortions.** We define multiple prevention-costs, as the *relative costs to the regulator* and consequent *market distortions*. *Relative costs to the regulator* are costs incurred by the regulator when e.g. implementing a preventive control strategy. *Market distortions* are potential damaging impacts on the market that result directly or indirectly from implementation of a preventive control strategy.

First, we detail relative costs to the regulator. When implementing a strategy $\mathcal{S}$ over time period $\mathcal{T}$, we assume the regulator may incur two types of costs : an *agent-cost* and a *flow-cost*. In common with others [Albert *et al.*, 2000], though explicitly rather than implicitly, we assume the agent-cost is proportional to F_n for any strategy $\mathcal{S}$. The flow cost is also expected to increase with F_n but to be highly variable depending on the underlying strategy $\mathcal{S}$. Here, in line with Moslonka-Lefebvre *et al.* [2013], we assume the flow-cost is proportional to $(\Sigma_a \text{fs}_a)(F_n;\mathcal{S})\ (\mathcal{T})$, which is the total flow share of the fraction F_n of agents targeted in strategy $\mathcal{S}$. In practice, since we are interested in ranking a set of strategies, we measure *relative rather than absolute costs* : we hence directly track F_n and $(\Sigma_a \text{fs}_a)(F_n;\mathcal{S})(\mathcal{T})$, which have support in $[0,1]$, as proxies for economic risk.

Second, we specify market distortions. Any preventive strategy implemented by the regulator is ex-

pected to cause market disruptions. Here, by analogy with our market categories, we consider two types of economic risk proxies to measure disruptions to the market : *disruption to the overall flow polarity* and *disruption to the overall flow share. Disruption to the overall flow polarity*, denoted $\mu(F_n;\mathcal{S})(\mathcal{T})$, is measured as the relative mismatch between overall in- and out-flow :

$$\mu(F_n;\mathcal{S})(\mathcal{T}) = \left| \frac{\left(\Phi^{in}(F_0;\mathcal{S}) - \Phi^{in}(F_n;\mathcal{S})\right)(\mathcal{T}) - \left(\Phi^{out}(F_0;\mathcal{S}) - \Phi^{out}(F_n;\mathcal{S})\right)(\mathcal{T})}{\left(\Phi^{in}(F_0;\mathcal{S}) - \Phi^{in}(F_n;\mathcal{S})\right)(\mathcal{T}) + \left(\Phi^{out}(F_0;\mathcal{S}) - \Phi^{out}(F_n;\mathcal{S})\right)(\mathcal{T})} \right| , \qquad (5.6)$$

where $\Phi^{in}(F_k;\mathcal{S})(\mathcal{T})$ and $\Phi^{out}(F_k;\mathcal{S})(\mathcal{T})$ are the total in- and out- flow distorted when a fraction F_k of agents are targeted according to strategy $\mathcal{S}$. As specified in 5.2.2, flow conservation implies $\Phi^{in}(F_0;\mathcal{S})(\mathcal{T}) = \Phi^{out}(F_0;\mathcal{S})(\mathcal{T})$. By construction, $\mu(F_n;\mathcal{S})(\mathcal{T})$ takes values in $[0,1]$. *Disruption to the overall flow share* is assessed by $(\Sigma_a \mathrm{fs}_a)(F_n;\mathcal{S})(\mathcal{T})$, the total flow share targeted and hence potentially disrupted when a fraction F_n of agents is targeted according to strategy $\mathcal{S}$. Notice that the same value of $(\Sigma_a \mathrm{fs}_a)(F_n;\mathcal{S})(\mathcal{T})$ is used as a proxy for two differing quantities : flow-cost (incurred by the regulator while implementing a strategy) and disruption to the overall flow share (which reflects a particular market disruption induced by the regulator's intervention).

Practical implementation of the MCDA. Multiple criteria decision analyses are carried out over time period $\mathcal{T}$, which we set to start in 2009 for cattle and 2010 for swine. To account for potential delays in the collection of data necessary to calculate flow polarity fp_a (5.2) and flow share fs_a (5.3), we explore two contrasting cases : the *real-time scenario* or the *deferred scenario*, respectively, where the regulator has access to real-time data or deferred data, respectively, so fp_a and fs_a can be calculated based on $\mathcal{T}$ or based on $\mathcal{T} - \delta t$ (with δt representing the delay of data collection), respectively. We set $\delta t = 1$ year, a deliberately large value for data collection. For each strategy explored, the relative decrease in the LSCC size is evaluated at increasing values of F_n. For each value of F_n, we keep track of the proxy for prevention-effectiveness (the relative decrease in the LSCC size (5.5)), and of the four proxies for prevention-costs : the two relative costs to the regulator (relative agent-cost F_n and flow-cost $(\Sigma_a \mathrm{fs}_a)(F_n;\mathcal{S})(\mathcal{T})$, where $\propto$ denotes proportionality) ; and two measures of market distortions (disruption to the overall flow polarity $\mu(F_n;\mathcal{S})(\mathcal{T})$ (5.6) and disruption to the overall flow share $(\Sigma_a \mathrm{fs}_a)(F_n;\mathcal{S})(\mathcal{T})$).

5.3 Results and Discussion

5.3.1 Analyses and representation of trade networks using market categories

Considering the French datasets for livestock-movements of cattle and swine, we categorise agents with similar market characteristics, according to position along the supply chain, quantified by flow polarity (Fig 5.1A,C), and market leadership, quantified by flow share (Fig 5.1B,D). Agents with similar

ranks in flow polarity and flow share belong to the same market category (Fig 5.2A). Since market categories are relatively invariant over time (section D.2.1, Fig. D.1), we focus our subsequent analyses based on 2009 for cattle and 2010 for swine.

We notice suppliers and demanders are overly-represented in the cattle and swine markets respectively (Fig 5.1A). Since demanders act as epidemiological dead-ends, our analyses suggest that the cattle market is riskier than the swine market, a result in agreement with Rautureau *et al.* [2012]. According to the distribution of flow share, the swine market is less scattered than the cattle market (Fig 5.1B). For instance, the median flow share is $10^{-4.4}$ in swine in 2010 and only $10^{-6.0}$ in cattle in 2009, i.e. the median flow share in the swine market is 40 times larger than in the cattle market. This indicates a larger economic risk associated with the failure of a typical swine agent compared with the failure of a typical cattle agent.

To shed light on the practical meaning of flow polarity and flow share, we analyse the correspondence between standard market groups and our market categories (Fig 5.1C-D). For cattle, we remark that farms and trading agents correspond to suppliers and wholesalers, while for swine, farms and trading agents rather correspond to wholesalers-to-demanders (Fig 5.1C). For both cattle and swine, French exchanges with the rest of the world essentially correspond to exports, i.e. the rest of the world acts as a strict demander. Concerning flow share for cattle and swine, farms and trading agents correspond to nichers-to-followers and followers-to-leaders respectively (Fig 5.1D). The rest of the world represents a considerable flow share (about 13% for cattle and 4% for swine) and can be described as a major market leader for both trade systems.

We compare the cattle and swine markets using the market categories defined in Fig 5.2A. In the cattle market, the proportion of leaders is small, irrespective to the flow polarity. The most represented market categories are SN and SF, with no significant difference between nichers and followers. In the swine market, the largest proportions of agents are in the SF and DF market categories, illustrating the fact that followers exhibit a polarised activity most of the time. These distributions are modified when scrutinising the Largest Strongly Connected Component (LSCC). Indeed, for both cattle and swine, the proportion of agents belonging to the LSCC increases with increasing flow share (Fig 5.2B), implying that the epidemiological risk associated with leaders is larger than that associated with nichers. This trend is confirmed by the distribution of average betweenness centrality amongst market categories (section D.2.2, Fig. D.2). For a given flow share, both risk indicators, i.e. the LSCC and betweeness centrality, typically have larger values for agents with negligible flow polarity, which suggests that wholesalers are probably stronger epidemiological drivers than suppliers, a finding in agreement with the theoretical results reported in Pautasso *et al.* [2010]; Moslonka-Lefebvre *et al.* [2012b]. We also notice, in line with Rautureau *et al.* [2012], that the proportion of agents belonging to the LSCC is larger in cattle than in swine, which suggest there is greater epidemiological risk associated with trade in cattle markets.

The use of market categories also enables comparison of connection patterns in the cattle and swine markets (Fig. 5.3). In the cattle market, the total trade flow is relatively large for the WF and WL categories. Also, exchanges to and from wholesalers (irrespective of flow polarity) and within categories WF and WL are clearly over-represented in this market, thus generating structural loops and potentially infectious feedbacks. In contrast, there is a larger number of exchanges in the swine market than in the cattle market that involve direct transactions from suppliers to demanders, particularly from SF to DF, which leads to a trading structure with a limited number of potentially infectious feedbacks. Again, this result suggests that the swine market is less prone to epidemic spread compared with the cattle market.

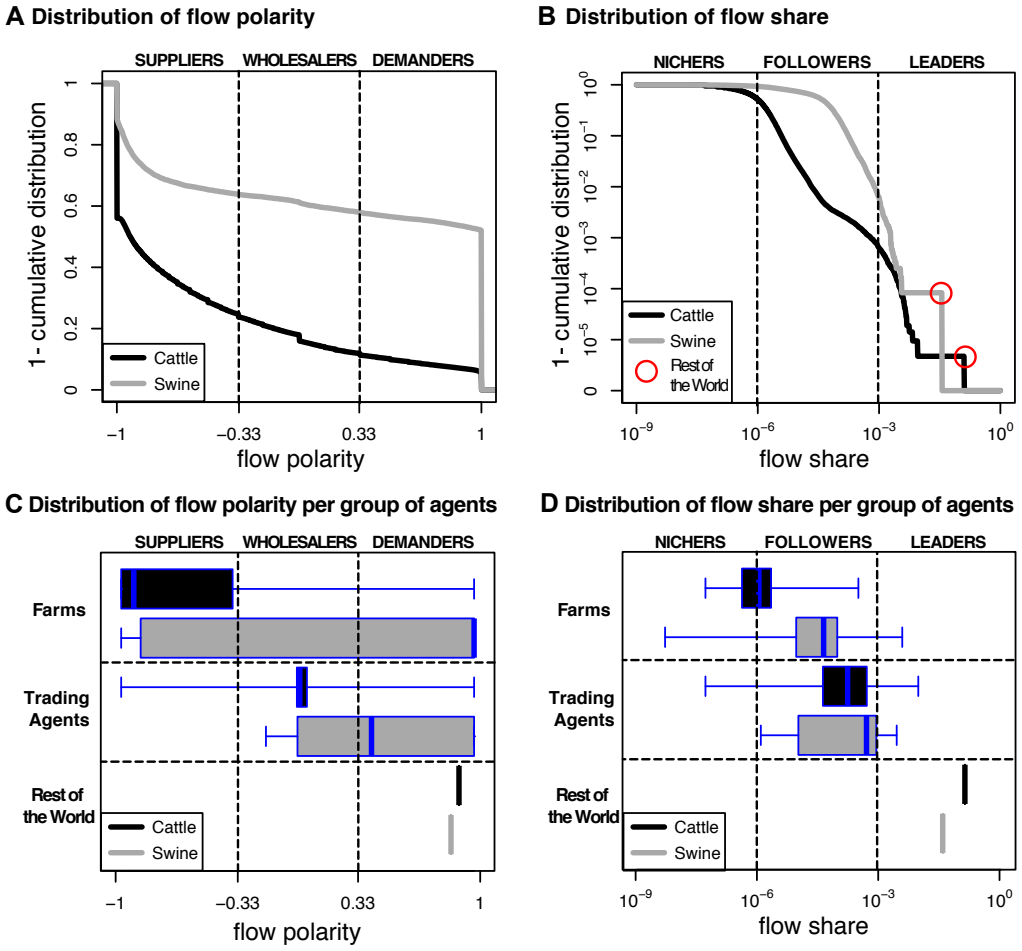


FIGURE 5.1 – **Categorisation of economic agents : flow polarity and flow share. A)-B)** Distributions of flow polarity and flow share in the cattle and swine markets. Flow polarity and flow share of an economic agent quantify its *position along the supply chain* and *market leadership* respectively. Qualitatively, flow polarity (5.2) can be used to define *supplier*, *wholesaler* and *demander*, categories of agents corresponding to values that are negative, symmetrically distributed about zero, and positive, respectively. Similarly, flow share (5.3) can be used to define *nicher* (low values), *follower* (mid values) and *leader* (high values). **C)-D)** Distributions of flow polarity and flow share per group of agents in the cattle and swine markets. Groups are either sets of French *farms* or *trading agents*, or a single entity aggregating all agents located outside of France, namely the *rest of the world*. Flow polarity and flow share are calculated over year 2009 for cattle and throughout year 2010 for swine.

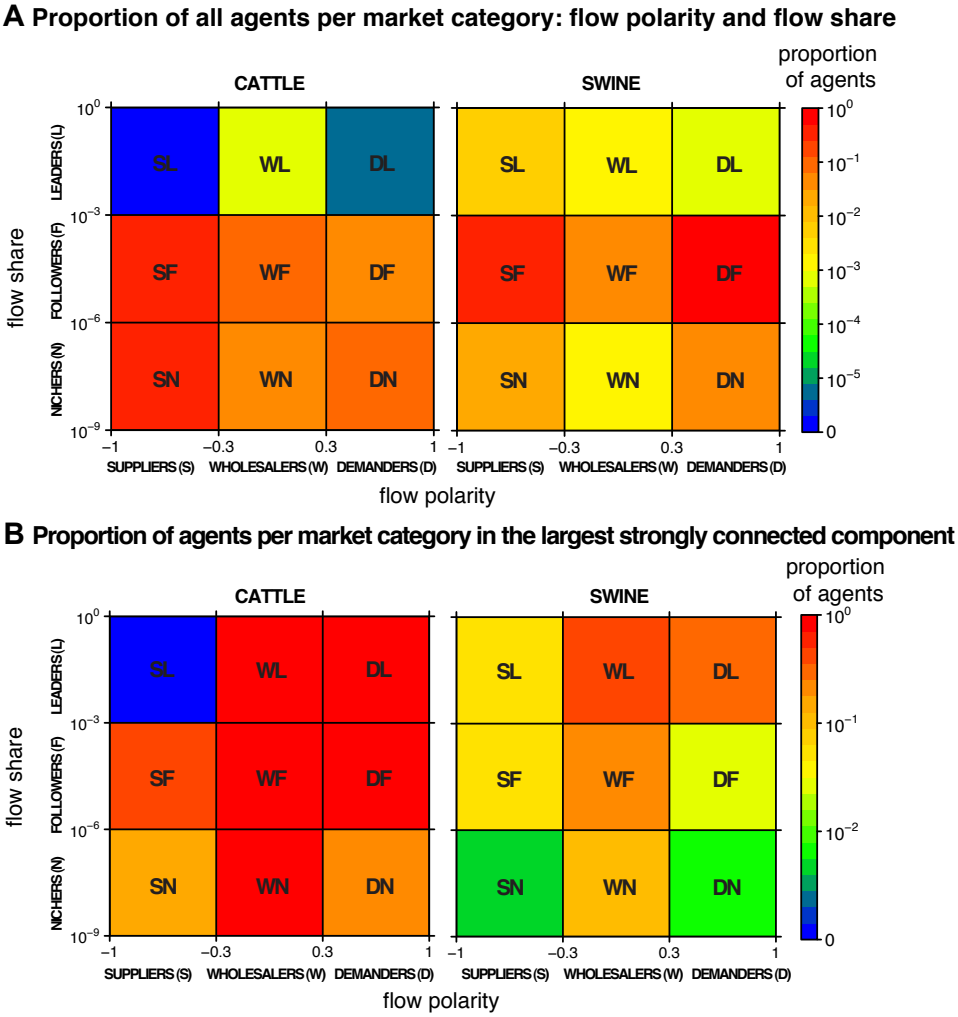


FIGURE 5.2 – **Market categories defined according to flow polarity and flow share and used to assess joint economic-epidemiological risk. A**) Proportion of agents in the whole population that are in each *market category*, i.e. agents with the given ranges of flow polarity and flow share. **B**) Proportion of agents in the largest strongly connected component (LSCC) that are in each market category. The LSCC is a proxy for both the probability of an outbreak and the epidemic final size. The risk of outbreak increases as the number of agents that are in the LSCC increases. To ease interpretation, flow polarity and flow share are discretised on a 3×3 grid, leading to a total of 9 market categories. The corresponding non-discretised marginal distributions of flow polarity and flow share for the cattle and swine markets are available in (Fig 5.2A). Flow polarity and flow share are calculated over year 2009 for cattle and year 2010 for swine.

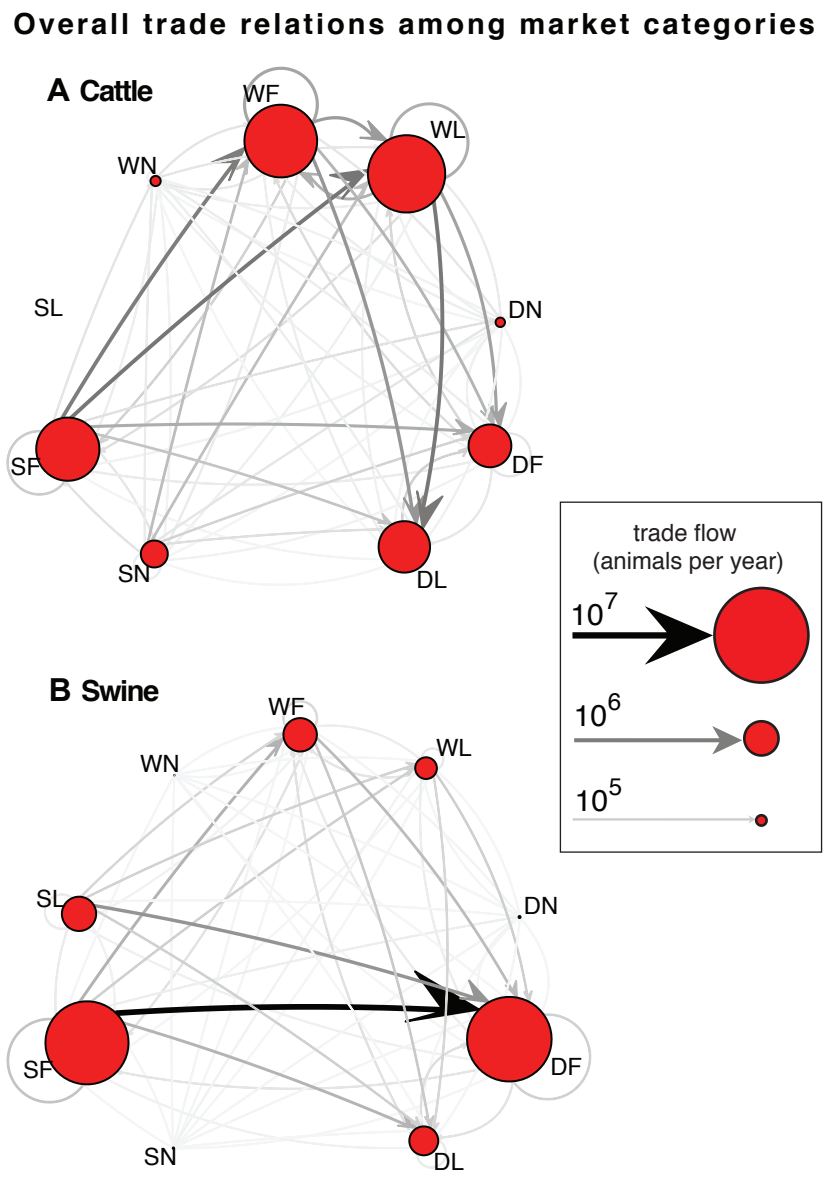


FIGURE 5.3 – **Trade networks of the cattle and swine livestock markets A)-B)** Data are yearly aggregated (2009 for cattle and in 2010 for swine). Each node encompasses all the agents in each market category, as defined on the 3-by-3 grid in Fig. 5.2. The size of each node is proportional to the yearly aggregated trade flow summed up over all agents in each market category. Widths and colors intensity of directed links (arrows) are proportional to corresponding yearly aggregated trade flow from / to each aggregated category, summed up over all the agents in the categories involved in in- or out- movements.

5.3.2 Identification of targeted control strategies based on risk assessment per agent per market category

Particular preventive strategies are selected from preliminary risk assessments (5.2.3). By definition, trade flow per capita increases with flow share for both cattle and swine, implying a larger economic-epidemiological risk per capita associated with leaders compared with nichers. Flow polarity does not seem to influence trade flow significantly, especially in cattle (Fig. 5.3). Also note that in cattle, wholesalers have comparatively larger trade flows than the other categories. This outcome suggests

that wholesalers are stronger epidemiological drivers than suppliers, a finding in agreement with the theoretical results reported in Pautasso *et al.* [2010]; Moslonka-Lefebvre *et al.* [2012b].

At the scale of market categories, WL are more connected (Fig. 5.3), have a larger betweenness centrality (section D.2.2) and are more likely to belong to the largest strongly connected component than SN (Fig. 5.2). This latter point renders WL representative both in the cattle and in the swine markets, although for the swine market WL agents are not involved in large volumes of trade (Fig. 5.3B). The swine market is driven mainly by trade flows from SF and SL to DF, which suggests that the epidemiological risk could be confined to only a few market categories. We therefore expect the cattle market to be at a greater epidemiological risk compared with the swine market, a finding in agreement with the results reported in Section 5.3.1 and in a previous study [Moslonka-Lefebvre *et al.*, 2013].

Taken together, our results corroborate, in agreement with the literature, that wholesalers-leaders (WL) appear to act as infectious super-spreaders. WL, as market leaders, are also associated with major economic risks in case of failure. Like WL, suppliers-nichers (SN) can act as infection sources and are associated with epidemiological risk. However, in contrast with WL, SN have minor market importance and are less likely to induce market disruptions when disturbed. From a network perspective, WL (SN) can therefore be described as "hubs" ("anti-hubs"), i.e. agents with a large (low) number of links compared with the average number of links per agent [Barrat *et al.*, 2004a]. We therefore evaluate two contrasting strategies : the preferential targeting of hubs, i.e. WL, and the preferential targeting of anti-hubs, i.e. SN. The strategies aiming at targeting WL first (SN first) are referred to as *the WL strategies* (the *SN strategies*). In practice, we set $z^{suppliers} = z^{demanders} = 1$, $z^{nichers} = 0$ and $z^{leaders} = 1$ in (5.4) for the WL strategies, and $z^{suppliers} = 1$, $z^{demanders} = 0$, $z^{nichers} = 1$ and $z^{leaders} = 0$ in (5.4) for the SN strategies.

5.3.3 Evaluation of the targeted control strategies using MCDA

Based on the results from the risk assessment (5.3.2), our multiple-criteria decision analyses (MCDA) focus on strategies aiming at targeting preferentially WL agents compared with strategies preferentially targeting SN agents. The relative performances of these strategies are compared on both markets using various criteria (as defined in 5.2.3).

We start by introducing results from the regulator's point of view (i.e. quantifying agent-costs and flow-costs). WL strategies appear always to be more effective than SN strategies provided that the overall prevention cost is driven by the fraction of agents targeted (plain black curves in Fig. 5.4A,C, 5.5A,C). Under these conditions, we recover the commonly accepted wisdom that preferentially targeting the most central nodes in a heterogeneous network is the best way to mitigate an epidemic [Albert *et al.*, 2000], where centrality of the nodes is determined by having large betweeness centrality and probability of belonging to the LSCC.

However, if the overall prevention cost is driven instead by the fraction of targeted flow, the SN stra-

tegies can be more efficacious than the WL strategies as seen in cattle (plain grey curves in Fig. 5.4A,C) but not in swine (plain grey curves in Fig. 5.5A,C). An example of a strategy driven by the fraction of targeted flow (total animals to be protected per time unit) would be an initiative whereby diagnostic tests would be distributed to farmers purchasing livestock in order to test biological samples from the animals purchased. This apparently counterintuitive result stems from the structure of the cattle network (Fig. 5.3A). and may not be so surprising : when applying tests at purchase, it may be better to cover both a large geographical and topological space (the "area" covered by a network), i.e. to dispatch a constant number of tests to a very large number of premises with very small flow shares, rather than to a very small number of holdings with very large flow shares.

From the market's point of view, when targeting agents to prevent epidemics, the regulator will necessarily induce distortions to the market (Fig. 5.4B,D, 5.5B,D). Typical disruptions of the market include shifts in price, removal or later reintroduction of suppliers and demanders and/or of local and global depletions in supply and demand stocks. Disruptions can be induced by infection and subsequent eradication of contaminated stock for sanitation or other preventive measures [Moslonka-Lefebvre *et al.*, 2013]. At first sight, it seems impossible to attain an optimal situation where epidemics are at low risk without affecting the market. Some strategies appear better than others though : while the SN strategies can induce fewer distortions than the WL strategies for most levels of prevention-effectiveness in cattle (Fig. 5.4B,D), the WL strategies are always the best in swine (Fig. 5.5B,D). These results suggest that cattle and swine markets, while both corresponding to heterogeneous livestock-exchange networks, require differing preventive measures.

Introduction of delays in the collection of data to design preventive strategies has little effect on our results (dashed versus plain curves in Fig. 5.4-5.5). Our conclusions are also not affected by the use of an alternative measurement of prevention-effectiveness with an epidemiological risk proxy accounting for the time-varying nature of the network (i.e. the succession of transactions and hence of network links over time) nor by the inclusion of time lags in the implementation of preventive strategies (section D.2.3).

Taken together, our results suggest that the WL strategies are not always the best. In particular for the cattle market, when costs of prevention are driven by the number of animals to target, the SN strategies perform better than the WL strategies.

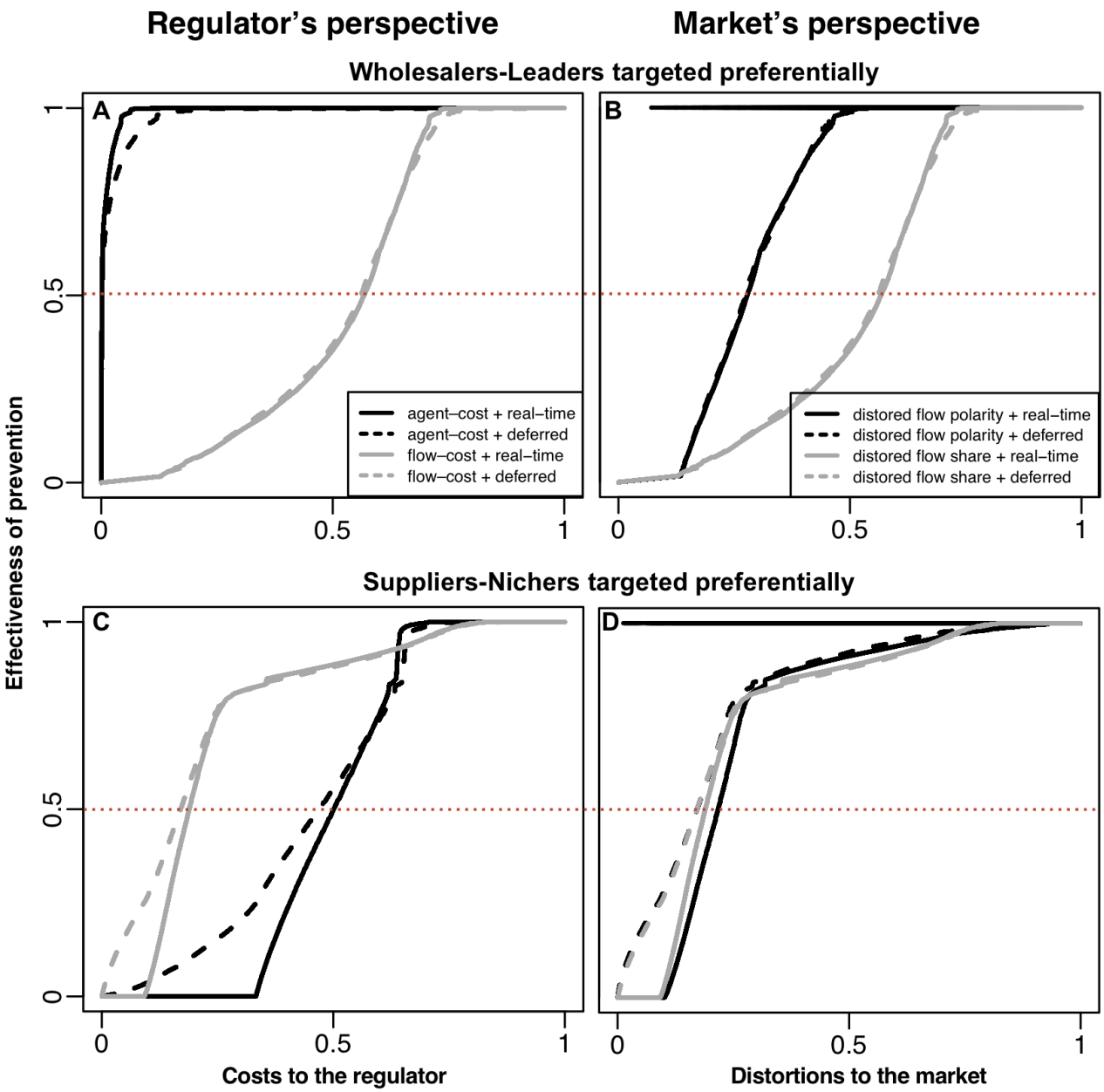


FIGURE 5.4 – **Multiple-criteria decision analyses (MCDA) of contrasting targeted control strategies in the cattle livestock market.** MCDA of strategies targeting wholesalers-leaders first (**A**-**B**, with $z^{suppliers} = 1$; $z^{demanders} = 1$; $z^{nichers} = 0$; $z^{leaders} = 1$ in (5.4)) and suppliers-nichers first (**C**-**D**, with $z^{suppliers} = 1$; $z^{demanders} = 0$; $z^{nichers} = 1$; $z^{leaders} = 0$ in (5.4)). For each strategy, the x-axis quantifies prevention-costs, i.e. the relative costs to the regulator (**A**-**C**) and relative distortions to the market (**B**-**D**) to reach a given prevention-effectiveness against epidemics as depicted on the y-axis (e.g. the red dotted lines to reach 50% of prevention-effectiveness). Prevention-effectiveness is measured by the proportion of agents removed from the largest strongly connected component (5.5). Preventive strategies in the cattle market are implemented over year 2009. Market categories are defined either over 2009 (real-time information available on agents, plain curves) or over 2008 (deferred information available on agents, dashed curves). Each case corresponds to 100 replicate simulations (notice the weak variability).

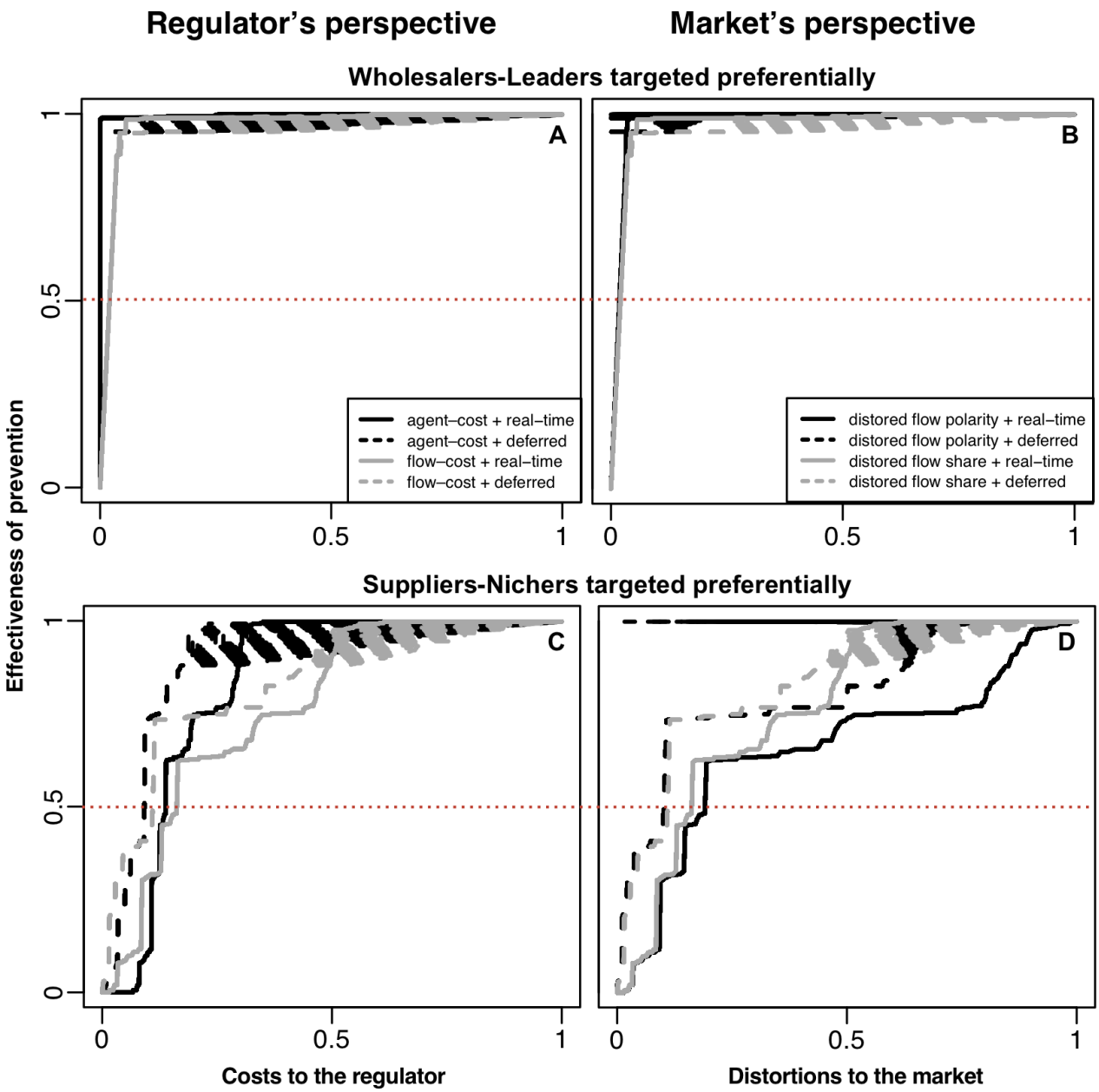


FIGURE 5.5 – **MCDA of contrasting targeted control strategies in the swine livestock market.** Targeting strategies in the swine market are implemented over year 2010. Market categories are defined either over year 2010 (real-time information available on agents, plain curves) or over January 2010 (deferred information available on agents, dashed curves). The rest of the legend is identical to Fig. 5.4.

5.4 Conclusion

Regulators tasked with managing disease outbreaks are generally constrained by limited resources Klein *et al.* [2007]. Therefore, prioritising interventions under limited resources is essential to achieve effective prevention and control of epidemic outbreaks. Typical targeted preventive measures include vaccination of risky agents, and risk-based surveillance such as blood-tests for identifying cryptic infectious cases, in particular for purchased animals in livestock exchange markets. Targeted prevention is particularly relevant for a national regulator aiming at eradicating or controlling a disease in order to generate subsequent commercial benefits or to maintain a disease-free status to satisfy the requirements of the international animal health code. Acquiring or maintaining an internationally-recognised disease-free status is associated with major benefits such as capacity to export livestock and reduction of control burdens. For instance, according to the French ministry in charge of Agriculture, surveillance measures of bovine tuberculosis, including on-farm visits and systematic animal testing, cost the French State as much as 20 million euros over the period 2010-2011 [MAAF, 2013]. In this context, our analyses suggest that while the risk of epidemic introduction in France due to contaminated livestock-imports appears limited, the economic risk associated with potential sanitary bans on French exports is important and could lead to major market disruptions.

While we build on an already rich literature that applies network analyses to inform health policies, in particular for the animal sector, we depart from existing studies by introducing a market-based categorisation to analyse and protect trade networks propagating epidemics. Our market-based categorisation, which we have found to be relatively stable over time, can intuitively describe market structure and interaction mechanisms. It can also be used to quantify joint economic-epidemiological risks, and hence to evaluate prevention strategies that target particular market categories, thereby concentrating resource application to confined sectors of the system at risk. In particular, when both the standpoints of the regulator and of the market are taken into account, we find that targeting preferentially suppliers-nichers, which are anti-hubs, can, in some cases, outperform the preferential targeting of wholesalers-leaders, i.e. hubs. The preferential targeting of hubs appears to be systematically more effective when we only consider the regulator's point of view and assume that intervention costs are proportional to the number of economic agents to be protected. In summary, our study suggests that multiple perspectives should be adopted when evaluating targeted preventive strategies, a finding with general implications for epidemiological and ecological studies aiming at prioritising interventions for maintaining healthy and diverse (eco-)systems.

Achieving the best epidemiological outcome under a constrained regulatory budget has been addressed by others, for example, through optimal control theory Rowthorn *et al.* [2009]. A typical objective of optimal control theory consists in finding an optimal amount of treatment at each time step to minimise the total number of infected animals during the course of an epidemic without exceeding a

fixed budget. However, prior to the centric-anayses introduced here, the influence of epidemics and subsequent regulatory measures on market functioning at the microeconomic scale were largely unknown Moslonka-Lefebvre *et al.* [2013]. Although we do not consider here a coupled dynamic model of infectious disease and economics dynamics , our study constitutes a first step towards understanding the likely impacts of epidemics on trade. At the interface between the data-motivated approach adopted here and the proof-of-concept approach exposed in Moslonka-Lefebvre *et al.* [2013], the elaboration of an agent-based, economic-epidemiological model integrating temporal feedbacks will be the subject of future work. While we have focused our applications to animal health policy, our empirical formulation to identify market categories can aid the analysis of highly complex networks with multiple node types, and directed, weighted and dynamical links. We believe that the framework we have proposed can provide wider valuable insights to uncover the mechanisms underpinning joint disease and exchange dynamics.

Data accessibility

French livestock- and swine-exchange data were respectively provided by the French Ministry in charge of Agriculture (FMA) (contact point : bicma.sdspa.dgal@agriculture.gouv.fr) and the professional union BDPorc (contact point : administrateur@bdporc.fr), which operates under contract for the FMA. Data collection and analyses are subject to a confidentiality agreement available upon request from the designated contact points.

Competing interests

The authors have no competing interests.

Authors' contributions

MML designed the study, carried out data analysis and interpretation, and drafted the manuscript ; CAG and HM coordinated the study and helped draft the manuscript ; CB and PE helped draft the manuscript ; JANF and EV designed the study, participated in data analysis and interpretation, and helped draft the manuscript. All authors contributed to data processing and handling, and gave final approval for publication.

Acknowledgments

Many thanks to Julien Fosse, Etienne Geoffroy, Bhagat Lal Dutta, François Moutou, Emilie Moyne and Stéphane Robin for helpful comments and insights. We are grateful to the French Ministry in

charge of Agriculture and to the professional union BDPorc for granting us access to the cattle and swine datasets respectively.

Funding

EV, HM and MML would like to thank the French Ministries in charge of Agriculture and Environment, the INRA MIA Department for financial and operational support. MML benefited from a Dufrenoy grant from the French Academy of Agriculture. JANF and CAG were funded by DEFRA and USDA. CB, EV, and PE acknowledge financial support provided by the French Research Agency, program Investments for the future, project ANR-10-BINF-07 (MIHMES) and by the European fund for the regional development (FEDER) of Région Pays-de-la-Loire. EV and PE acknowledge financial support provided by the metaprogramme GISA of INRA, project PrediCatT.

Troisième partie

. . .to epidemic models with adaptive behaviour and constrained interactions

Foreword to part III

Interaction constraints faced by agents exert a fundamental influence on epidemic dynamics (part II). Interaction constraints arise because exchanges are costly and resources are limited, implying an agent cannot interact continuously and uniformly with every other agents. But constraints do not explain *how* exchanges occur in the first place. In addition, any crisis impacting exchange processes can possibly modify interaction between agents in an adaptive-fashion.

To ease the study of interaction constraints and subsequent infection processes, we assumed so far agents were passive, which is clearly at odds with available data and common sense (section 2.2). Rather, exchanges are driven by multiple motivations, and exchange-driven epidemics not only results in biological and economic damages to contaminated agents, but can also alter disease-conducive exchanges through complex behavioural responses. For instance, when agents face an epidemic burden, they can take individual and collective protective measures. As an example, some farm managers can spontaneously decide to vaccinate their herds as a measure of epidemic prevention. If the resulting vaccination coverage is insufficient, national and international authorities can generate incentives or enforce regulation to increase the overall number of vaccinated herds [Klepac *et al.*, 2013]. Epidemiological models increasingly account for such risk-aversion behaviour in response to pathogen outbreaks [Funk *et al.*, 2010]. But agents' behaviours do not boil down to risk-aversion due to complex interdependencies involving adaptation of social or economic mechanisms, such as those observed in infection-conducing markets. When trade is disrupted by a major outbreak resulting in an economic crisis, some agents may be tempted to take additional risks for making extra profit or simply to compensate the losses incurred by the incapacitation of their trading partners due to infection and subsequent sanitation measures.

In this part, we aim at shedding light on joint exchange and epidemic dynamics. We hence relax the assumption of passive agents and present cases where epidemics can also disrupt the disease-conducive exchange networks by affecting adaptive mechanisms governing exchanges. More precisely, we account for complex adaptive behaviour encountered in markets, which we integrate into coupled market-epidemiological (ME) models. We also explore the particular case of adaptive market behaviour without epidemic, and study its influence on trade dynamics. Interaction constraints are now subsumed into adaptive behaviour, and exchanges are described by a dynamical process resulting from agents' decisions to trade. It follows we can explore the economic-epidemiological implications of a particular constraint, namely the frictional constraint or trade friction (Fig. 2.1F), which creates a trade-off between the frequency and intensity of market transactions. Building upon the SIRS compartmental model (section 1.1), we start by conceiving and exploring a ME model at the macroeconomic scale, which we apply to livestock movements in France (chapter 6). The macro ME model suggests friction exerts a leading influence on joint trade and infection dynamics. To confirm our initial results, we

extend the macroeconomic framework at the microeconomic scale. We can hence account for dynamical networks of infectious contacts among agents (chapter 7). Thanks to detailed available data on the trade of cattle, we can clarify the meaning of some model parameters and state variables, and study the influence of heterogeneity in local friction and global economic spillovers on economic-epidemiological dynamics. We emphasise our two ME models, which describe market processes occurring at macroscopic and microscopic levels, make use of various ideas reported in part II, notably some socio-economic mechanisms presented in appendix A and the market categories as defined in chapter 5.

Chapitre 6

A mechanistic model to explore the interplay between exchange and epidemic dynamics in markets with homogeneous agents.

This chapter presents an article in minor revision for *Journal of Theoretical Biology* entitled *Epidemics in markets with trade friction and imperfect transactions* [Moslonka-Lefebvre *et al.*, 2013].

While it is widely accepted that trade can drive disease epidemics and other biological invasions [Noremark *et al.*, 2011], the interaction of these processes with the inherent dynamics of markets remains unclear. Markets can propagate diseases among market agents (e.g. farms) through the exchange of contaminated products (e.g. animals). Conversely, market dynamics are influenced by complex adaptive behaviour of trade agents in response to regulation and individual awareness of epidemics [Funk *et al.*, 2009]. Epidemiological models increasingly account for reductions in infectious contact, such as risk-aversion behaviour in response to pathogen outbreaks [Funk *et al.*, 2010]. However, responses in market dynamics clearly differ from simple risk-aversion, as market behaviour is driven by complex motivations (see section 2.2).

Here, we aim at modelling the overall influence of market dynamics on the dynamics of infectious-disease epidemics, and in turn, of epidemics on market dynamics. We hence develop a coupled market-epidemiological (ME) framework at the macroeconomic scale to model trade dynamics and subsequent epidemics (Fig. 6.1A). To represent this process, we link a model of an economic market system and a model of an epidemiological system. Each model dynamics can exist per se, i.e. epidemics can occur in host populations unaffected by markets, and markets often operate without disease outbreaks through trade routes. However, by building a system that links the dynamics of these subsystems we can study their interdependencies. We study the ME model using a bottom-up approach : we start by analysing

a novel market model which explicitly account for contacts among agents in the absence of epidemics (Fig. 6.1B), and then explore the integrated ME model where epidemics and trade influence each other (Fig. 6.1A-C).

Initially assuming an absence of epidemics, we develop a frictional-trade market (FTM) model for the dynamics of markets where goods of a single type are exchanged for money at a single price per good. In contrast with standard economic models [Mas-Colell *et al.*, 1995], the FTM model explicitly accounts for *transient interaction events* among economic agents, which eventually shape the epidemiological contact structure underlying the ME model. We consider a market made of N_S suppliers and N_D demanders. We assume there are $N_{S\cap D}$ wholesalers, i.e. agents that are both supplying and demanding. We assume trade occurs directionally from suppliers to demanders (as observed in chapter 4) through *transactions*. A *transaction* is an instant delivery event from a supplier to a demander. At the market-level, transactions take place at a certain rate Θ, *the transaction rate*. Transactions are extremely important because they form a dynamic contact structure among agents driving exchanges and potential disease-transmission events. During a transaction, an *average batch size* of q goods is exchanged. It follows *trade flow* Φ, which represents the number of products traded from suppliers to demanders during a time period divided by the same period of time, is given by the product $\Theta \times q$. For a given Φ, Θ and q are hence related by a trade-off. For instance, the French cattle- and swine-exchange markets roughly have the same Φ, but q is substantially larger in swine than cattle (Fig. 2.3). In the FTM model, we control the trade-off between Θ and q by a parameter called *trade friction*, denoted κ, and which is largely used in economics to model labour markets [Pissarides, 2011]. κ compounds multiple constraints, such as the search for trading partners and the logistics of delivery, and balances the frequency and intensity of market transactions : increasing friction will decrease the transaction rate but increase the average number of goods exchanged per transaction (see the frictional constraint, Fig. 2.1F). The FTM model is closed by explicitly specifying the mechanistic processes underlying Θ and q. We assume q is determined by the potentially imperfect matching of supply and demand stocks, while Θ is driven by the current price of goods and is limited by friction. We parameterise the FTM model from French exchange-network datasets involving cattle and pigs.

We find that, in the FTM model, friction can increase market equilibration time by several orders of magnitude without necessarily altering the long-term state of trade-flow. κ hence characterises a market's inherent dynamics and response to disturbance caused, for example, by disease outbreaks.

To investigate how markets and epidemics influence each other, we model the spread and control of infectious diseases in markets by incorporating the SIRS compartmental model (section 1.1) into our FTM model. We call this aggregate the *market-epidemiological* (ME) model. In order to compare and integrate our framework with the literature, we include in the model adaptive risk aversion (RA) behaviour by the market agents. The resulting ME model remains compartmental, but exhibit a beha-

viour qualitatively different from its purely epidemiological counterpart. The ME model considers two types of feedback loops between markets and epidemics (Fig. 6.1A). Following the literature, we assume epidemics disrupt markets through removal of infectious agents, their subsequent re-introduction after sanitation measures, and RA. In turn, markets influence epidemics through dynamical trade, which stems from agents' decisions to exchange. More precisely, infection is driven by exchanges from suppliers to demanders (chapters 4–5), where exchanges determine the two components of the rate of infection transmission associated with trade pathways (see subsection 1.2.1), namely the contact rate and the probability of infection per contact. From an epidemiological perspective, a transaction constitutes the elementary epidemiological contact provided trade is the only route of transmission. It follows the contact rate is given by the average transaction rate per active demander, i.e. Θ divided by the number of demanders who are either susceptible or infected. The probability of infection per contact is an increasing function of q, which represents the epidemiological intensity of a given transaction. It follows the influence of trade friction on epidemics is not evident, since friction changes the balance between Θ and q, which both contribute to the rate of transmission.

We show that, when trade is the dominant route of transmission, market friction can be a significantly stronger determinant of epidemics than risk-aversion behaviour. In particular, when trade is the only transmission route, the basic reproduction number of the ME model (R_0) is necessarily lower than :

$$R_0 \leq \frac{1}{\gamma} \frac{N_{S \cap D} \Phi^*}{\kappa N_D N_S} \ , \tag{6.1}$$

where γ is the rate of removal from the market, and Φ^* represents trade flow at equilibrium. Relation (6.1 implies the existence of critical level of friction above which epidemics do not occur. This result suggests that highly frictional markets cannot sustain epidemics. In particular, we expect trade of swine ($\kappa \approx 70$) to be less likely to sustain epidemics than trade of cattle ($\kappa \approx 3$), a prediction in agreement with other studies [Rautureau *et al.*, 2012]. In addition, (6.1) suggests the maximal delay in removal of infected agents that still allows mitigation of epidemics increases with friction. Our results are robust to model specificities and can hold in the presence of non-trade disease-transmission routes. We suggest policy for minimising contagion in markets could adjust the delay of intervention depending on market friction and, whenever possible, generate incentives for larger-volume, less-frequent transactions, increasing trade friction without necessarily affecting overall trade flow.

The macro ME model suggests friction exerts a leading influence on joint trade and infection dynamics. For simplicity, the encounter process between agents is assumed here to be governed by a homogenous mixing process. An explicit trade-agent model would account for transactions involving identified pairs of agents who jointly decide to exchange goods against money. A key open question we investigate in chapter 7 concerns the conditions under which realistic levels of friction can also mitigate epidemics propagated on such network-based markets.

Epidemics in markets with trade friction and imperfect transactions.

Mathieu Moslonka-Lefebvre[1,2,3,#], Hervé Monod[1], Christopher A. Gilligan[2], Elisabeta Vergu[1,*] and João A. N. Filipe[2,*]

[1] INRA, UR 0341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France.

[2] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, United Kingdom.

[3] AgroParisTech, F-75005 Paris, France.

[#] Corresponding author, email : mathieu@moslonkalefebvre.com

[*] These authors contributed equally to this work

Abstract

Market trade-routes can support infectious-disease transmission, impacting biological populations and even disrupting the disease-conducive trade. Epidemiological models increasingly account for reductions in infectious contact, such as risk-aversion behaviour in response to pathogen outbreaks. However, responses in market dynamics clearly differ from simple risk aversion, as are driven by other motivation and conditioned by "friction" constraints (a term we borrow from labour economics). Consequently, the propagation of epidemics in markets of, for example livestock, is frictional because there are time and cost limitations to the production and exchange of potentially-infectious goods. Here we develop a coupled economic-epidemiological model where transient and long-term market dynamics are determined by trade friction and agent adaptation, and can influence disease transmission. The market model is parameterised from French exchange-network datasets involving cattle and pigs. We show that, when trade is the dominant route of transmission, market friction can be a significantly stronger determinant of epidemics than risk-aversion behaviour. In particular, there is a critical level of friction above which epidemics do not occur, which suggests that highly frictional markets cannot sustain epidemics. In addition, the maximal delay in removal of infected agents that still allows mitigation of epidemics increases with friction. Our results are robust to model specificities and can hold in the presence of non-trade disease-transmission routes. We suggest policy for minimising contagion in markets could adjust the delay of intervention depending on market friction and, whenever possible, generate incentives for larger-volume, less-frequent transactions, increasing trade friction without necessarily affecting overall trade flow.

Keywords : behavioural response ; economic epidemiology ; epidemic threshold ; trade networks

6.1 Introduction

While it is widely accepted that trade can drive disease epidemics and other biological invasions, the interaction of these processes with the inherent dynamics of markets remains unclear. Economic markets can propagate diseases among market agents (e.g. farms) through the exchange of contaminated products (e.g. animals). Conversely, market dynamics are influenced by complex adaptive behaviour of trade agents in response to regulation and individual awareness of epidemics. Markets that contribute to infectious disease epidemics include livestock trade of cattle [Rautureau *et al.*, 2011], swine [Lentz *et al.*, 2011], and sheep [Kiss *et al.*, 2006] ; prostitution [Rocha *et al.*, 2011] ; and airline transportation [Colizza *et al.*, 2006]. Other types of epidemics occur through exchange of information on the Internet [Lloyd & May, 2001] and exchange of debt in financial markets [May & Arinaminpathy, 2010; Haldane & May, 2011].

The likely possibility that there may be an interaction or feedback loop between epidemic dynamics and host behaviour has generally not been considered in studies for identifying effective strategies for the control of infectious diseases. Recent modelling studies, however, have explored the epidemiological impact of one particular response that can cause such an interaction, namely, adaptive risk-aversion (RA) behaviour [Funk *et al.*, 2009, 2010; Durham & Casman, 2012; Nicolaides *et al.*, 2013]. RA behaviour is a form of disease prevention where asymptomatic hosts reduce exposure to infection by reducing their contact rate (e.g. by staying home) and/or their probability of infection per contact (e.g. by wearing protective masks) ; it implies that hosts have some information about a given disease outbreak and act on their own initiative rather than relying on community-wide measures by regulatory bodies. If this behaviour is determined by the ongoing perception of a variable risk, then it is said to be 'adaptive' RA. In the literature, RA has been expressed as a simple function of disease prevalence or outbreak awareness [Funk *et al.*, 2010], or evaluated via complex optimisation of a host's economic weighing between the benefits of interacting with other hosts and the cost of infection that may be acquired through such contacts [Fenichel *et al.*, 2011; Morin *et al.*, 2013]. Naturally, epidemiological models that neglect RA behaviour tend to overestimate the probability of occurrence and severity (e.g. infectious peak and cumulative cases) of epidemics [Funk *et al.*, 2009, 2010; Fenichel *et al.*, 2011; Morin *et al.*, 2013]. To the best of our knowledge, epidemiological modelling studies have focused on adaptive human behaviour that is altered solely in response to awareness of outbreaks.

In this paper, we investigate the epidemiological effects and implications for disease control of more general human adaptive behaviour, which may be difficult to anticipate. We focus on markets of goods, where the dynamics of potentially-infectious contacts are driven, primarily, by economic decisions very different from those underlying disease RA. Specifically, we aim at modelling the influence of market dynamics on the dynamics of infectious-disease epidemics, and in turn, the influence of epidemics on market dynamics. Indeed, when an epidemic shock occurs in a market, the subsequent actions and

behaviour may either help to restore or further disturb the balance between supply and demand. Sanatory regulation and RA aimed at reducing infectious contacts can diminish supply and demand. Conversely, the responses exhibited by market agents can have either positive or negative impacts on disease dynamics. For example, agents that try to establish alternative but potentially-infectious trade relationships could outweigh the effect of regulation and RA efforts, i.e. the effort of individual agents to adjust their own supply and demand to changing price could worsen disease outbreaks. Furthermore, the establishment of trade relationships, which underpin the epidemiological contacts, is conditioned by physical impediments such as the minimum time and effort involved. These resources are limited by *i*) producing profitable goods (e.g. reproduction and growth of livestock), *ii*) searching business partners and cutting deals (e.g. a buyer needs to find a seller with whom to trade a given number of goods at a given market price), and *iii*) delivering goods (e.g. organising transport from a buying to a selling holding). In labour economics, such interaction constraints shaping relationships between work sellers and work buyers are known as 'friction' (see the model of Diamond, Mortensen and Pissarides [Pissarides, 1985; Mortensen & Pissarides, 1994; Pissarides, 2011]). We transpose this concept to exchange-markets that can conduce infectious diseases. Therefore, by limiting the development of potentially-infectious trade contacts, friction may have a suppressive effect on epidemics. Phenomena such as friction and adjustment in supply and demand illustrate that human behaviour in response to disease epidemics that are are supported by trade does not simplify to regulation and RA.

To the best of our knowledge, existing mathematical models of market dynamics do not seem to represent explicitly the variety of transient non-equilibrium dynamics that occur when a market is disturbed and until it eventually reaches a steady-state equilibrium (see section E.1) ; therefore, they may not fully incorporate processes and parameters that establish a time scale for market steady-state equilibration, which is expected to vary widely among markets for rapidly changing external conditions such as those induced by epidemic outbreaks. In order to represent the joint-dynamics of markets and epidemics at appropriate and mutually-consistent time scales (see Section 6.2.1), we have developed a novel economic-market model, the *frictional-trade market (FTM) model* (see Section 6.2.2), where transient and long-term dynamics are determined by the level of trade friction and agents' decisions to supply or demand goods. Subsequently, we integrate market and epidemiological processes in a *market-epidemiological (ME)* modelling framework where trade influences disease transmission and disease control actions affect trade (see Section 6.2.3).

We first study the behaviour of the FTM model in the absence of epidemics (see Section 6.3.1). Then, we investigate how market dynamics affect epidemic development, and, conversely, how epidemics disrupt short and long-term market dynamics (see Section 6.3.2). We consider two forms of response to disease outbreaks taken from the literature : the removal (inactivation) of market agents found to be infected by regulators and their later re-introduction or replacement, and an adaptive RA behaviour of market agents. Therefore, we highlight differences in concept and impact on epidemic de-

velopment between market dynamics and RA ; market dynamics are influenced by centrally-regulated actions and by collective behaviour that drives changes to supply and demand in response to changing conditions, while RA behaviour is determined solely by individual decision-making. Finally, we extend our study beyond an isolated (e.g. national) market, by contrasting scenarios where infectious diseases are propagated through trade pathways with differing degree of openness to international trade and non-trade disease-transmission pathways. We expect our central results to apply to a range of different types of markets, and illustrate applications to cattle and swine livestock markets in France.

6.2 Market-epidemiological modelling framework

6.2.1 Overview

We develop a novel theoretical framework for the propagation of infectious diseases in economic markets where the exchanged goods can transmit an infectious organism between market agents (Fig. 6.1A). In order to represent this process, we link a model of an economic market system and a model of an epidemiological system. Each model dynamics can exist per se, i.e. epidemics can occur in host populations unaffected by markets, and markets often operate without disease outbreaks through trade routes. However, by building a system that links the dynamics of these subsystems we can study their interdependencies. As the epidemiological model we use is a simple adaptation of a standard compartmental epidemiological model, it is introduced later with brief explanation. The dynamic economic-market model, however, is novel, and is derived in detail. A key property of this model is its coefficient of friction, which characterises a market's inherent dynamics and response to disturbance caused, for example, by disease outbreaks (see the 'Results' section).

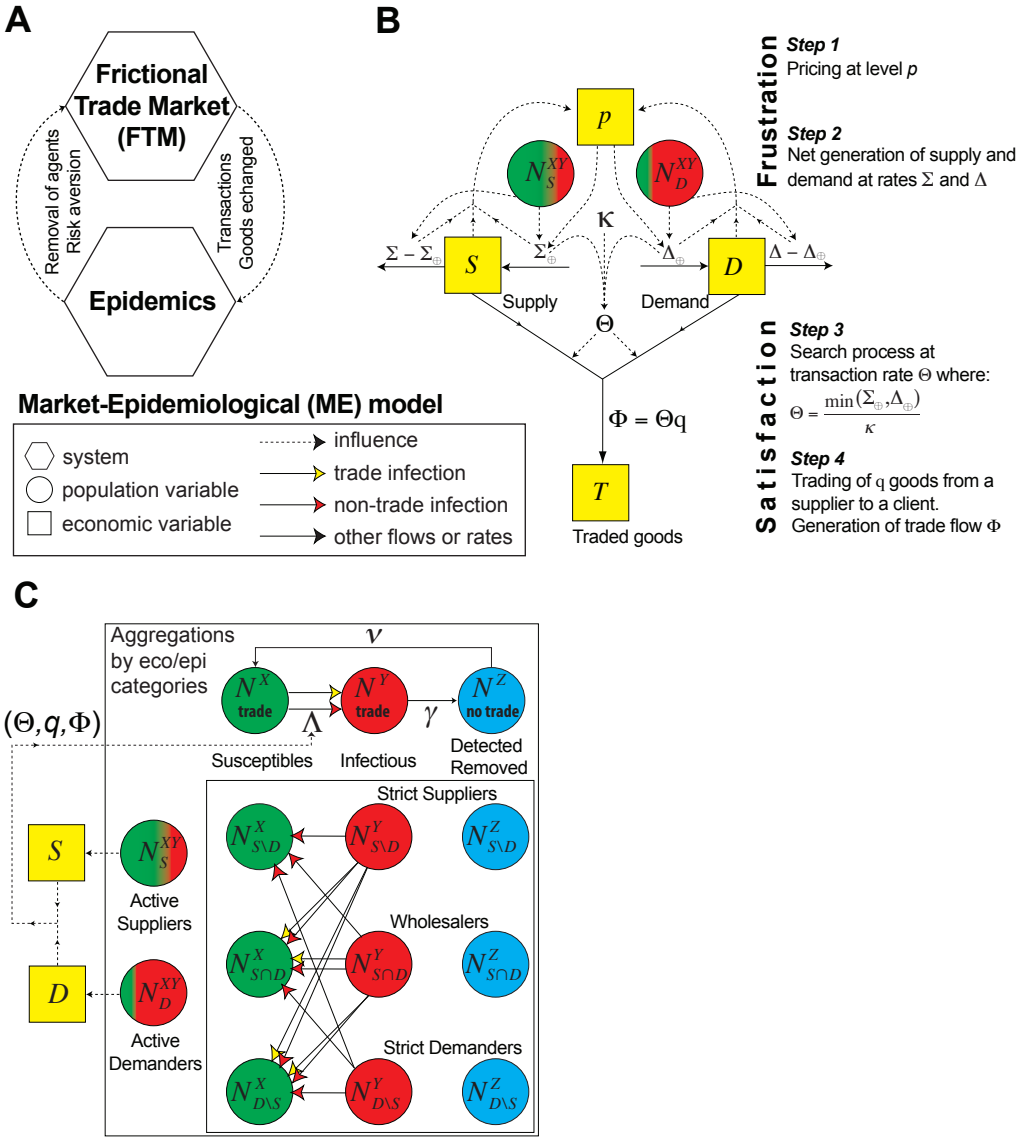


FIGURE 6.1 – **Joint market-epidemiological modelling framework. A**) General structure and key links between the market and epidemiological subsystems. **B**) Components of the frictional-trade market (FTM) model. **C**) Components of the market-epidemiological (ME) model. Yellow (red) arrows represent direction of disease transmission through trade (non-trade) routes. The market agents host the disease-causing pathogen. N_A represents the number of agents of arbitrary type A that can vary over time. $N_{S\setminus D}$, $N_{S\cap D}$ and $N_{D\setminus S}$ denote the numbers of *strict suppliers*, *wholesalers* (supplier and demander) and *strict demanders* respectively. From an epidemiological point-of-view, each agent can be in the *Susceptible* (X), *Infectious* (Y), or *Removed* (Z) state. Susceptible agents become infectious at rate $\Lambda(t)$; infectious agents are removed at rate γ ; and removed agents re-enter the market (recover) at rate ν. Here, 'removal' means that an infectious agent is detected and removed from the market by a regulator and becomes *inactive*. Each agent, whether a strict supplier, wholesaler, or strict demander, can be in each of the epidemiological states. For example, N_S^X and N_S^Y denote the number of susceptible and infectious suppliers, respectively. Therefore, there are $N_S^{XY} = N_{S\cap D}^{XY} + N_{S\setminus D}^{XY}$ active suppliers. Accumulated supply and demanded stocks S and D quantify the *willingness to trade* or *frustration* of suppliers and demanders respectively. Supply stock consists of goods available for sale, while demanded stock represents potential goods that buyers want to purchase and are virtual. The *trade flow* Φ out of the supply and demanded stocks, aggregates all exchanges between supply and demand agents that agree to trade some of their stock at price p. Transactions occur at *transaction rate* Θ and in each transaction an *average transaction stock* q is exchanged. At *the reference market* steady-state equilibrium, q is equal to κ, *the coefficient of friction*. If trade flow were kept at the corresponding equilibrium level Φ^*, then the transaction of goods would occur at a lesser frequency $\Theta^* = \Phi^*/\kappa$ if the number of goods exchanged per transaction $q = \kappa$ would increase (see 'Results').

6.2.2 A frictional-trade market (FTM) model

Here we develop a FTM model for the dynamics of markets (without epidemics) where goods of a single type are exchanged for money in *transactions* between suppliers (sellers) and demanders (buyers), and all transactions at given time are based on a single price per good. The model tracks the dynamics of *extensive* state variables (i.e. with a 'size' or 'scale') such as stocks, number of agents, and trade flow, and *intensive* state variables (i.e. global indices) such as price of goods. Usually, price, trade flow, and number of agents are observable market quantities, while overall stocks are not observable.

Our model market is defined, at each time t, by an overall *price* per good $p(t)$, and overall accumulated *supply stock* $S(t)$ and *demanded stock* $D(t)$ (see Table 6.1 for an explanation of relationship with common economic terminology and concepts). Accumulated supply stock consists of goods available for sale but not yet sold, while demanded stock represents potential goods that buyers want to purchase but have not yet purchased. These accumulated stocks quantify the *willingness to trade* of suppliers and demanders. These accumulated stocks quantify the suppliers' and demanders' *willingness to trade*. The supply stock is part of the *inventory stock* (e.g. total animals within farms), which also includes goods not for sale owned by sellers. The demanded stock cannot exceed the *physical capacity* of demanders (e.g. maximum number of animals), which is similar to a carrying capacity in ecology. Our FTM model integrates many complex mechanisms and dynamical feedbacks that have seemingly not been explored concurrently in the literature (see sectionE.1). To keep our model parsimonious, we hence model market dynamics at a whole-market level, where agent-level stocks and transactions are approached by average values per agent. Inspired by well-known population dynamics models (e.g. [May, 1977; Durrett & Levin, 1994]), our whole-market-scale model is a population-level description of agents and stocks with mass-action interactions (transactions). At this simplified level of description, we define the market model through temporal change in overall stocks. Each *stock* (S and D) is created at a specific *net generation rate* and depleted through a *trade flow*, represented by the rate equation :

$$\frac{d\,[\text{stock}]}{dt} = [\text{net generation rate}] - [\text{trade flow}]\ . \tag{6.2}$$

The *net generation rate*, of supply stock Σ or demanded stock Δ, is composed of : 1) a *generation rate*, $\Sigma_\oplus(p, N_S)$ or $\Delta_\oplus(p, N_D)$, respectively, that depends on current price p and numbers of supply agents N_S and demand agents N_D, 2) a net *loss rate* L, e.g. spoilage of supply goods and loss of demanders interest (positive loss) or multiplication of goods such as reproduction of livestock (negative loss), and 3) an *external flow* of stock E, e.g., import or export of raw materials or goods :

$$\begin{aligned}
[\text{net generation rate}] &= [\text{generation rate}] - [\text{loss rate}] + [\text{external flow}]\ , \\
\Sigma &= \Sigma_\oplus(p, N_S) - L_S + E_S \quad (\text{net supply flow})\ , \\
\Delta &= \Delta_\oplus(p, N_D) - L_D + E_D \quad (\text{net demanded flow})\ .
\end{aligned} \tag{6.3}$$

Following the economic literature (see section E.1 for details), generation rates are defined as :

$$\begin{aligned} \Sigma_{\oplus}(p, N_S) &= N_S \ \sigma_0 \ p^{\varepsilon_S} \ , \\ \Delta_{\oplus}(p, N_D) &= N_D \ \delta_0 \ p^{-\varepsilon_D} \ , \end{aligned} \tag{6.4}$$

where σ_0 and δ_0 are the reference *per-agent* generation rates in supply and demanded stocks at the reference price $p = p_0 = 1$, and $\varepsilon_S \geq 0$ and $\varepsilon_D \geq 0$ are the price elasticities of supply and demand respectively. ε_S and ε_D are assumed constant and are defined, respectively, as the relative changes in the generation rates of supply and demanded stocks in response to the relative change in price. Notice that $\Sigma_{\oplus}(p, N_S)$ increases while $\Delta_{\oplus}(p, N_D)$ decreases with increasing price. Furthermore, we assume that the loss rates are directly proportional to stocks, i.e. $L_S = r_S S(t)$ and $L_D = r_D D(t)$ with r_S, r_D constants, and that external flows (E_S, E_D) are constant.

The *trade flow* $\Phi(t)$ out of the supply and demanded stocks, aggregates all exchanges between supply and demand agents that agree to trade some of their stock (i.e. exchange of supply stock for demander's money). Transactions occur at *transaction rate* $\Theta(t)$ and in each transaction an *average transaction stock* $q(t)$ is exchanged ; therefore, the trade flow is :

$$\begin{aligned} [\text{trade flow}] &= [\text{transaction rate}] \ [\text{average transaction stock}] \ , \\ \Phi(t) &= \Theta(t) \ q(t) \ . \end{aligned} \tag{6.5}$$

First we define q. The average per-agent supply and demanded stocks are $S(t)/N_S$ and $D(t)/N_D$. Here we take the number of market agents to be constant in time, but later, when considering epidemics in markets, we allow for removal of infectious agents and their subsequent re-introduction after sanitation measures. In this model we assume that, once a pair of supply and demand agents has been identified and agreed to transact, they exchange the maximum possible per-agent stock (a best-possible match) : $S(t)/N_S$ if there is excess accumulated demand ($D(t)/N_D \geq S(t)/N_S$), and $D(t)/N_D$ if there is excess accumulated supply ($S(t)/N_S \geq D(t)/N_D$). Hence, the average transaction stock, conditional on best-possible matching, is :

$$q(t) = \min\{S(t)/N_S; D(t)/N_D\} \ . \tag{6.6}$$

The transaction rate $\Theta(t)$ in (6.5) is determined by a driving factor, the *reference transaction rate*, and a *limiting factor* that compounds multiple constraints, such as the time-consuming *production* of goods, the *search* for trading partners, and the logistics of stock *delivery*. Hence :

$$\begin{aligned} [\text{transaction rate}] &= \frac{[\text{reference transaction rate}]}{[\text{limiting factor}]} \ , \\ \Theta(t) &= \frac{\min\{\Sigma_{\oplus}(p, N_S); \Delta_{\oplus}(p, N_D)\}}{\kappa} \ . \end{aligned} \tag{6.7}$$

We assume that the *reference transaction rate* is determined predominantly by present decisions to increase stocks and, thus, by the generation rates in (6.3)-(6.4) ; more specifically, it is determined by the current maximum possible rate of exchange of indivisible goods between the two sides of the market, $\min\{\Sigma_\oplus(p,N_S),\Delta_\oplus(p,N_D)\}$. Here, the reference transaction rate is not determined by the net generation rates in (6.3) because these include loss rates L and external flows E that are likely to depend on the overall stocks S and D, which usually are imperfectly know or unquantifiable ; however, the current price $p(t)$, which determines the production rates, is known to market agents. In addition, we represent the limiting factor of the transaction rate by a dimensionless coefficient κ that, through physical analogy, we call coefficient of *friction* or *inverse fluidity* of the market (see more below).

To finalise the specification of our market model, we need to specify price dynamics. The *net willingness to trade* or *frustration* of agents at a given time is the excess in accumulated demanded stock, $D(t)-S(t)$. While different relationships between price and other state variables can be specified, we assume for definiteness, and in agreement with literature (see Section E.1), that changes in log price are directly related, via a dimensionless coefficient μ, to changes in net willingness to trade,

$$\begin{aligned}\frac{d[\text{price}]}{dt} &= \mu\frac{d[\text{net willingness to trade}]}{dt}[\text{price}]\ ,\\ \frac{dp}{dt} &= \mu\frac{d(D-S)}{dt}\ p\ ,\end{aligned} \tag{6.8}$$

which can be solved explicitly as an exponential relationship (see (6.9)).

With these specific assumptions, and in the absence of disturbances such as epidemics, our FTM dynamics are defined by the equations (represented diagrammatically in Fig. 6.1B) :

$$\begin{aligned}\frac{dS}{dt} &= \overbrace{\underbrace{N_S\sigma_0 p^{\varepsilon_S}}_{\Sigma_\oplus} - r_S S + E_S}^{\Sigma} - \overbrace{\underbrace{\frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}}_{\Theta}\underbrace{\min\{\tfrac{S}{N_S};\tfrac{D}{N_D}\}}_{q}}^{\Phi}\ ,\\ \frac{dD}{dt} &= \overbrace{\underbrace{N_D\delta_0 p^{-\varepsilon_D}}_{\Delta_\oplus} - r_D D + E_D}^{\Delta} - \overbrace{\Theta q}^{\Phi}\ ,\\ p(t) &= p(0)\exp\{\mu\,[D(t)-S(t)-(D(0)-S(0))]\}\ .\end{aligned} \tag{6.9}$$

The special case where $r_S = r_D = 0$ and $E_S = E_D = 0$ is referred to as the *reference market*. Hereafter, we use a star in superscript to denote market variables at steady-state equilibrium in the reference market in the absence of disturbances (e.g. Φ^* is the reference trade flow at steady-state equilibrium).

At a market (macroscopic) level, our model has two determinants of trade flow (see (6.5)-(6.7)) : the coefficient of friction κ, and the average transaction stock q.

We explain first the interpretation and significance of friction. An increase in the coefficient of friction κ reduces the transaction rate (see (6.7))as in physical systems, friction is a macroscopic

manifestation of microscropic resistance to movement. The microscopic-level constraints that underlie friction in trade (transactions) are a result of 'viscosity' in *production* of tradable good, *search* of partners, and *delivery* of stock ; therefore, we may consider that κ has at least three components :

$$\kappa = \kappa_{\text{production}} + \kappa_{\text{search}} + \kappa_{\text{delivery}} \ . \tag{6.10}$$

As an example, let us consider livestock markets and suppose κ_{delivery} is the dominant component and 1000 animals are produced, consumed and traded per farmer per year. If a minimum of $\kappa = 10$ animals are delivered in a single shipment, the transaction rate is at most $\Theta = 1000/10 = 100$ per farmer per year. If, however, the nature of the animals (for instance pigs are smaller than cattle) and transportation mean it is more viable to ship a minimum of 100 animals, then the transaction rate would be at most $\Theta = 1000/100 = 10$ per farmer per year. This illustration involves simplifications ; in practice the macroscopic coefficient κ is unlikely to associate so directly with a microscopic quantity like minimum shipment size. In our reference market, κ turns out to be the average stock exchanged per transaction when the market is at steady-state equilibrium q^* (Results section *Trade without stock loss : the reference market*) ; this suggests κ must exceed, but can still be arbitrarily larger than the minimum shipment size. In markets of money and financial products, for example, goods can be subdivided almost indefinitely, so we expect κ to be potentially close to 0, which translates into an almost frictionless market, in line with the high liquidity of monetary and financial markets. In contrast, for indivisible goods such as livestock, the minimum shipment size is at least one, so we expect $\kappa \geq 1$. Due to a combination of production constraints and costly transportation of single animals, farmers often exchange batches of animals rather than individual animals, which further implies that $\kappa > 1$. For simplicity, κ is estimated assuming the French cattle and swine markets are in steady-state equilibrium in years 2009 and 2010 respectively. Temporal analyses at both monthly and yearly scales indeed suggest that these markets seem to be in steady-state equilibrium in years 2005-2009 for cattle and year 2010 for swine (unpublished data). While we estimate κ under the assumption of market steady-state equilibrium, we expect friction to exist both at equilibrium and outside of equilibrium (see Table 6.1 for explanation of non-equilibria accounted for by our FTM model). We also expect the coefficient of friction to be independent of whether the system is or not at steady-state equilibrium at least in the model as formulated ; although we expect the effect of friction to be more apparent outside of equilibrium. Moreover, we expect estimation of friction to rely only weakly on the accuracy of the equilibrium assumption because the model is only weakly sensitive to linear variation in κ (Fig 6.2A). We find $\kappa = q^* \approx 3$ for cattle and $\kappa = q^* \approx 72$ for pigs, where q^* is mean number of animals exchanged per transaction at steady-state equilibrium (calculated at the reference market steady-state equilibrium, see Material and Methods and Section E.2). In fact, in line with this interpretation of the coefficient of friction, we find from direct inspection of the datasets that the empirical average $\langle q \rangle$, calculated based

on pairs of trading agents, is $\langle q \rangle \approx 3$ (SD ≈ 6) for cattle in 2009 and $\langle q \rangle \approx 63$ (SD ≈ 102) for swine in 2010, where SD is the empirical standard deviation.

The second determinant of trade flow is the average transaction stock q (see (6.5) and (6.6)). A characteristic of this market model is that the match between supply and demand is generally imperfect, in the sense that there can be *residual* supply or demanded stock after each transaction (imposed by the min function in (6.6) when there is excess accumulated stock per agent). These residuals lead to a transient 'excess frustration' in market agents whose duration depends on the 'fluidity' (or conversely, the 'friction') of the market. As trade flow depletes both stocks S and D and is an observable quantity, the cumulative trade flow over a period, $T([t_0, t_f]) = \int_{t_0}^{t_f} \Phi(t)dt$, is a measurable indicator of the evolution of satisfaction (or frustration) of market agents. Detailed analyses of our market model (Section E.3.5 and Fig. E.3) show that supply stock and demanded stock in the model's steady-state equilibrium are typically lower than the inventory stock of cattle in holdings recorded in France. These results support our definitions of supply and demanded stocks and suggest that the inventory stock of a farm is close to its maximal capacity. Our explicit representation of market transactions driven by imperfect and frictional individual-level supply and demand, from which potentially long-lasting out-of-steady-state market dynamics can emerge depending on the coefficient of friction (Results section *Market dynamics without shocks - effect of trade friction*), seems fundamentally different from some economic models (see Table 6.1 for explanations of potential differences relative to terminology or concepts in economics literature, Section E.1 for a comparative review of existing market models, and Section E.3.1 for further detail on the microeconomic foundations of our FTM model).

TABLE 6.1 – **Relationship between terminology in the frictional-trade market (FTM) model and in economics**

Terminology adopted (notation/equation)	**Meaning**	**Related economic concept**	**References**
(Accumulated) supply stock (S) ((6.2) and (6.9))	amount of goods available for sale (e.g. in a farm, may be a fraction of the total number of animals)	a part of the inventory, capital, or stock of natural resources (e.g. in a farm, may be the total number of animals)	[Clark, 1976; Barro & Sala-i Martin, 2003; Khan & Thomas, 2003; Caputo, 2005]
Supply flow, or rate of generation of supply stock ($\Sigma_\oplus$, (6.4))	increase in amount of goods available for sale per unit time	(in excess or equal to) the amount of goods sold per unit time, known as 'supply'	[Fair & Jaffee, 1972; Hahn, 1982; Mas-Colell *et al.*, 1995; McCauley, 2009]
(Accumulated) demanded stock (D, (6.2) and (6.9))	number of goods wanted for purchase but not yet purchased, including accumulated demand that has so far not been satisfied (e.g. farmers who have been losing livestock during an undetected disease outbreak have accumulated the will to purchase D replacement animals ; depending on current conditions of supply of stock, there may be multiple subsequent transactions that gradually satisfy this demand, but which can spread over an economically-significant time period	·	·
Demand flow, or rate of generation of demand for purchase of stock ($\Delta_\oplus$, (6.4))	increase in potential amount of goods to be purchased per unit time	in excess or equal to the amount of goods purchased per unit time (rate of depletion of stocks not including losses), known as 'demand', where 'demand' is often assumed to equal 'supply'	[Fair & Jaffee, 1972; Hahn, 1982; Mas-Colell *et al.*, 1995; McCauley, 2009]
Reference transaction rate ($\min\{\Sigma_\oplus;\Delta_\oplus\}$, (6.7))	reference number of transactions per unit time when as few as one good can be exchanged per transaction	·	·
Trade friction ($\kappa \geq 0$, (6.7) and (6.10))	physical impediments to the trade of goods : constraints in production, effort and delay in the search for business partners and negotiation of deals, and logistical constraints in the delivery of goods. In steady-state equilibrium κ is the average number of goods exchanged per transaction, implying that $\kappa \geq 1$ for most systems where it is not possible to trade less than one unit of goods (6.7)	transaction costs ; friction in labour markets, e.g. intensity of search and matching	[Klaes, 2008; Pissarides, 1985; Mortensen & Pissarides, 1994; Williamson & Masten, 1999; Klos & Nooteboom, 2001; Tesfatsion, 2001; Pissarides, 2011]
Transaction rate ($\Theta = \frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}$, (6.7))	number of stock exchanges in the market (deliveries from sellers to buyers) per unit time	'matching function' in labour market economics : number of successful matches per unit time between job vacancies and unemployed workers	[Pissarides, 1985; Mortensen & Pissarides, 1994; Pissarides, 2011]
Average transaction stock (q, (6.6))	average number of goods exchanged per transaction which we currently set to the maximal match between current supply stock and demanded stock in each transaction	·	·
Trade flow ($\Phi = \Theta\ q$, (6.5))	market-level amount of goods actually exchanged per unit time	related to actual quantity exchanged at a given point in time, but not identical when the period during which exchanges occurred was not specified (see also Section E.3.3)	[Fair & Jaffee, 1972]
Occurrences of 'non-equilibrium' in the FTM model	*Overall rates of change are not zero*, i.e. non-negligible rate of change of market state variables : supply stock ($\frac{dS}{dt}$), demanded stock ($\frac{dD}{dt}$), or price ($\frac{dp}{dt}$) in (6.9)	absence of steady-state or inter-temporal equilibrium	[Clark, 1976; Pindyck, 1982; Barro & Sala-i Martin, 2003; Khan & Thomas, 2003; Caputo, 2005; Nowak, 2006; McCauley, 2009]
	the two possibilities below may associate with the occurrence of a non-steady-state :		
	Imperfect transactions, i.e. the supply and demanded stocks of trading partners do not necessarily match, causing residual stocks that subsequently accumulate (q (6.6) is the smallest of $\frac{S}{N_S}$ and $\frac{D}{N_D}$)	disequilibrium	[Fair & Jaffee, 1972; Quandt, 1988]
	Imbalance between net generation rates of supply and demanded stocks, i.e. net rates of generation of supply stock and demanded stock differ ($\Sigma \neq \Delta$, (6.3)). The level of imbalance depends on how much the market is competitive, as determined by parameter μ in our model which sets a timescale for price dynamics ((6.8) and (6.9)), and may or not preclude an eventual steady state	Imperfect competition, potentially leading to unbalanced long-run profits between selling and buying firms. Potential causes of such market failures include : epidemic outbreak, monopolistic behaviour, externalities, free-riding, information asymmetries, transaction costs, and price rigidities	[Mas-Colell *et al.*, 1995; Williamson & Masten, 1999; Klos & Nooteboom, 2001; Klein *et al.*, 2007; Klaes, 2008; Silvestre, 2008]

We focus on notions that may help readers with economics background. A dot (·) means the corresponding concept does not appear to be defined in economics (to the best of our knowledge).

6.2.3 Market-epidemiological (ME) model with risk aversion

To investigate how disease epidemics and economic markets can influence each other, we model the spread and control of infectious diseases in markets by incorporating a standard epidemiological (E) model into our FTM model. We call this aggregate *market-epidemiological* (ME) model. In order to compare and integrate our framework with the literature, we include in the model adaptive risk aversion (RA) behaviour by the market agents.

The market agents are the population hosting the disease-causing pathogen. We use notation $\mathcal{N}_A$ to represent a set of agents of arbitrary type A, and N_A to represent their number, which can vary over time. An agent can be a *strict supplier*, a *wholesaler* (supplier and demander), or a *strict demander* [Pautasso *et al.*, 2010]. The corresponding sets of agents are $\mathcal{N}_{S\setminus D}$, $\mathcal{N}_{S\cap D}$, and $\mathcal{N}_{D\setminus S}$; and, the total number of agents is $N = N_{S\setminus D} + N_{S\cap D} + N_{D\setminus S}$. The markets is hence composed of $N_S = N_{S\setminus D} + N_{S\cap D}$ suppliers and $N_D = N_{S\cap D} + N_{D\setminus S}$ demanders. We use a standard 'SIRS' epidemiological model [Anderson & May, 1991] where each agent (host) can be in the *Susceptible* (X), *Infectious* (Y), or Removed (Z) state (the notation XYZ is preferred to SIR to avoid confusion with the market model notation). Susceptible agents become infectious at rate $\Lambda(t)$ (the force of infection); infectious agents are removed at rate γ; and removed agents re-enter the market (recover) at rate ν (Fig. 6.1C). Here, 'removal' means that an infectious agent is detected and removed from the market by a regulator and becomes *inactive*. The infectious period ($1/\gamma$) is, therefore, the average time during which an infectious agent remains *active*, which is determined by the swiftness of the regulators (we assume removal occurs immediately after detection and before the end of the biological infectious period). The recovery period ($1/\nu$) is the quarantine and sanitation time during which an infectious agent remains *inactive*, and is generally determined by regulators and agents. Each agent, whether a strict supplier, wholesaler, or strict demander, can be in each of the epidemiological states (Fig. 6.1C). For example, $N^X_{S\cap D}$ and $N^Y_{S\cap D}$ denote the number of susceptible and infectious wholesalers, respectively. Therefore, there are $N^{XY}_{S\cap D} = N^X_{S\cap D} + N^Y_{S\cap D}$ *active* wholesalers and $N^{XY}_S = N^{XY}_{S\cap D} + N^{XY}_{S\setminus D}$ active suppliers.

In specifying how disease spreads in the market model, we consider the general case where the pathogen can be transmitted both through trade routes (tr) and non-trade routes ($\bar{tr}$). We assume that transmission through trade occurs in the direction of transactions, i.e. through the shipment of contaminated stock from active infected suppliers ($\mathcal{N}^Y_S$) to active non-infected demanders ($\mathcal{N}^X_D$), while transmission through non-trade routes occurs from active infected agents ($\mathcal{N}^Y$) to active non-infected agents ($\mathcal{N}^X$). We also allow for import of contaminated stock though external flow E (see (6.3)). In this case, the force of infection on demanders has terms associated with transmission through trade and non-trade routes, and a risk-aversion factor :

$$\Lambda(t) = \left[\Lambda_{tr}(t) + \Lambda_{\overline{tr}}(t)\right] P_{RA}(t) \ . \tag{6.11}$$

The term for trade routes is $\Lambda_{tr}(t) = \beta_{tr} \cdot N_S^Y / N_S^{XY}$, i.e. a rate of transmission via trade β_{tr}, times the probability of transacting with an active supplier that is infected, N_S^Y / N_S^{XY}. The rate $\beta_{tr} = P_{tr}(q) \cdot \Theta(t)/N_D^{XY}$, i.e. the probability $P_{tr}(q)$ of acquiring infection from the average transaction stock $q(t)$ shipped from an infected supplier during a single transaction, times the transaction rate per active demander $\Theta(t)/N_D^{XY}$. If each of the q units of stock has similar and independent probability of infection ϕ, the probability of no infection from the stock is $(1-\phi)^q$, and the probability that the demander is infected during a single transaction by at least one unit of stock is $P_{tr}(q) = 1-(1-\phi)^q$. Similarly, the force of infection via non-trade routes, is $\Lambda_{\overline{tr}}(t) = \beta_{\overline{tr}} \cdot N^Y / N^{XY}$, the rate of transmission $\beta_{\overline{tr}}$ per active agents via yet unspecified transmission routes, times the probability of contacting an infected active agent, N^Y / N^{XY}. Finally :

$$\begin{aligned} \Lambda_{tr} &= \overbrace{\underbrace{[1-(1-\phi)^q]}_{P_{tr}(q)} \frac{\Theta}{N_D^{XY}}}^{\beta_{tr}} \frac{N_S^Y}{N_S^{XY}} \ , \\ \Lambda_{\overline{tr}} &= \beta_{\overline{tr}} \ \frac{N^Y}{N^{XY}} \ . \end{aligned} \tag{6.12}$$

Following the literature [Funk *et al.*, 2010], we include RA in our model by allowing agents to reduce their probability of infection per transaction, or per non-trade contact, according to the level of disease detection. We assume that RA reduces the probability of infection per contact by a factor $0 \leq P_{RA} \leq 1$ given by :

$$P_{RA} = \left(1 - \frac{N^Z}{N}\right)^{\alpha} , \tag{6.13}$$

where $\alpha \geq 0$, and N^Z is the number of inactive (detected) agents. When $\alpha = 0$ there is no RA ($P_{RA} = 1$), while in the limit $\alpha \to \infty$ RA is maximal ($P_{RA} = 0$).

In addition to how epidemics affect the active agent population, we consider how epidemics affect their stocks. When an infected supplier is removed (at rate γ) its share of the stock, S/N_S^{XY} is also removed ; hence the rate of removal of infected supplier stock is $\rho_S = \gamma N_S^Y (S/N_S^{XY})$. As the generic stock-loss rates in the FTM model without epidemics (equations (6.3) and (6.9)) are analogous to ρ_S, in this paper we neglect losses other than ρ_S by setting $r_S = r_D = 0$.

Finally, the ME model is defined by the dynamics of market stocks (which generalizes (6.9)) and agents under an epidemic (we show wholesaler equations as regard to epidemics, see Section E.4 for other agents) :

$$
\begin{aligned}
\frac{dS}{dt} &= \overbrace{\underbrace{N_S^{XY}\sigma_0 p^{\varepsilon_S}}_{\Sigma_\oplus} - \gamma\frac{N_S^Y}{N_S^{XY}}S + E_S}^{\Sigma} - \overbrace{\underbrace{\frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}}_{\Theta}\underbrace{\min\{\tfrac{S}{N_S^{XY}};\tfrac{D}{N_D^{XY}}\}}_{q}}^{\Phi} \ , \\
\frac{dD}{dt} &= \overbrace{\underbrace{N_D^{XY}\delta_0 p^{-\varepsilon_D}}_{\Delta_\oplus} - \gamma\frac{N_D^Y}{N_D^{XY}}D + E_D}^{\Delta} - \overbrace{\Theta q}^{\Phi} \ , \\
\frac{dN_{S\cap D}^X}{dt} &= \nu N_{S\cap D}^Z - \Lambda N_{S\cap D}^X \ , \\
\frac{dN_{S\cap D}^Y}{dt} &= \Lambda N_{S\cap D}^X - \gamma N_{S\cap D}^Y \ , \\
\frac{dN_{S\cap D}^Z}{dt} &= \gamma N_{S\cap D}^Y - \nu N_{S\cap D}^Z \ ,
\end{aligned}
\tag{6.14}
$$

where the market price $p(t)$ is given by (6.8) and the forces of infection by (6.11)-(F.20).

We will compare epidemics in the ME model with those in simpler models. In particular, we will consider the disease-free market steady-state equilibrium ($q^* = \kappa$ and $\Theta^* = \Phi^*/\kappa$)) as a system without explicit market dynamics, whose trade transmission rate (in (6.12)) is :

$$
\beta_{tr}^* = P_{tr}(q^*)\frac{\Theta^*}{N_D^{XY}} = [1-(1-\phi)^\kappa]\frac{\Phi^*}{\kappa N_D^{XY}} \ . \tag{6.15}
$$

In addition, when comparing with the literature on epidemiological models with RA, we will take the limit of (6.15) in a frictionless (immediate steady-state equilibration) market ($\kappa \to 0$)

$$
\beta_{tr}^0 = \ln\left(\frac{1}{1-\phi}\right)\frac{\Phi^*}{N_D^{XY}} \ , \tag{6.16}
$$

which is identical to known functional forms of the transmission rate (see e.g. [Keeling & Rohani, 2008]). As in previous epidemiological models incorporating host adaptive behaviour driven by health economics and other factors [Funk *et al.*, 2009; Fenichel *et al.*, 2011; Morin *et al.*, 2013], we allow for RA-driven reduction in transmission rate ((6.16) or, more generally, that in (6.12)) through the RA factor (F.20), as in (6.11). However, by incorporating frictional-market dynamics, our model differs from those in the epidemiological literature, which, to our knowledge are comparable to frictionless markets.

6.3 Insights on market dynamics out-of-steady-state equilibrium and their interaction with epidemics

To help understanding the implications of our new theoretical framework, we study the FTM and ME models using a bottom-up approach (Table 6.2). We first analyse our FTM model in the absence of epidemics (Fig. 6.1B), and then explore the integrated ME model where epidemics and trade influence each other (Fig. 6.1A-C). The FTM and ME models are described mathematically by systems of non-linear ordinary differential equations ((6.9) and (6.14) respectively). We derive key analytical insights on the FTM and the ME models (see Sections E.3 and E.4 respectively) using standard tools in the study of dynamical systems : existence and uniqueness theorem ; steady-state equilibria, stability and bifurcation analyses.

TABLE 6.2 – **Overview of the models investigated in our study.**

	frictional trade	price dynamics	stock loss	external flows	epidemics	risk aversion
Market dynamics without shocks (FTM)						
Reference market	✓	✓	·	·	·	·
Trade with stock loss (special case)	✓	·	✓	·	·	·
Trade with stock loss (general case)	✓	✓	✓	·	·	·
Trade with external flows	✓	✓	·	✓	·	·
Numerical illustrations	✓	✓	✓	·	·	·
Market dynamics with epidemics						
Frictionless epidemiological model	·	·	·	·	✓	✓
Market-epidemiological (ME) model	✓	✓	✓	✓	✓	✓

Ticks (dots respectively) represent the mechanisms included (not included respectively) in each model.

Our FTM and ME models incorporate a variety of parameters, the values of which may not be well known. To assess the robustness of our findings to uncertainty or variation in these parameters, as well as uncertainty about initial conditions, we carried out two global sensitivity analyses (GSA) on key economic and/or epidemiological outputs. Specifically, we did one GSA for the FTM model, and another one for the ME model. We ranked the relative importance of parameters and initial conditions of interest using an improved version of the Morris method, a suitable GSA technique for high-dimension models [Campolongo *et al.*, 2007]. In a nutshell, the improved Morris method can discriminate the sign and overall influence of factors at a low computational cost and minor risk of error (see Section E.3).

To parameterise the FTM model and estimate the range of values of the coefficient of friction that encompasses real markets (Table 6.3), we analysed trade records in two livestock markets : cattle (BDNI dataset) and swine (BDPorc dataset). The BDNI and BDPorc datasets are respectively managed by the French ministry in charge of Agriculture and the French professional union BDPorc. Each dataset details the movement of livestock within France among all economic agents involved in the supply chain, from strictly breeding farms to slaughterhouses, with various categories of wholesalers in between (e.g. breeding-fattening farms, strictly fattening farms, markets, dealers). To estimate model parameters, we extracted and reconstructed a table of individual transactions from a subset of each dataset. This table details, for each transaction : the supplier-demander pair, the date and associated volume of

goods exchanged (see Section E.2 for details on the specific subset of the dataset that we have used). Traceability is imposed by the regulator at the scale of individual animals for cattle and batches of animals for swine. As a result, transactions could be directly extracted from the BDNI dataset but not from the BDPorc dataset. We explain in Section E.2 how detailed transactions can be inferred from the BDPorc dataset.

TABLE 6.3 – **Parameter values calculated for French cattle and swine markets.**

notation	meaning	cattle	swine	unit
κ	coefficient of friction	3.4	71.7	none
Φ^*	trade flow	7,578,476	8,075,973	animals per year
Θ^*	transaction rate	2,224,182	112,683	deliveries per year
σ_0	per-agent generation rate of supply stock	39	1,474	animals per year
δ_0	per-agent generation rate of demanded stock	64	761	animals per year
N_S	number of suppliers	193,354	5,480	none
N_D	number of demanders	118,503	10,619	none
$N_{S\setminus D}$	number of strict suppliers	88,761	1,314	none
$N_{S\cap D}$	number of wholesalers	104,593	4,166	none
$N_{D\setminus S}$	number of strict demanders	13,910	6,453	none

6.3.1 Market dynamics without shocks - effect of trade friction

The FTM model introduces the notion of imperfect transactions with friction, bringing together differing economic models (see Section E). To assess the impacts of friction on transient and long-run trade dynamics *per se*, we analyse the FTM model in the absence of epidemic shock (system (6.9)). We start by exploring the reference market where stock losses and external flows are neglected. We then analyse the cases of non-negligible stock losses ($r_S \neq 0$ or $r_D \neq 0$) and symmetric imports ($E_S = E_D = E \geq 0$). In all cases, transients and steady-states are investigated analytically and numerically.

Trade without stock loss : the reference market

In the *reference market* we neglect stock loss ($L_S = L_D = 0$, i.e. $r_S = r_D = 0$) and external flows ($E_S = E_D = 0$), and denote the steady-state equilibrium value of state variables with a *star* in superscript. This market has an infinite number of steady-state equilibria (see Section E.3 for proof) :

$$\begin{aligned} &(S^* = S^*_{min}; D^* \geq D^*_{min}; p = p^*) \ , \\ &(S^* \geq S^*_{min}; D^* = D^*_{min}; p = p^*) \ , \end{aligned} \tag{6.17}$$

where $S^*_{min} = \kappa N_S$ and $D^*_{min} = \kappa N_D$ are the minimal supply and demanded stocks for which the market is in steady-state equilibrium, and $p^* = (\frac{N_D \delta_0}{N_S \sigma_0})^{\frac{1}{\varepsilon_S + \varepsilon_D}}$ is the unique equilibrium price, obtained by solving $\Sigma_\oplus(p^*) = \Delta_\oplus(p^*)$. The steady-state equilibria in supply and demand depend on the initial

conditions (see Section E.3) ; there is hence an infinite number of unstable equilibria (S^*, D^*) with a switched fixed point : either $S^* = S^*_{min}$ or $D^* = D^*_{min}$. Since p^* is unique, trade flow at steady-state equilibrium is unique and given by :

$$\Phi^* = \Sigma_\oplus(p^*) = \Delta_\oplus(p^*) = [N_S\sigma_0]^{\frac{\varepsilon_D}{\varepsilon_S+\varepsilon_D}}[N_D\delta_0]^{\frac{\varepsilon_S}{\varepsilon_S+\varepsilon_D}} \ . \tag{6.18}$$

The market converges asymptotically to reference flow ($\Phi(t) \to \Phi^*$) and price ($p(t) \to p^*$) for any initial conditions or external perturbations (see Section E.3 for proof). The famous law of supply and demand (LSD) is a particular case (see Section E.3 for proof). The LSD implies that supply should equal demand (both in terms of accumulated stocks and flows of stock) when price is equilibrated ; which is a very special case of our model with unique steady-state equilibrium $S^* = D^* = \max\{S^*_{min}, D^*_{min}\}$ (Fig. E.2D). We study this case analytically, and then return to the general model formulation (system (6.9)) as there is little empirical support for the LSD [McCauley, 2009].

Trade with stock loss : detailed analysis of a special case

To study market transient behaviour, we consider initial conditions and parameter values that enable us to solve system (6.9) analytically. We set $[S(t_0) = D(t_0) \geq 0; p(t_0) = p^*]$ at initial time t_0 and track trade flow Φ until steady-state equilibration. This set of initial conditions is compatible with the LSD since $S(t_0) = D(t_0)$ and $p(t_0) = p^*$. Once accumulated over time at rates $\Sigma_\oplus$ and $\Delta_\oplus$, supply and demanded stocks are converted through trade (Φ) and losses (at rates $r_S S$ and $r_D D$). For simplicity, we consider symmetrical losses ($r_S = r_D = r$). In this case, equations (6.9) are symmetrical, which, with the above initial conditions ensures stocks remain symmetrical ($S(t) = D(t)$ for $t \geq 0$) and price remains constant, $p(t) = p^*$ (see Section E.3). Therefore, system (6.9) reduces to :

$$\dot{S} = \Phi^* - (r + \frac{a\Phi^*}{\kappa})S \ , \tag{6.19}$$

where $a = \min\{\frac{1}{N_S}, \frac{1}{N_D}\}$ is a dimensionless constant, and can be solved analytically to give (see Section E.3) :

$$\Phi(t) = \Phi^* \frac{\Phi^*}{\Phi^* + \frac{r\kappa}{a}}[1 - e^{-t(r+\frac{a\Phi^*}{\kappa})}]. \tag{6.20}$$

In the long term ($t \to \infty$), flow diverges ($\Phi \to \infty$) if $r \leq -\frac{a\Phi^*}{\kappa}$, while, if $r > -\frac{a\Phi^*}{\kappa}$ flow converges to :

$$\Phi_{eq} = \Phi^* \frac{\Phi^*}{\Phi^* + \frac{r\kappa}{a}} \ , \tag{6.21}$$

where eq in subscript denotes steady-state equilibrium values in the general case (in contrast with the special case of the reference market where steady-state equilibrium is denoted by a star in superscript). In the more realistic case where losses are strictly positive ($r > 0$), trade flow is sub-optimal ($\Phi_{eq} < \Phi^*$),

i.e., friction κ and r have an overall negative impact on steady-state equilibrium flow. If, however, the losses were negative with $r \in]-\frac{a\Phi^*}{\kappa}, 0[$, trade flow would be over-optimal ($\Phi_{eq} > \Phi^*$). When losses are negligible ($r = 0$), we recover the reference market flow ($\Phi_{eq} = \Phi^*$).

Trade with stock loss : the general case

We now consider the more general case of a market with asymmetric positive losses ($r_S \geq 0$ and $r_D \geq 0$), and no external flows ($E_S = E_D = 0$), and consider arbitrary initial conditions $[S(0); D(0); p(0)]$ (see Section E.3 for derivations). From (6.9), we deduce that the steady-state equilibrium flow is always suboptimal :

$$\Phi_{eq} \leq \min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} \leq \Phi^* = \min\{\Sigma_{\oplus}^{*}, \Delta_{\oplus}^{*}\} \ . \tag{6.22}$$

Then, two cases arise. When the steady-state equilibrium limit of the average transaction stock $q_{eq} \equiv \min\{\frac{S_{eq}}{N_S}, \frac{D_{eq}}{N_S}\}$ is bounded by the per-agent supply (i.e. $q_{eq} = \frac{S_{eq}}{N_S}$), we find :

$$\Phi_{eq} = \Sigma_{\oplus}^{eq} \frac{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\}}{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} + r_S \kappa N_S} \ . \tag{6.23}$$

Conversely, when q_{eq} is bounded by the per agent demand ($q_{eq} = \frac{D_{eq}}{N_D}$), we have, by symmetry, that q_{eq} is given by (6.23) but with N_S and r_S replaced by N_D and r_D respectively. Importantly, equation (6.23) generalises the special case of equation (6.21) and, likewise, imply that r_S, r_D and κ have a negative impact on steady-state equilibrium flow. Since the dynamics with stock loss are not fully analytically tractable in the general case, we resort to extensive numerical simulations to confirm the key influence of r_S, r_D and κ on trade dynamics (see Global Sensitivity Analysis (GSA) of the FTM model in the Section E.3).

Trade with external flows

To examine the impact of external flows on market dynamics, we consider, for simplicity, positive and symmetric external inflows ($E_S = E_D = E > 0$) and neglect losses ($r_S = r_D = 0$). Symmetry ensures that the steady-state equilibrium price and trade rate are the same as in the reference market, $p_{eq} = p^*$ and $\Theta_{eq} = \Theta^*$; while trade flow and average stock exchanged per transaction increase to $\Phi_{eq} = \Phi^* + E$, the solution to $\frac{dS}{dt} = \frac{dD}{dt} = 0$, and $q_{eq} = \kappa \frac{\Phi^* + E}{\Phi^*} > \kappa$, as $q = \Phi/\Theta$. Similarly to (6.17), steady-state equilibrium supply and demanded stocks have two infinite sets of possible values : either $S_{eq} = \kappa N_S \frac{\Phi^* + E}{\Phi^*}$ and $D_{eq} \geq \kappa N_D \frac{\Phi^* + E}{\Phi^*}$, or $S_{eq} \geq \kappa N_S \frac{\Phi^* + E}{\Phi^*}$ and $D_{eq} = \kappa N_D \frac{\Phi^* + E}{\Phi^*}$. Hence, external flow increases trade flow through the average stock exchanged, but not the transaction rate, which is determined by number of agents, trade friction, and price.

Numerical illustrations

To confirm and extend the analytical insights on the impact of trade friction and losses on trade dynamics, we now explore the market model numerically. We use $[S(0) = 0; D(0) = 0; p(0) = 1.2\ p^*]$ as initial condition, and consider symmetric losses ($r_S = r_D = r$) and no external flows ($E_S = E_D = 0$). As the initial price is not equilibrated, trade flow equilibrates over time in a way that depends on market characteristics such as trade friction κ and stock loss rate r (Fig. 6.2). Increasing κ, drastically slows down market steady-state equilibration (Fig. 6.2A-B). In addition, the steady-state equilibrium flow depends on r. Without stock losses ($r = 0$), flow converges to the reference market level Φ^* (Fig. 6.2A). However, equilibrium flow is sub-optimal, $\Phi_{eq} \leq \Phi^*$, when there are positive losses ($r > 0$) and friction is large enough ; conversely, flow is over-optimal when losses are negative $r < 0$ and friction is large (Fig. 6.2B). Scenarios with a wide range of loss rates r_S and r_D, external flows E_S and E_D, and initial conditions, explored via GSA, confirm these findings (see Section E.3), which also agree with our previous analytical findings ((6.21) and (6.23)). Note that a negative loss rate corresponds to exponential inflows of supply and demanded stocks, a scenario that may appear to contrast real markets ; we include it to show the general scope of the model. Overall, we find that, in the model, friction can increase market steady-state equilibration time by several orders of magnitude, while stock loss and external stock flow can alter the long-term steady-state equilibrium of the market. We expect these parameters to play a central role in understanding trade dynamics in markets and, therefore, in the epidemiology of trade-driven diseases.

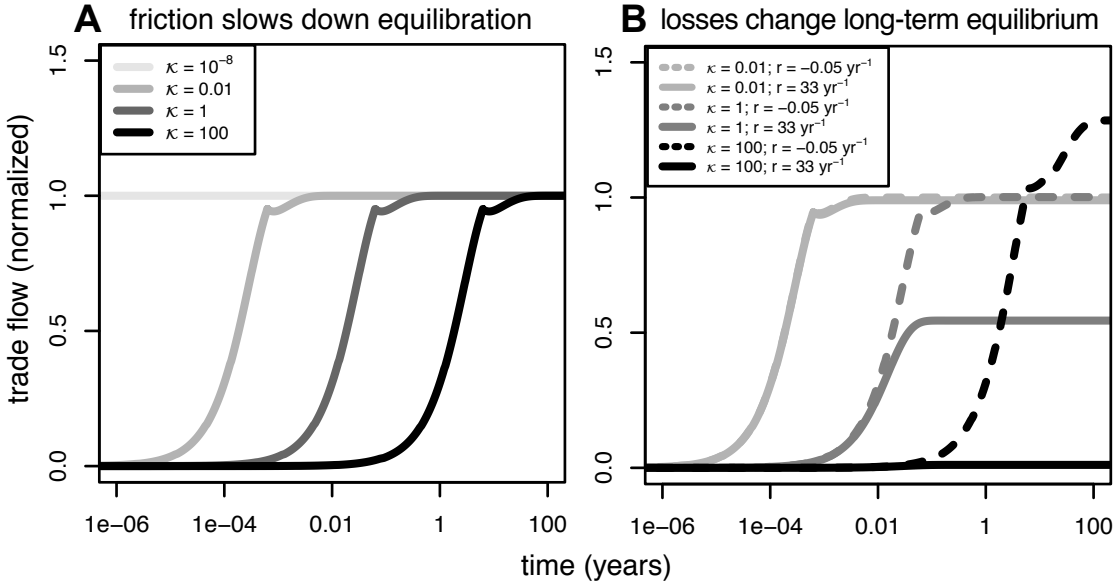


FIGURE 6.2 – **The influence of frictional-trade on transient and long-term market dynamics.** Evolution of normalised trade flow ($\frac{\Phi(t)}{\Phi^*}$) for variable levels of friction (from $\kappa = 10^{-8}$ in light gray to $\kappa = 100$ in dark) when losses are negligible ($r = 0\ yr^{-1}$, **A**) or non-negligible ($r \in \{-0.5, 33\}\ yr^{-1}$ in dashed and plain lines respectively, **B**). Initial conditions are set to $[S(0) = 0, D(0) = 0, p(0) = 1.2p^*]$. We take $\mu = 10^{-6}$, $\varepsilon_S = 0.4$ and $\varepsilon_D = 0.8$. Other parameters are set from the French cattle estimates (see Table 6.3).

6.3.2 Market dynamics with epidemic shocks

We now explore the dynamics of the ME model, where epidemics and trade dynamics can influence each other (Fig. 6.1A-C and system (6.14)).

Relative impacts of trade friction and adaptive risk aversion for contrasted delays in enforcement of regulation

We parameterise our model to investigate epidemics with contrasted delays in enforcement of regulation, and explore the impacts of frictional-trade and RA behaviour on epidemic dynamics (Fig.6.3). We assume that the market is has reached steady-state equilibrium before epidemic onset and trade is the only path of pathogen transmission. The trade-transmission rate has either a frictionless-market value β_{tr}^0 (equation (6.16) ; grey in Fig.6.3) or its corresponding frictional-market value β_{tr} (equation (6.12) ; black in Fig.6.3). When market friction is very low (Fig.6.3A-C), infection reaches the same endemic level with or without friction. The inclusion of RA (dashed lines) reduces the number of infected agents, and does so similarly with or without friction (Fig.6.3A-C). This reduction in infections in the frictionless market is in agreement with the literature (e.g. [Funk *et al.*, 2010]), and is expected as RA decreases the force of infection in response to an outbreak (6.11). As expected intuitively, increasing the delay in enforcement of regulation results in a lower effect of RA (Fig.6.3A-C). When market friction has a significant level and provided the speed of regulation is quick enough (Fig.6.3B), the endemic level is considerably lower in the frictional than in the frictionless market. Compared with the frictionless market, delaying regulation enforcement does not change the endemic level but slows down the epidemic (Fig.6.3D). Again, the inclusion of RA has a similar effect on the endemic level with and without friction, but this is comparatively less important than the effect of a significant increase in friction (κ from 0.01 to 1, Fig.6.3A-C and Fig.6.3B-D). Overall, our results suggest trade friction can mitigate trade-driven disease transmission significantly, possibly more than in Fig.6.3B-D, and to a significantly greater extent than RA behaviour, as our analyses of cattle market data suggest $\kappa = 3.4 > 1$. In addition, the combined effects of market friction and RA can lead to epidemic elimination (Fig.6.3B) when trade is the only pathway of transmission and the delay to enforce regulation is small enough.

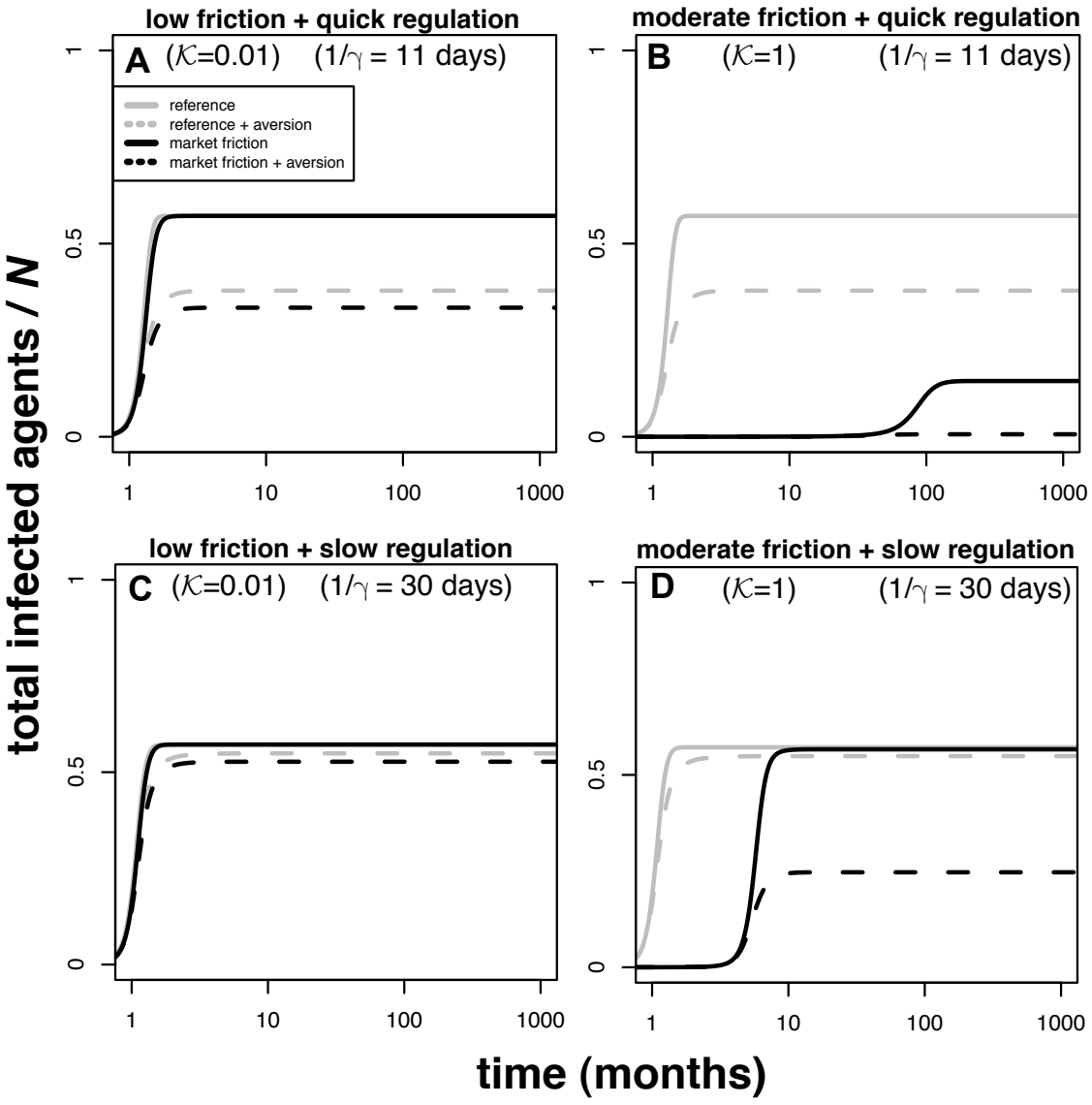


FIGURE 6.3 – **Impacts of frictional-trade dynamics with risk aversion for contrasted delays in enforcement of regulation.** Evolution of total infected agents normalised by the number of agents as function of time for various types of frictional-trade dynamics with risk aversion and contrasted delays in enforcement of regulation. We assume trade is the only path of transmission ($R_0^{tr} = R_0$) and parameterise our model to replicate an XYZ outbreak ($\phi = 0.9780$, $\gamma > 0$, $\nu = 0$). Friction is either low ($\kappa = 0.01$, **A-C**) or moderate ($\kappa = 1$, **B-D**). The trade transmission rate is either set to its frictionless reference value β_{tr}^0 (in grey ; see also (6.16)) or to its frictional-market value β_{tr} (in black ; see also (6.12)). RA is either negligible ($\alpha = 0$, plain lines) or non-negligible ($\alpha = 8$, dashed line). Enforcement of disease regulation is either quick ($1/\gamma = 11$ days, **A-B** ; corresponding to a reference value of $R_0 = 4$ without friction) or slow ($1/\gamma = 30$ days, **C-D** ; corresponding to a reference value of $R_0 = 11$ without friction). Initial conditions are set to start from a steady-state equilibrated market $[S(0) = \kappa N_S \frac{\Phi^* + E}{\Phi^*}, D(0) = \kappa N_S \frac{\Phi^* + E}{\Phi^*}, p(0) = p^*]$ and one agent per market category is initially infected (one strict supplier, one wholesalers, one strict demander). Other parameters are as in Fig. 6.2.

Impact of market friction level on epidemics

Increasing trade friction reduces the severity (Fig.6.3 and GSA of the ME model in Section E.4) and magnitude of the peak (Fig.E.5 and GSA in Section E.4) of epidemics when trade routes are the only pathway of transmission. Trade friction is also a key determinant of the epidemic threshold as assessed by the basic reproduction number R_0, a fundamental epidemiological summary. R_0 is the average number of susceptible agents infected by a single infectious agent propagated in an initially disease-free agent population [Anderson & May, 1991]. In a deterministic framework, the pathogen eventually dies out if $R_0 \leq 1$; while if $R_0 > 1$, the pathogen eventually invades the population. In the general case with both trade and non-trade pathogen transmission (6.11), R_0 is given by (see Section E.4) :

$$\begin{aligned} R_0 &= \frac{R_0^{tr} + R_0^{\overline{tr}} + \sqrt{(R_0^{tr} - R_0^{\overline{tr}})^2 + 4R_0^{tr}R_0^{\overline{tr}}\frac{N_S N_D}{N N_{S\cap D}}}}{2} \ , \\ R_0^{tr} &= \frac{\beta_{tr}}{\gamma}\frac{N_{S\cap D}}{N_S} \ , \\ R_0^{\overline{tr}} &= \frac{\beta_{\overline{tr}}}{\gamma} \ , \end{aligned} \tag{6.24}$$

where R_0^{tr} ($R_0^{\overline{tr}}$ respectively) is the value of R_0 when trade (non-trade) provides the only pathway of pathogen transmission. When trade is the only transmission route ($R_0^{\overline{tr}} = 0$), and inserting expression (6.12) for β_{tr}, yields (noting that $N_D^{XY} = N_D$ in this context) :

$$R_0 = \frac{P_{tr}(q_{eq})}{\gamma}\frac{N_{S\cap D}\Theta_{eq}}{N_D N_S} \leq \frac{1}{\gamma}\frac{N_{S\cap D}\Phi^*}{\kappa N_D N_S} \ , \tag{6.25}$$

since $P_{tr}(q_{eq}) \leq 1$ and $\Theta_{eq} = \dfrac{\min\{\Sigma_\oplus^{eq}, \Delta_\oplus^{eq}\}}{\kappa} \leq \dfrac{\Phi^*}{\kappa}$. Therefore, R_0 vanishes in the limit when the market friction is large :

$$\lim_{\kappa\to\infty} R_0 = 0 \tag{6.26}$$

Result (6.26) stands for *any* modelling choice for $P_{tr}(q)$, including our current $P_{tr}(q) = [1 - (1-\phi)^q]$. In addition to its mechanistic interpretation, this choice has the advantage of yielding a finite value for R_0 (equation (6.25)) in the limit of negligible friction ($\kappa \to 0$, when $\beta_{tr} \to \beta_{tr}^0$ given by (6.16)) :

$$R_0 = \frac{\ln(\frac{1}{1-\phi})}{\gamma}\frac{N_{S\cap D}\Phi^*}{N_D N_S} \ , \tag{6.27}$$

which allows comparison with existing epidemiological models that implicitly assume $\kappa = 0$. Increasing trade friction can cause decrease in R_0 up to the critical point where $R_0 < 1$ (Fig. 6.4A). Provided that the delay in enforcing regulations is small enough, this result also stands when trade is not the main transmission pathway (Fig. 6.4B). Therefore, accounting for trade friction is central to the estimation of epidemic thresholds in markets. This finding is confirmed by a GSA of R_0 in response to variation of its composing parameters (see GSA in Section E.4). We can also use our expression of R_0 to rank

the relative risk of sustaining an epidemic for various markets. As an example, French swine markets are characterized by a larger coefficient of friction (κ = 71.7; see Table 6.3) than the French cattle market (κ = 3.4). Since R_0^{tr}(swine)/R_0^{tr}(cattle) ≈ 0.8 for ϕ and γ kept constant, trade of swine is less likely to sustain epidemics than trade of cattle. This result would appear counter-intuitive for typical epidemiological models, because trade flow is larger in swine than cattle (Table 6.3).

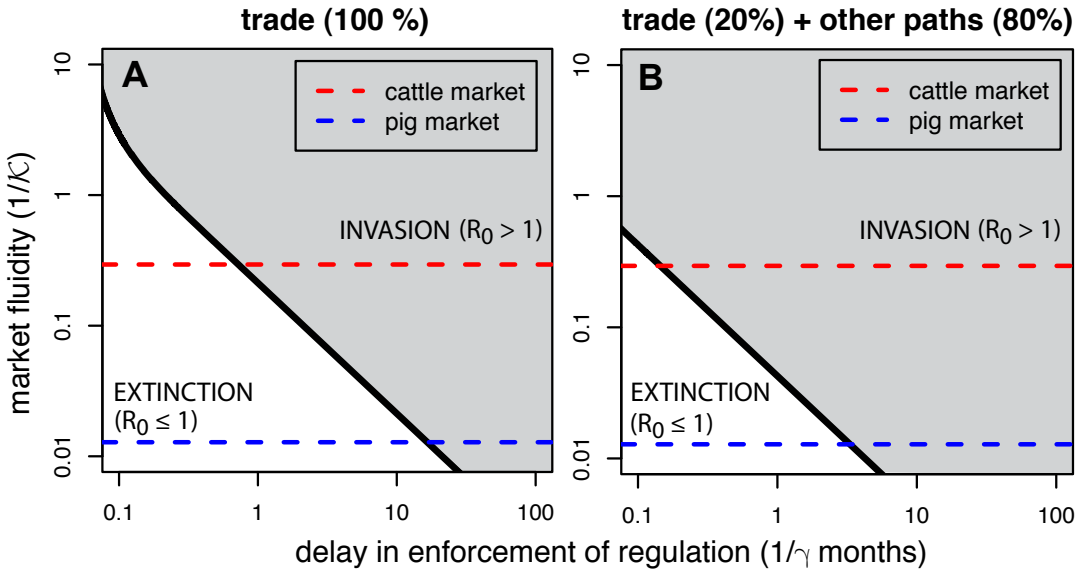


FIGURE 6.4 – **Maximal delay in enforcement of regulation that still allows prevention of epidemics depending on market fluidity and other paths of transmission.** Probability of invasion P_I for various levels of market fluidity $\frac{1}{\kappa}$, delays in enforcement of regulation $\frac{1}{\gamma}$ and intensity of other paths of transmission (**A-B**). For each panel, the black curve represent the equation $R_0 = 1$ that separates the $(\frac{1}{\kappa}, \frac{1}{\gamma})$ space into two subspaces : the area under (above) the curve leads to an extinction (invasion) of the disease, i.e. $P_I = 0$ ($P_I = 1$). The epidemic is either only caused by trade (the black curve is given by : $R_0^{tr}(\kappa_c, \gamma_c) = 1$ $R_0 = 1$; **A**) or by trade and other paths of transmission (the black curve is given by : $R_0^{tr}(\kappa_c, \gamma_c) = 0.2$ $R_0 = 0.2$; **B**), where $\frac{1}{\gamma_c}$ is the maximal delay to enforce regulation corresponding to the critical amount of friction $\frac{1}{\kappa_c}$, the only value of $\frac{1}{\kappa}$ for which $R_0 = 1$. The dashed red (blue) line represent the estimated level of market fluidity of the French cattle market (French swine market). Other parameters are as in Fig. 6.3.

Open international market versus closed national market

When we consider an open market by including imports in the model, we find that international trade can boost epidemics moderately in comparison with closed national markets (Fig. E.6). Imports increase stocks, and thus the average stock exchanged q (6.6) and probability of infection $P_{tr}(q) = 1 - (1 - \phi)^q$ per transaction. However, imports do not affect the transaction rate Θ. Therefore, as the force of infection involves the product $P_{tr}(q)\Theta$ and $P_{tr} \leq 1$, the effect of imports on the force of infection, and thus on epidemics, is limited. Another limiting factor is the current level of control measures, i.e. removal of infected agents after a given period of infectiousness.

Markets with different types and intensities of disease regulation

In our model, regulation to control or prevent disease spread in markets can be implemented in two ways : by removing infected agents at rate γ or by limiting their re-introduction, after sanitation, at rate ν. We show that epidemics can be mitigated by increasing the exit rate of infected agents and/or decreasing their reentry rate (Fig. 6.3 and Fig. E.6B). The GSA suggests that increasing the removal rate γ may be more efficient at mitigating disease than decreasing the rate of agent re-introduction ν (see Section E.4). In particular, epidemics can be mitigated (Fig. 6.3B-D) and even eradicated by increasing γ (Fig. 6.4A-B). Our analytical summaries of R_0 (equations (6.24)) provide estimates of the maximal delay in the enforcement of regulation $1/\gamma_{max}$ that still allows prevention of epidemics for various types of markets and combinations of transmission pathways.

6.4 Discussion

Market trade routes propagate epidemics differently from other transmission pathways due to the unique contact structure that emerges from the willingness of agents to sell and buy goods (Fig. 6.1). This willingness arises from the business motivation of economic agents and inherent features of trade dynamics, which may differ from other epidemic-conducive human behaviour. Trade markets involve recurrent interaction events (transactions) between suppliers and demanders (consumers are demanders), and in each transaction there is a variable volume of goods exchanged (q) that may be contaminated and lead to the spread of infection. This notion of 'transaction' is somewhat different from that in the economics literature, particularly from the notion of 'match' in labour economics (see Section E.1). Epidemiologically, it essential to precise this concept beyond an economic perspective, because transactions of goods, in addition to their market function also provide a dynamic contact structure that can support disease transmission among agents. In our model, the frequency of trade, i.e., transactions, is limited by multiple constraints that we summarise in a single coefficient of friction, κ. Causes of such friction include production of exchangeable goods, the search for a trading partner, and the logistics of stock delivery. Beyond questions about interpretation, this coefficient can be easily measured as a characteristic feature of markets ; we have estimated κ relatively easily from data on trade flow of livestock (see Section E.2 for application to exchanges of swine and cattle in France). In relation to the concepts in economics of matching friction, and other types of friction induced for example by transaction costs (see Table 1), the concept of trade friction used by us accounts additionally for production constraints, e.g., that the minimum tradable unit is one living animal. This tangible mechanism has significant epidemiological consequences, as well as economic implications, and is absent in models that implicitly assume the exchange of a continuous amount of goods, including 'infinitesimal pieces of animals'. We believe that the notion of frictional dynamical transactions is an improvement over existing market models (see Section E.1 for details) that may help to better understand the interactions between trade

and disease epidemiology for many infectious diseases.We believe that the notion of frictional dynamical transactions is an improvement on existing market models (see Section E.1 for details) that may contribute to a better understanding of the interactions between trade and the epidemiology for many infectious diseases.

Taken together, our findings show that frictional markets are associated with a specific response to infectious diseases that contrasts with the response of other complex systems that sustain epidemics and are often assumed to be frictionless. In particular, the coefficient of friction κ is a central parameter governing trade and disease dynamics : κ can increase market steady-state equilibration time by several orders of magnitude (Fig. 6.2) and suppress trade-driven disease transmission to a significantly greater extent than RA behaviour (Fig. 6.3). Our findings hence suggest that while frictional markets can still be responsible of disease introductions, trade itself cannot sustain epidemics in such markets. This counter-intuitive implication should be taken with care until confirmed by subsequent studies grounded on agent-based models.

The outcomes of our model suggest that to minimise contagion in markets, κ could be increased to allow for larger-volume, less-frequent transactions, without necessarily affecting overall trade flow, and therefore, business activity. However, increasing friction may be difficult to achieve in practice due to practical constraints underlying κ (e.g. trucks have a limited size). The model also suggests that in markets with a given level of friction κ, international trade (Fig. 6.3C) and regulatory measures (Figs. 6.3D and 6.4A-B) have strong, but contrasting, influence on trade-driven epidemics. In today's globalised world, a key question is how to mitigate epidemics efficiently in open international markets. The GSA of the ME model suggests that increasing the intervention rate γ is more efficient at mitigating disease than decreasing the business re-establishment rate ν, irrespective of imports (see Section E.4). In other words, an alternative policy for minimising contagion in markets could adjust the delay of intervention $1/\gamma$ depending on κ.

From a cross-disciplinary perspective, the FTM model can be used to approximate market dynamics when there is lack of empirical data.The model can also be used to predict the impact and potential non-equilibrium in trade dynamics caused by disturbances such as unexpected disease outbreaks, for example in livestock markets (Fig. E.7). Three forms of non-equilibrium are accounted for by the FTM model (see Table 6.1 for detail). One form is when the rates of change in supply stock, demanded stock, and price are not zero, i.e. when the system is not in steady-state or inter-temporal equilibrium. A second form is when there is disequilibrium in state variables at a given time such as a mismatch between supply and demanded stocks of trading agents ; this may be caused by past disturbances such as epidemic outbreaks and abrupt loss or increase in supply or demanded stocks. We allow for this possibility by limiting the stock exchanged in each transaction to the match between supply and demanded stocks (q). A third form of equilibrium is when there is a difference between the rates of generation of supply stock and demanded stock at given market price, which is a form of imperfect

competition, which could be caused by epidemics or market failures. The level of competitiveness in our model is influenced by parameter μ, which sets a timescale for price dynamics but does not necessarily preclude an eventual steady state (see Table 1 for further details) We indeed expect that disease epidemics where transmission is predominantly driven by trade are likely to cause asymmetric economic shocks, in the form of persistent excess of either supply or demanded stock. For example, quarantine measures could cause limited accumulated supply and thus excess accumulated demanded stock, while disease-control measures or consumer fright could cause excess accumulated supply of goods. Moreover, with the exception of wholesalers, disease transmission will occur from suppliers to demanders rather than in the opposite direction (Section E.4). It follows that suppliers and demanders would generally be differently impacted by disease, and would also be differently targeted by disease-control measures. The specific outcomes would depend on the specific market and disease. Therefore, accounting for non-equilibrium in trade dynamics is of central importance for understanding market-epidemic feedbacks. However, some approaches have often assumed that all transactions occur without mismatch between supply and demand [e.g. Mas-Colell *et al.*, 1995], i.e. without focusing on differences between stocks and flows that are crucial in a market-epidemiological context. In other words, these studies consider that there is no accumulation of stock of goods or of demand for goods over time by implicitly assuming that both suppliers and demanders are fully satisfied through transactions where there is perfect match between what is available for sale and what is wanted for purchase. In contexts of market disturbance caused by disease outbreak it is essential to allow for such unbalance over the time scale of the outbreak or longer. Hence, following disequilibrium approaches from economics Fair & Jaffee [1972], we have allowed for imperfect match between supply and demanded stocks ; this required defining accumulated stocks in addition to what we called flows (or 'rates of generation') of supply and demanded stock, which, formally, are the default supply and demand in economics (Table 6.1). We note that while 'demanded stock' seems not to be considered in economics, there are good reasons for considering that demand can accumulate over time when we do not assume the current needs of buyers are fully satisfied in one time step but remain to be satisfied over an economically-significant period of time. For example, farm owners who experience accumulating losses of livestock due to an ongoing undetected infectious disease, would increase demand for livestock for sometime to replace ongoing animals deaths ; alternatively, their would interrupt business ; either route would have a lasting unbalancing influence on the wider market.

Further on the usefulness of a market-epdemiological model, predictive assessments of market resiliency to disturbance such as disease outbreak are important for health economics. However, economic models applied to the study of epidemics often focus on descriptive assessments that neglect temporal dynamics (e.g. [Soliman *et al.*, 2010]) or predictive equilibrium-based discrete-time approaches (e.g. [Zhao *et al.*, 2006]) that can be at odds with the multiple and contrasting time scales encompassed by economic and epidemiological processes. It is important to model in continuous time to adequately

resolve the multiple processes of infection and the market transactions. Even when economic model of trade operate in continuous time, they apparently do not consider explicit contact structures among economic agents (e.g. [Caputo, 2005]), which render them inapplicable to describe epidemics. Epidemiologists, ecologists and network-scientists have a long-standing expertise in formalising such contact structures, which suggest the tools developed by those communities could also benefit to the study economic systems, especially when submitted to systemic shocks (see also [May *et al.*, 2008]). The use of R_0 can also be of importance in economics ; for example, in assessing whether trade can drive epidemics that have the potential to damage trade. Specifically, the ME model shows cases where, depending on whether R_0 is less or greater than 1, contagion in markets can lead to either full recovery or long-lasting recession of the economic sector. However, interestingly, when considering the dynamics of a perturbed market per se, i.e. the FTM model without coupling to an epidemiological model, we typically find that the market is robust to disturbance as it returns rapidly to the original steady-state equilibrium. A fundamental aspect, therefore, lies in the feedback between dynamical processes that are often left separated because they associate with different disciplines or research communities. We note, however, that R_0 has limitations in its usefulness, and in particular is not a good predictor of the severity of an outbreak. For example, our model suggests that in face of an epidemic scenario ($R_0 > 1$) (6.25) swift regulation can eliminate the epidemic and mitigate the impact on the market when there is risk aversion and moderate friction, but not without risk aversion (Fig. 6.2B and Fig. S6 black-dashed versus plain-black curves). However, risk aversion is a behavioural adaptation that occurs over the course of the epidemic, and as such is not incorporated in the calculation of R_0, which focuses on the early stages of the epidemic. Another example is the case of an outbreak occurring in markets outside of equilibrium (i.e. prior to supply stock, demanded stock, and/or price having reached steady-state equilibrium. We find that the market's initial supply and demanded stocks generally have minor impact on trade and epidemics. However, an initially non-equilibrated price can strongly influence the size of the outbreak until the market reaches a steady-state (GSA of the ME model ; Section E.4). Therefore, the same epidemiological system, characterised by a single R_0 value, can lead to contrasted epidemic and market outcomes, depending on adaptive risk-aversion behaviour and the market conditions at the time of epidemic onset. Both of these examples highlight the importance of accounting for adaptive mechanisms of interacting but apparently disparate dynamics ; which bares resemblance to ecological dynamics.

From an economic perspective, our study only stands for emerging or re-emerging epidemics since agents adapt their behaviour in response to outbreaks but do not really anticipate them. Further work of interest would transpose the current ME model to endemic diseases that are often anticipated by market agents. Agent anticipation could be implemented using explicit inter-temporal optimisation functions that are widely used in economics (e.g. by [Zhao *et al.*, 2006] ; see Section E.3.1 for details and presentation of other potential extensions inspired by the economic literature). In addition, we

could explicitly account for inventories and capacities of economic agents [Khan & Thomas, 2003], for example, by dissociating animals within a farm (an inventory) from animals for sell within a farm (a supply stock) ; or by dissociating the maximal number of animals that can be present within a farm (a capacity) from the number of animals expected to be bought (demanded stock). Our model can also be extended to account for heterogeneous contact structure such as individual-based networks of trading agents (e.g. [Atalay *et al.*, 2011]). An explicit trade-agent model would account for transactions involving identified pairs of agents who jointly decide to exchange goods against money. For simplicity, this pairing process is assumed here to be governed by a homogenous mixing process. However, where trade flow and the number of business partners are positively correlated at agent level, which is likely in many cases, we expect heterogeneous contact structure to boost epidemic development in comparison with homogeneous settings [Kamp *et al.*, 2013]. A key open question we are currently investigating concerns the conditions under which realistic levels of friction can also mitigate epidemics propagated on heterogeneous markets.

Acknowledgments

Many thanks to Samuel Alizon, Olivier Allais, Hugues Beyler, Caroline Bidot, Pauline Ezanno, Yann Kervinio, Julien Fosse, Christine Fourichon, Etienne Geoffroy, Natacha Go, Bhagat Lal Dutta, François Moutou, Emilie Moyne, Marco Pautasso, Louis-Georges Soler and Stéphane Robin for helpful comments and insights, and to two anonymous reviewers for helpful comments on a previous draft. We are grateful to the French Ministry in charge of Agriculture and to the professional union BDPorc for granting us access to the cattle and swine datasets respectively. EV, HM and MML would like to thank the French Ministries in charge of Agriculture and Environment, the INRA MIA Department for financial and operational support. This study benefited from a Dufrenoy grant from the French Academy of Agriculture. JANF and CAG were funded by DEFRA and USDA.

Chapitre 7

Modelling trade and trade-driven epidemics in markets with heterogeneous agents

As a direct extension to chapter 6, this chapter presents preliminary approaches and findings to address specific questions arising when various sources of heterogeneities in the characteristics of agents and their interactions are taken into account. Most results we present were obtained by Sébastien Geeraert, as part of a Master internship carried out under our supervision (see appendix F, in French). Our personal contributions lies in the formulation of models of trade and trade-driven epidemics in markets with heterogeneous agents, conception of numerical experiments to fill current knowledge gaps, and suggestion of modelling refinements in the light of Sébastien Geeraert's work.

While providing interesting insights on the economic mechanisms likely to drive epidemics in trade pathways, most importantly trade friction, the frictional trade market (FTM) and the market–epidemiological (ME) models explored in chapter 6 rely on the strong assumption that suppliers (demanders) are all identical to a representative supplier (demander). Though necessary to develop an overall intuition on the mutual influence of exchange and epidemic dynamics, assuming markets are composed of homogeneous agents appears at odds with trade data [e.g. Atalay *et al.*, 2011; Vernon, 2011]. In actual fact, economic agents and their interactions can exhibit marked differences in degree, weight, overall direction of exchange or friction, as typically observed in livestock-exchange markets propagating infectious diseases (chapters 5 and 6). We identified two knowledge gaps at the crossroad of economics and epidemiology that can be addressed by accounting for heterogeneous features encountered in markets : *i*) uncovering plausible behavioural mechanisms to explain and predict trade dynamics as observed with livestock exchanges ; *ii*) exploring to which extent a market with diverse friction can sustain infection transmission.

Here, we attempt to fill the identified knowledge gaps by integrating market and epidemiological

processes at the microeconomic/microscopic scale. We hence extend the macroscopic FTM and ME models to describe trade and trade-driven epidemics with adaptive behaviour in markets with heterogeneous agents. The resulting frameworks are called the *heterogeneous FTM* (hFTM) and *heterogeneous ME* (hME) models. When confronted with cattle-exchange data at the agent-level, the hFTM and hME models shed light on possible behavioural mechanisms underlying trade dynamics (section 7.1) and appear to confirm the leading influence of friction on economic-epidemiological dynamics (section 7.2).

7.1 A modelling framework to shed lights on trade dynamics

Adequate targeting of disease prevention and control measures requires to identify those agents most at risk of acquiring and/or transmitting infection. Network-based centrality measures are commonly used to assess such risks, but exchange networks required to calculate such measures may be unavailable due to missing data or delays in data collection [Lal Dutta *et al.*, 2014]. In the case of French livestock-exchange data, delays in data collection do not appear to affect recommended polices for preventing epidemics (chapter 5). But this particular result is unlikely to stand for any exchange network. To enable effective epidemic risk predictions in the general case, Valdano *et al.* [2014] propose a new framework where the loyalties of agents in their partnerships, as estimated from past data, are used to model the future shape of contact networks. The framework allows to adequately reproduce certain features of empirical dynamical exchange networks, but is not based on mechanistic assumptions. While adequate for disease prevention, such predictions may prove problematic when evaluating the efficacy of mitigation strategies. In contrast with preventive measures which attempt to prevent epidemics before they occur, mitigation strategies aim at reducing the severity of ongoing outbreaks. Since epidemics and exchanges are dynamical processes mutually influencing each other, models neglecting the mechanisms driving exchanges in the first place are unlikely to provide accurate epidemiological predictions for mitigation strategies. It follows understanding how exchanges operate at the agent-level remains fundamental to extrapolate exchange networks for disease control.

Here, we present a mechanistic modelling framework to shed light on a particular type of exchange at the agent-level, namely trade of goods. As a case study, we attempt to explain observed trade dynamics on a small subset of the French cattle-exchange market (section F.3). Assuming an absence of market disruption such as epidemics, we develop a deterministic network-based model of trade dynamics directly building upon the Frictional Trade Market (FTM) framework (chapter 6). We call this generalised FTM model the heterogeneous FTM (hFTM) model, which While we focus our applications on epidemics, we believe the hFTM can prove valuable to uncover the mechanisms underpinning adaptive trade dynamics. We explore the influence of contrasted behavioural assumptions on trade dynamics simulated with the hFTM model to assess which one(s) most closely match observed trade data. Our preliminary findings suggest the hFTM model is a good candidate model to infer future trade patterns and investigate joint exchanges and epidemic dynamics.

7.1.1 The heterogeneous Frictional Trade Market (hFTM) model

Here, we detail in turn the components of the hFTM model, from the microscopic production of supply and demand to equations governing the evolution of supply and demand stocks. In contrast with the macroscopic FTM model, transactions and exchanges are now explicitly described at the level of agents' pairs.

Microscopic production of supply and demand

By analogy with the FTM model, we define the rates of production in supply and demand of agent i, denoted σ_i and δ_i respectively, as :

$$\begin{aligned} \sigma_i(p) &= \sigma_0^i \left(\frac{p}{p_0}\right)^{\varepsilon_S} , \\ \delta_i(p) &= \delta_0^i \left(\frac{p}{p_0}\right)^{-\varepsilon_D} , \end{aligned} \tag{7.1}$$

where $\sigma_0^i \geq 0$ and $\delta_0^i \geq 0$ are the baseline production rates in supply and demand of agent i, $\varepsilon_S \geq 0$ and $\varepsilon_D \geq 0$ are the price elasticities of supply and demand, p is the unit price of the focus good traded and p_0 is the numéraire, i.e. the price of one unit of money (in practice, we assume $p_0 = 1$ to simplify calculations without loss of generality).

By construction, a strict supplier (strict demander) is an agent i with $\sigma_0^i > 0$ and $\delta_0^i = 0$ (with $\sigma_0^i = 0$ and $\delta_0^i > 0$). A wholesaler is an agent i for which $\sigma_0^i > 0$ and $\delta_0^i > 0$. The market importance of an agent i is given by $(\sigma_0^i + \delta_0^i)$.

Elementary transaction rates, partitioning functions and microscopic friction

Let θ_{ij} represent the *elementary transaction rate* from seller i to demander j, i.e. the current number of transactions from seller i to demander j per time unit.

We denote by f_{ij} (g_{ij}) the share of σ_i allocated by i to j (the share of δ_j allocated by j to i). We recall the production rate in supply of i (in demand of j) is given by σ_i (δ_j) (7.1). By construction, we have $\sum_j f_{ij} = 1$ and $\sum_i g_{ij} = 1$, and we will hence refer to f_{ij} and g_{ij} as the *partitioning functions*.

Let κ_{ij} represent the *microscopic friction* associated with the directed link (i, j). As in the FTM model, we assume the elementary transaction rate from i to j is given by :

$$\theta_{ij} \quad = \quad \frac{\min\{f_{ij}\ \sigma_i; g_{ij}\ \delta_j\ \}}{\kappa_{ij}} \ . \tag{7.2}$$

Since an indivisible good cannot be further divided, the current maximal possible rate of exchange of indivisible goods from i to j is $\min\{f_{ij}\ \sigma_i\ ; g_{ij}\ \delta_i\}$, i.e. the value of θ_{ij} when $\kappa_{ij} = 1$. As in the FTM model, we notice friction $\kappa_{ij} \geq 1$ aggregates all barriers leading to relation $\theta_{ij} \leq \min\{f_{ij}\ \sigma_i\ ; g_{ij}\ \delta_i\}$. However, κ_{ij} is now defined at the scale of the directed pair of interacting agents (i, j). Under reasonable assumptions, κ_{ij} simply corresponds to the average number of goods exchanged per transaction from agent i to agent j, which can be easily calculated from available data. Finally, we can remark from (7.2) that the partitioning functions f_{ij} and g_{ij} are crucial quantities since they capture agents' preferences in interacting preferentially with certain trade partners.

Microscopic evolution of supply, demand and price

Let S_i and D_i ($S = \sum_i S_i$ and $D = \sum_i D_i$) represent the elementary stocks of supply and demand of agent i (aggregated stocks of supply and demand). As with the FTM model, we assume the batch size associated with transaction θ_{ij} is given by the maximum possible stock that can be exchanged from i to j, i.e. $q_{ij} = \min\{S_i; D_j\}$.

The hFTM model is finally given by :

$$\begin{aligned} \frac{dS_i}{dt} &= \sigma_i(p) - \sum_{j\neq i} \theta_{ij}\ q_{ij}\ , \\ \frac{dD_i}{dt} &= \delta_i(p) - \sum_{j\neq i} \theta_{ji}\ q_{ji}\ , \\ p(t) &= p(0) e^{\mu[D(t)-S(t)-(D(0)-S(0))]}\ , \end{aligned} \tag{7.3}$$

where $p(0)$, $S(0)$, $D(0)$ are the initial price and stocks of supply and demand, and $\mu \geq 0$ is a dimensionless coefficient. Note that price dynamics is still defined at the macroscopic level since agents are assumed to be price-takers.

Let $\Phi_i^{\text{out}} = \sum_{j\neq i} \theta_{ij} q_{ij}$ and $\Phi_i^{\text{in}} = \sum_{j\neq i} \theta_{ji} q_{ji}$ denote the out- and in-trade flow of agent i. System (7.3) is formally equivalent to :

$$\begin{aligned} \frac{dS_i}{dt} &= \sigma_i(p) - \Phi_i^{\text{out}}\ , \\ \frac{dD_i}{dt} &= \delta_i(p) - \Phi_i^{\text{in}}\ , \\ \frac{dp}{dt} &= \mu p \sum_k (\delta_k - \sigma_k), \end{aligned} \tag{7.4}$$

which is more amenable to analytical derivations.

7.1.2 Analytical insights on the hFTM model

We now explore the hFTM model analytically.

Recovering the macroscopic FTM model from the hFTM model

At the population level, the FTM model assumes that supply and demand rates are created at global rates $\Sigma_\oplus(p, N_S)$ and $\Delta_\oplus(p, N_S)$ respectively (6.4). The FTM model is an aggregated economic model, which implies that each supplier and demander is identical to an average representative supplier and demander respectively. By setting $\sigma_0^i = \sigma_0$ and $\delta_0^i = \delta_0$, and $\sigma_i = \sigma$ and $\delta_i = \delta$ in (7.1), we notice the correspondence between the FTM and the hFTM models :

$$\begin{aligned} \sigma(p) &= \frac{\Sigma_\oplus(p, N_S)}{N_S} = \sigma_0 \left(\frac{p}{p_0}\right)^{\varepsilon_S}\ , \\ \delta(p) &= \frac{\Delta_\oplus(p, N_D)}{N_D} = \delta_0 \left(\frac{p}{p_0}\right)^{-\varepsilon_D}\ . \end{aligned} \tag{7.5}$$

We then highlight the implicit microeconomic assumptions on θ_{ij} that lead to Θ, the aggregated transaction rate of the FTM model (6.7). We assume agents share their respective production equally among their business partners but without allowing self-loops (i.e. a wholesaler cannot exchange goods with himself/herself), which translates into :

$$f_{ij} = \frac{1}{N_D - \langle \mathbb{1}_{i\in\mathcal{N}_D} \rangle} = \frac{1}{N_D - \frac{N_{S\cap D}}{N_S}} \quad \text{(supplier } i \text{ to demander } j\text{)} , \tag{7.6}$$

$$g_{ij} = \frac{1}{N_S - \langle \mathbb{1}_{j\in\mathcal{N}_S} \rangle} = \frac{1}{N_S - \frac{N_{S\cap D}}{N_D}} \quad \text{(demander } j \text{ to supplier } i\text{)} , \tag{7.7}$$

where $\langle \mathbb{1}_{i\in\mathcal{N}_D} \rangle$ and $\langle \mathbb{1}_{j\in\mathcal{N}_S} \rangle$ represent the probabilities for the supplier i and the demander j, respectively, to be a wholesaler. We assume $\kappa_{ij} = \kappa$ as in the FTM model. It follows from (7.2) :

$$\theta_{ij} = \frac{\min\{N_S\sigma(p); N_D\delta(p)\ \}}{\kappa(N_S N_D - N_{S\cap D})} . \tag{7.8}$$

The macroscopic transaction rate Θ is obtained by summing the elementary transaction rates θ_{ij} over all supplier-demander pairs without self-loops :

$$\Theta = \sum_{i,j;j\neq i} \theta_{ij} = \theta_{ij}(N_S N_D - N_{S\cap D}) = \frac{\min\{N_S\sigma(p); N_D\delta(p)\ \}}{\kappa} , \tag{7.9}$$

which is the definition of the transaction rate of the FTM model.

Finally, by summing the hFTM equations for the evolution of supply and demand stocks (7.3) over the N_S suppliers and the N_D demanders, we immediately recover the macroscopic FTM model (6.9).

hFTM model equilibria

We explore the potential equilibria of the hFTM model and denote putative equilibrium values with an "eq" sign in exponent.

From (7.4), we can rewrite the equation for price dynamics as :

$$\frac{dp}{dt} = \mu p \left(\left[\left(\frac{p}{p_0} \right)^{-\varepsilon_D} \sum_k \delta_0^i \right] - \left[\left(\frac{p}{p_0} \right)^{\varepsilon_S} \sum_k \sigma_0^i \right] \right). \tag{7.10}$$

By setting $\sum_k \delta_0^i = N_D\delta_0$ and $\sum_k \sigma_0^i = N_S\sigma_0$, we notice (7.10) is identical to the equation describing price dynamics in the FTM model (6.8). We conclude *i)* $p^{\text{eq}} = p_0 = 1$ is the only price equilibrium of the hFTM model, and *ii)* p^{eq} is globally stable (see **??** for a proof).

At market equilibrium, we necessarily have $\frac{dS_i}{dt} = \frac{dD_i}{dt} = 0$ in system (7.4), which imply $\sigma_i^{\text{eq}} = \Phi_i^{\text{out};\,\text{eq}}$ and $\delta_i^{\text{eq}} = \Phi_i^{\text{in};\,\text{eq}}$. Since $p^{\text{eq}} = 1$ is the only price equilibrium, we have $\sigma_i^{\text{eq}} = \sigma_0^i$ et $\delta_i^{\text{eq}} = \delta_0^i$. We conclude :

$$\sigma_0^i = \Phi_i^{\text{out};\,\text{eq}} \quad \text{and} \quad \delta_0^i = \Phi_i^{\text{in};\,\text{eq}} \tag{7.11}$$

where $\Phi_i^{\text{out};\,\text{eq}}$ and $\Phi_i^{\text{in};\,\text{eq}}$ can be estimated from trade data as the number of goods sold and bought by i during a time period and divided by the same time period. Following the ideas developed in chapter 5, we can hence characterise an agent i by its flow polarity and its flow share — as calculated from (7.11) — and consequently define market categories for the hFTM model (see Fig. F.2 for an example).

Provided we specify f_{ij} and g_{ij}, we can finally estimate microscopic friction κ_{ij} because elementary transactions θ_{ij} are observable. Choices of particular partitioning f_{ij} and g_{ij} are discussed in subsection 7.1.4.

Note that while the equilibrium price $p^{\text{eq}} = 1$ exists, is unique and globally stable, equilibria in S_i^{eq} and D_i^{eq} do not always exist (see subsection F.3.2 for a proof). As in the FTM model, equilibria in trade flows at the agent-level, provided they exist, are necessarily unique and converge by construction towards observed trade flow (7.11). However, we cannot prove the existence of $\Phi_i^{\text{out};\,\text{eq}}$ and $\Phi_i^{\text{in};\,\text{eq}}$ in the general case. We hence largely rely on numerical experiments.

Dynamical constraints on the batch size q_{ij}

We recall q_{ij} is defined as $q_{ij}(t) = \min(S_i(t), D_j(t))$. We denote by N the number of agents in the market.

Proposition 1. *Let $(q_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}$ represent real numbers. The following propositions are equivalent :*

(i). There exists vectors $S \in \mathbb{R}^N$ and $D \in \mathbb{R}^N$ respecting $\forall i, \forall j, q_{ij} = \min(S_i, D_j)$.

(ii). $\forall i, \forall j, q_{ij} = \min\left(\max_{1\le k\le N} q_{ik}, \max_{1\le k\le N} q_{kj}\right)$.

(iii). $\forall i, \forall j, \left(q_{ij} = \max_{1\le k\le N} q_{ik} \text{ or } q_{ij} = \max_{1\le k\le N} q_{kj}\right)$.

Démonstration. $(ii) \Rightarrow (iii)$: trivial.

$(iii) \Rightarrow (ii)$: Since $\max_{1\le k\le N} q_{ik}$ and $\max_{1\le k\le N} q_{kj}$ are larger than q_{ij}, if either of the two quantities worths q_{ij}, their min also worth q_{ij}.

$(ii) \Rightarrow (i)$: it is sufficient to set $S_i = \max_{1\le k\le N} q_{ik}$ and $D_j = \max_{1\le k\le N} q_{kj}$.

$(i) \Rightarrow (ii)$: $\forall i, \forall j, S_i \ge \min(S_i, D_j) = q_{ij}$ so for all i, $S_i \ge \max_{1\le k\le N} q_{ik}$. Along the same line, for all j, $D_j \ge \max_{1\le k\le N} q_{kj}$. We deduce :

$$q_{ij} = \min(S_i, D_j) \ge \min\left(\max_{1\le k\le N} q_{ik}, \max_{1\le k\le N} q_{kj}\right)$$

Since $\max_{1\le k\le N} q_{ik} \ge q_{ij}$ and $\max_{1\le k\le N} q_{kj} \ge q_{ij}$, the reverse inequality also stands :

$$\min\left(\max_{1\le k\le N} q_{ik}, \max_{1\le k\le N} q_{kj}\right) \ge q_{ij}$$

, which shows (ii). □

The conditions on q_{ij} implied by proposition 1 are restrictive. In particular, the matrix q made of elements q_{ij} contains one line or column constant and equal to $\min_{i,j} q_{ij}$. The observed q_{ij} are hence unlikely to correspond to the simulated q_{ij}. This limitation in the hFTM arise from the definition of q_{ij}. To overcome this issue, we present an alternative definition of the batch size in subsection 7.1.4.

7.1.3 Reproducing trade dynamics with the hFTM framework

We explore whether the hFTM model has the potential to adequately describe trade data. First, we describe the particular trade data employed in this chapter. Second, we explain how the hFTM model can be calibrated from trade data. Third, we explore simulated trade dynamics for contrasted scenarios in initial conditions of state variables and assess to which extent the hFTM model can

Data used

To ease the numerical explorations, we focus in this chapter on a small geographical subset of the French cattle-exchange data, namely the Finistère, a French Département. Finistère is part of Brittany, one of the most important French area for production and exchange of dairy cattle (see also appendix D). We focus on internal movements to the Finistère from year 2005 to year 2009 and neglect animal movements to slaughterhouses, which act as epidemiological dead-ends. Over years 2005 to 2009, the cattle-exchange network in the Finistère represents :

- 5798 agents, with 1723 strict suppliers and 398 strict demanders.
- 456420 animal movements ;
- 79751 transactions ;
- 55470 directed links ;
- 3252 agents in the largest strongly connected component (about 56 % of total agents).

Importantly, we assume trade data used for model parametrisation is at equilibrium, which appears reasonable for livestock-exchange in France (chapter 5). Values estimated from data are denoted with a hat sign ($\hat{}$) in superscript.

Model calibration to fit observed trade dynamics

The hFTM model is hence calibrated as closely as possible to observed trade conditionally on agents' observed neighbourhoods. We hence assume agents keep the same neighbours as observed from year 2005 to year 2009 in our livestock data.

We assume the partitioning f_{ij} and g_{ij} are chosen proportionally to the trade flow $\hat{\phi}_{ij}$ observed in

data, and are denoted with a star in superscript :

$$f_{ij}^{*} = \frac{\hat{\phi}_{ij}}{\sum_k \hat{\phi}_{ik}} = \frac{\hat{\phi}_{ij}}{\hat{\phi}_i^{\text{out}}} \ , \qquad g_{ij}^{*} = \frac{\hat{\phi}_{ij}}{\sum_k \hat{\phi}_{kj}} = \frac{\hat{\phi}_{ij}}{\hat{\phi}_j^{\text{in}}} \ . \tag{7.12}$$

We recall $\sigma_i = \hat{\phi}_i^{\text{out}}$ and $\delta_i = \hat{\phi}_i^{\text{in}}$ at equilibrium, and f_{ij} and g_{ij} are originally defined as the shares of σ_i and δ_i allocated by i to j and by j to i respectively. It follows f_{ij}^{*} and g_{ij}^{*} correspond to the values of f_{ij} and g_{ij} for which agents partition their productions so as to conclude optimal bargains, i.e. without losses of trade flow.

The elementary transaction rate θ_{ij}^{eq} (7.2) is hence given by :

$$\theta_{ij}^{\text{eq}} = \frac{1}{\kappa_{ij}} \min\left(f_{ij}^{*}\hat{\phi}_i^{\text{out}}, g_{ij}^{*}\hat{\phi}_j^{\text{in}}\right) = \frac{\hat{\phi}_{ij}}{\kappa_{ij}} \tag{7.13}$$

To have $\theta_{ij}^{\text{eq}} = \hat{\theta}_{ij}$ as observed in data, we need to define κ_{ij} as :

$$\kappa_{ij} = \hat{q}_{ij} = \frac{\hat{\phi}_{ij}}{\hat{\theta}_{ij}} \ , \tag{7.14}$$

which is strictly analogous to the value of friction κ in the macroscopic FTM model at equilibrium (6.7).

Dynamical exploration for various initial conditions

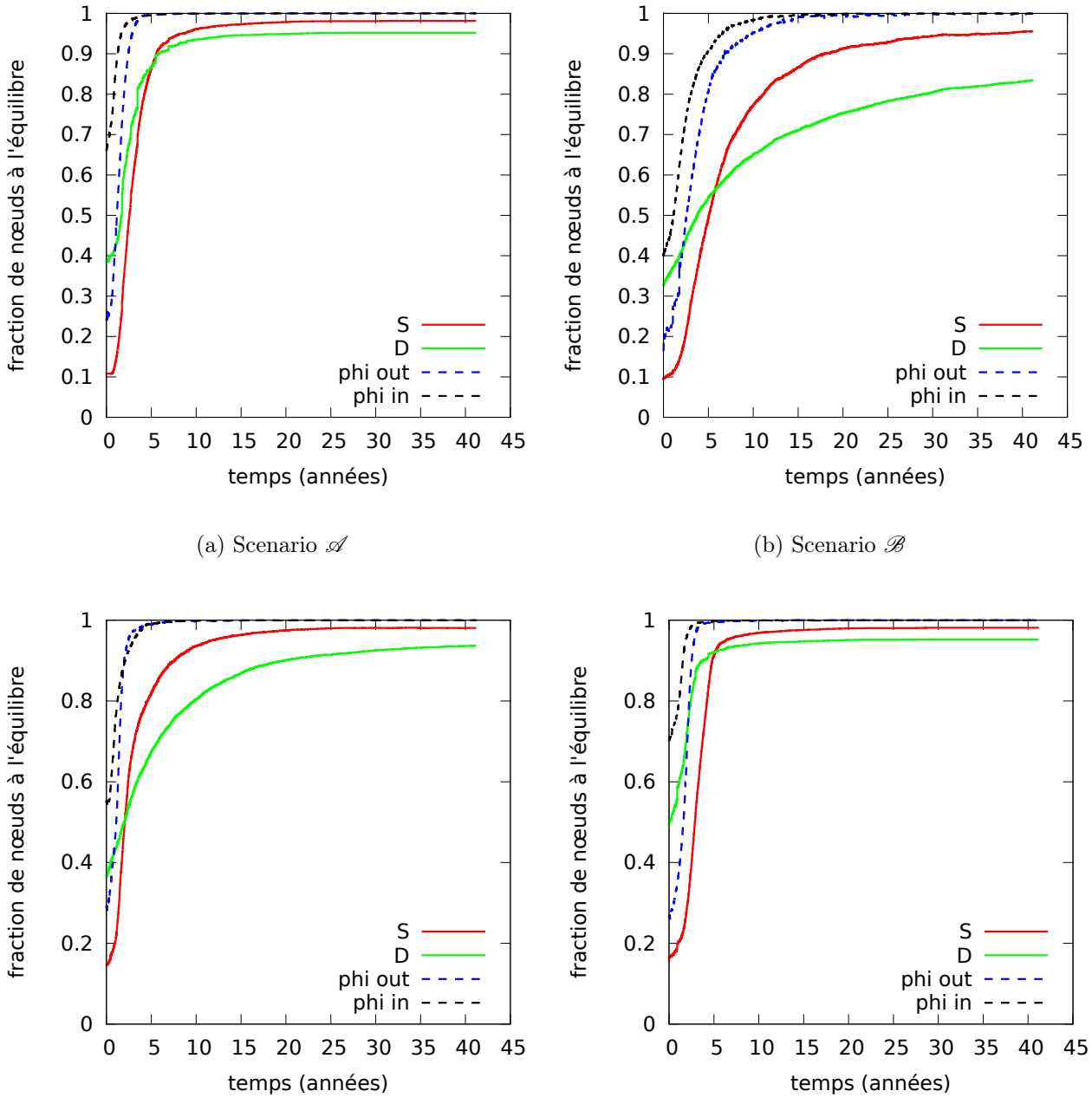
We now explore the hFTM dynamics with κ_{ij} given by (7.14), $f_{ij} = f_{ij}^{*}$ and $g_{ij} = g_{ij}^{*}$ (7.12).

Since price $p^{*} = 1$ is globally stable, we place ourselves at $p^{*} = 1$ without lack of generality and we consider four scenarios for initial stocks $S_i(0)$ and $D_i(0)$:

- Scenario $\mathscr{A}$: $S_i(0)$ and $D_i(0)$ are set to zero ;
- Scenario $\mathscr{B}$: $S_i(0)$ and $D_i(0)$ are equal and set to the average annual number of animals present within holdings i as calculated from data ;
- Scenario $\mathscr{C}$: $S_i(0)$ and $D_i(0)$ are equal to the cumulative annual total trade flow : $S_i(0) = D_i(0) = (1 \text{ year}) \times (\hat{\phi}_i^{\text{in}} + \hat{\phi}_i^{\text{out}})$;
- Scenario $\mathscr{D}$: $S_i(0)$ and $D_i(0)$ are asymmetric and set to $S_i(0) = (1 \text{ year}) \times \hat{\phi}_i^{\text{out}}$ and $D_i(0) = (1 \text{ year}) \times \hat{\phi}_i^{\text{in}}$.

For any initial stock explored, the hFTM model converges in trade flows but not always in stocks of supply and demand (Fig. 7.1). The speed of equilibration can considerably vary among scenarios. For instance, equilibration speed in scenario $\mathscr{B}$ is substantially lower than in scenario $\mathscr{A}$. Ranges of equilibria in S_i et D_i appear very narrow, indicating a poor dependency of the system final state on

initial values of state variables (Fig. 7.2). Overall, the hFTM model exhibits a behavior qualitatively identical to its macroscopic counterpart.



(a) Scenario $\mathscr{A}$

(b) Scenario $\mathscr{B}$

(c) Scenario $\mathscr{C}$

(d) Scenario $\mathscr{D}$

FIGURE 7.1 – Market dynamics for contrasted scenarios on initial stocks of supply and demand. The x-axis represents time expressed in years. The y-axis represents the fraction of agents i which reach an equilibrium in supply (S_i ; curve "S"), demand (D_i ; curve "D"), out-trade flow (ϕ_i^{out} ; curve "phi out") and in-flow (ϕ_i^{in} ; curve "phi in"). Further details are available in appendix F.

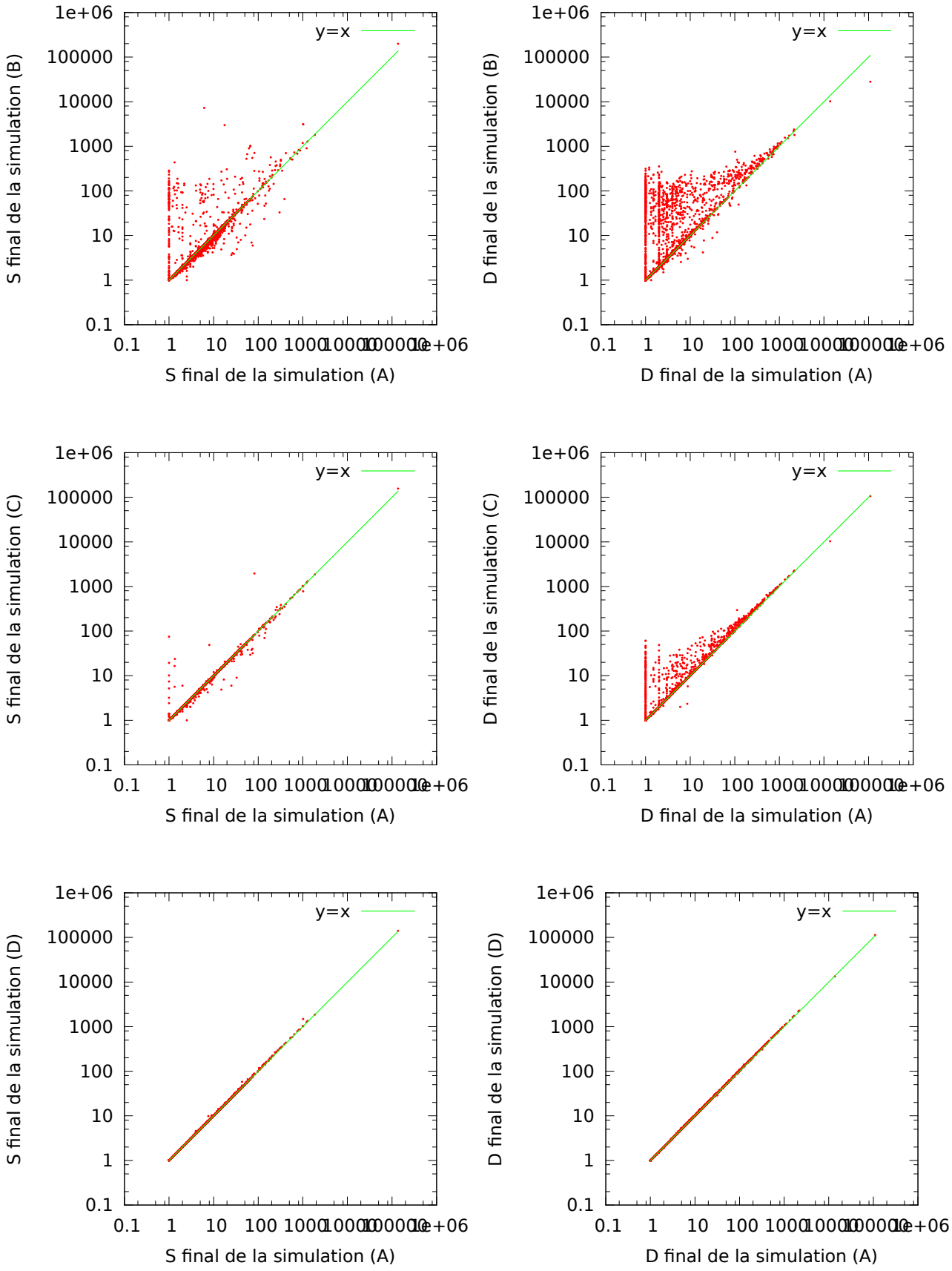
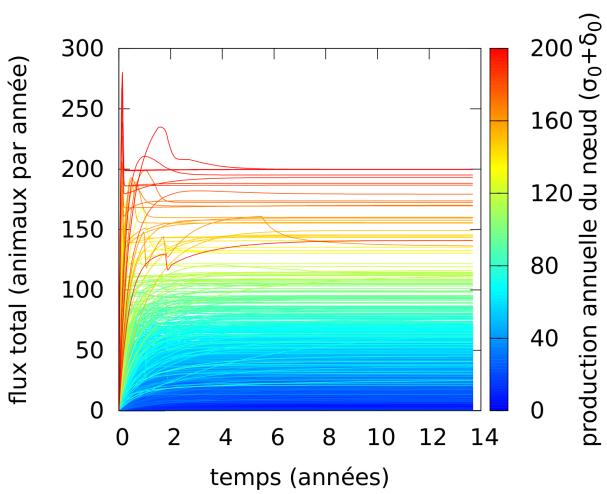


FIGURE 7.2 – **Comparison of equilibrium values of supply and demand for the four scenarios tested**. The x-axis (y-axis) presents the equilibrium values of S_i and D_i obtained under scenario $\mathscr{A}$ (under scenarios $\mathscr{B}$, $\mathscr{C}$ and $\mathscr{D}$). Further details are available in appendix F.

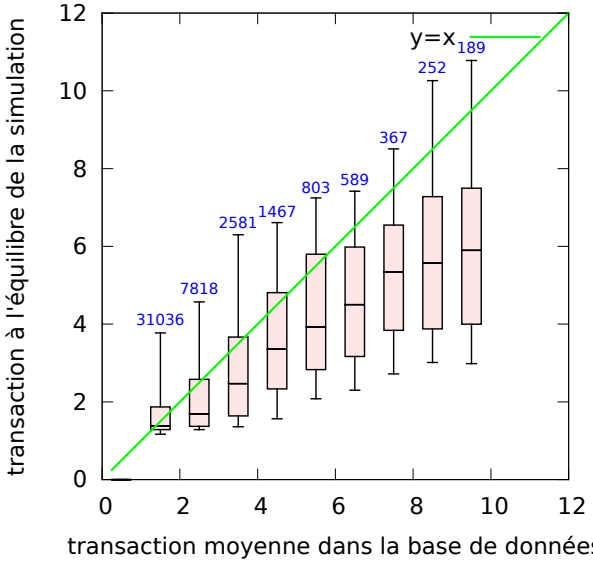
Comparison of typical hFTM outcomes with data

Since all scenarios in initial conditions exhibit a close behaviour, we focus our forthcoming analyses on scenario $\mathscr{A}$. Here, we assess to which extent the hFTM model calibrated from trade data can reproduce observed trade at equilibrium.

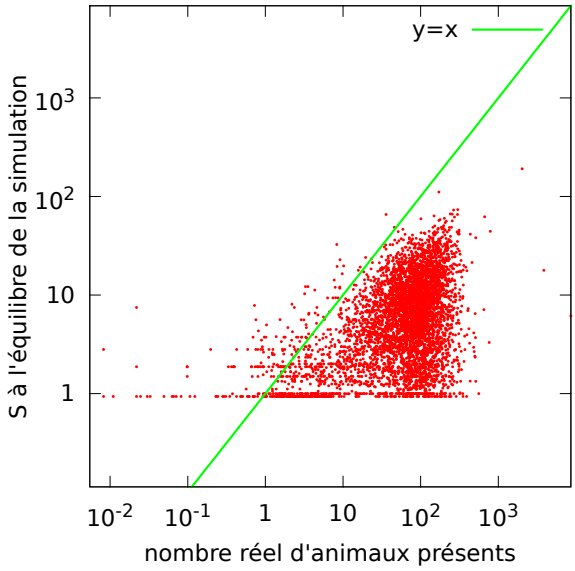
The hFTM model appears to perfectly match trade flow data at the agent-level as predicted analytically (Fig 7.3a). However, at the link-level, the simulated batch size at equilibrium q_{ij}^{eq} overestimates (underestimates) the real $\hat{q}_{ij}$ for low values (medium to large values) (Fig 7.3b). This realism issue probably arises from the definition of q_{ij} which leads to strong constraints in the values that q_{ij} can possibly take (see proposition 1). Supply and demand stocks at equilibrium are almost always lower than the actual number of animals present within holdings (Fig. 7.3c and 7.3d). This feature of our model is realistic, because an agent cannot sells or buys more goods than he/she owns or can accommodate.



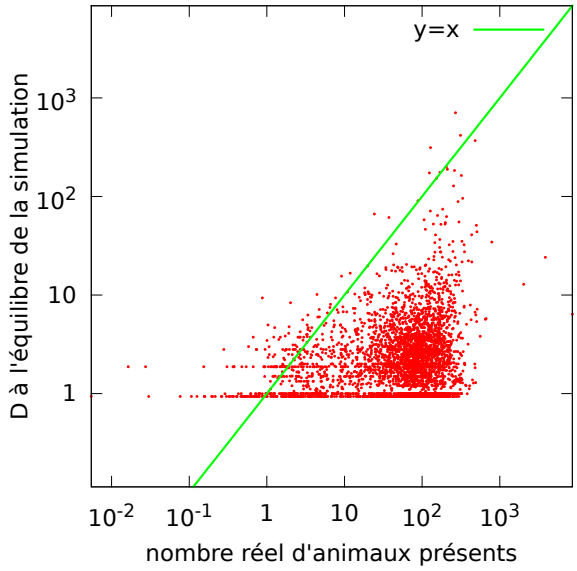
(a) Total trade flow per agent i ($\phi_i^{\text{out}} + \phi_i^{\text{in}}$) as function of time. Individual trajectories are colored from blue to red as function of the reference production rates in supply and demand ($\sigma_0^i + \delta_0^i$). We represent all trajectories for agents trading less than 200 animals per year (43 agents neglected).



(b) q_{ij}^{eq} (simulated) as function of $\hat{q}_{ij}$ (data). Values of $\hat{q}_{ij}$ were grouped by intervals of size 1 (values larger than 10 were neglected, which represent about 2 % of total directed links). The number of data per group are indicated in blue above each box.



(c) Supply stock at equilibrium for agent i (S_i^{eq}) as function of the observed annual average number of animal present within i



(d) Demand stock at equilibrium for agent i (D_i^{eq}) as function of the observed annual average number of animal present within i

FIGURE 7.3 – **Typical outcomes of scenario $\mathscr{A}$**. Evolution of total trade flow per agent i ($\phi_i^{\text{out}} + \phi_i^{\text{in}}$) colored from blue to red as function of ($\sigma_0^i + \delta_0^i$) (7.3a), comparison of q_{ij}^{eq} with $\hat{q}_{ij}$(7.3b) and comparison of agents' supply and demand stocks at equilibrium as function of observed annual average number of animal present within i (7.3c and 7.3d).

7.1.4 Explaining observed trade dynamics

We attempt to explain observed trade dynamics by comparing contrasted models for the batch size q_{ij} and the partitioning functions f_{ij} and g_{ij}. Importantly, those features of the hFTM model implies trade contacts emerges from socio-economic mechanisms describing how agents' prefer to exchange.

Potential drivers underlying batch size q_{ij}

The current definition of q_{ij}, i.e. $q_{ij}(t) = \min(S_i(t), D_j(t))$, assumes seller i and demander j exchange as much as they possible can whenever they transact. This appears a reasonable assumption, because residual stocks, which are probably costly to maintain for sellers and generate frustrations in demanders, are as low as they can be (chapter 6). However, comparisons with data do not seem to fully support this definition of q_{ij} (subsection 7.1.3).

In practice, we attempt to address this discrepancy between data and simulations by defining an alternative batch size q_{ij}^{new}, given by :

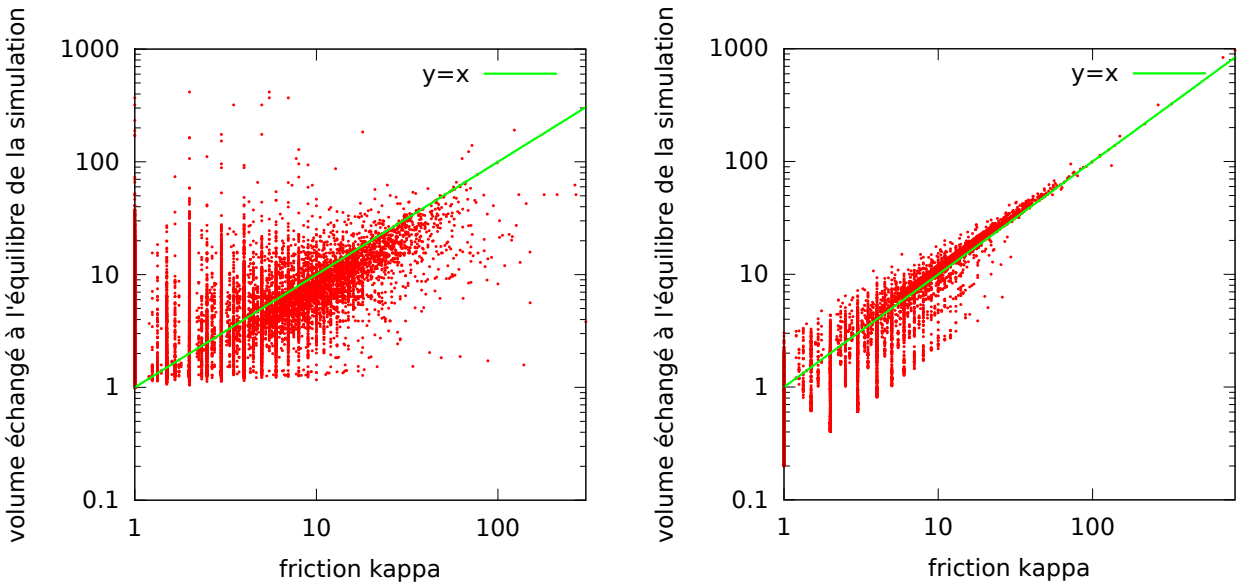
$$q_{ij}^{\text{new}}(t) = \min(f_{ij}S_i(t), g_{ij}D_j(t)) \ , \tag{7.15}$$

Since $f_{ij} \leq 1$ and $g_{ij} \leq 1$, we necessarily have $q_{ij}^{\text{new}} \leq q_{ij}$, which corresponds to a new model where agents do not exchange as much as they can at each transaction. Rather, they are assumed to exchange proportionally to their partitioning preferences encoded in f_{ij} and g_{ij}. This model is also plausible, because agents may provision stocks proportionally to the production share they reserve for their trading partners.

We find model q_{ij}^{new} is substantially closer to data than q_{ij} (Fig 7.4 and Table 7.1), which suggests agents exchange stocks proportionally to the production share they reserve for each of their partners.

TABLE 7.1 – Comparisons of simulations to data assuming the reference partitioning f_{ij}^* and g_{ij}^*. We carry out linear regressions of q_{ij} and q_{ij}^{new} at equilibrium (y) as function of $\hat{q}_{ij}$ (x), where a and b are constants. The values represent the fraction of variance explained (r^2). Note that we find $a \approx 1$ for the q_{ij}^{new} model.

	y=ax+b	y=ax
q_{ij} model	0.540	0.564
q_{ij}^{new} model	0.971	0.972



(a) Original model q_{ij}

(b) New model q_{ij}^{new}

FIGURE 7.4 – Equilibrium values of q_{ij} and q_{ij}^{new} as function of observed values $\hat{q}_{ij}$. Only values for which S_i and D_j are equilibrated are represented. Initial conditions correspond to scenario $\mathscr{A}$ and we take $\kappa_{ij} = \hat{q}_{ij}$.

Potential drivers underlying partitioning f_{ij} and g_{ij}

So far, we did not try to explain the mechanisms underlying f_{ij} and g_{ij}. Here, we attempt to explain the factors leading to partitioning $f_{ij} = f^*_{ij}$ and $g_{ij} = g^*_{ij}$ conditionally on agents' observed neighbourhoods. In other words, we assume agents keep the same neighbours as observed from year 2005 to year 2009 in our livestock data. This simplifying assumption is plausible for livestock markets, which exhibit high loyalties in their trading relationships [Valdano *et al.*, 2014].

Understanding partitioning preferences with an adjusted partitioning model. We assume partitioning functions $f_{ij}(\alpha_i)$ and $g_{ij}(\beta_j)$ now depends on two parameters α_i and β_j according to :

$$\begin{aligned} f_{ij}(\alpha_i) &= \frac{\delta_j^{\alpha_i}}{\sum\limits_{k\in\hat{\Gamma}_i^{\text{out}}} \delta_k^{\alpha_i}} \ , \\ g_{ij}(\beta_j) &= \frac{\sigma_i^{\beta_j}}{\sum\limits_{k\in\hat{\Gamma}_j^{\text{in}}} \sigma_k^{\beta_j}} \ , \end{aligned} \tag{7.16}$$

where $\hat{\Gamma}_i^{\text{out}}$ and $\hat{\Gamma}_j^{\text{in}}$ are the out- and in- trading partners of i and j as observed in data. Notice α_i (β_j) reflects agent's i (agent's j) preferences in interacting preferentially with its demanders k (its suppliers k), depending on their productions of demand δ_k (productions of supply σ_k). We recall that at equilibrium, $\sigma_k = \hat{\phi}_k^{\text{out}}$ and $\delta_k = \hat{\phi}_k^{\text{in}}$.

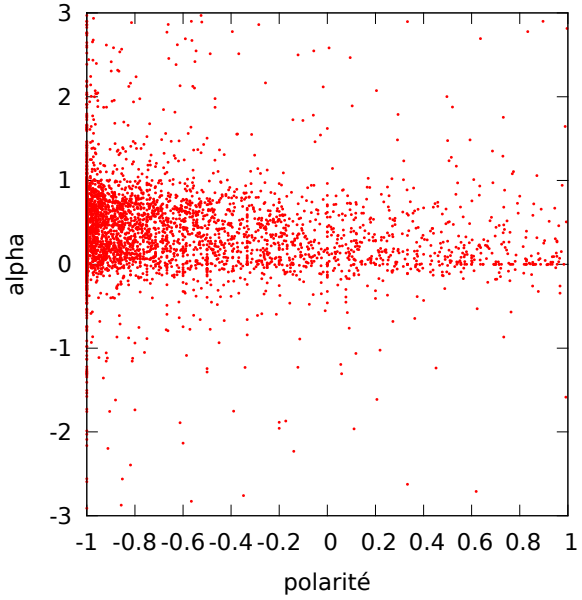
We adjust parameters $\alpha_i = \hat{\alpha}_i$ and $\beta_j = \hat{\beta}_j$ to best reflect data by minimising the residual sum of squares :

$$\begin{aligned} \hat{\alpha}_i &= \underset{\alpha_i}{\text{argmin}} \sum_{j\in\hat{\Gamma}_i^{\text{out}}} \left(f_{ij}(\alpha_i) - f^*_{ij}\right)^2 \ , \\ \hat{\beta}_j &= \underset{\beta_j}{\text{argmin}} \sum_{i\in\hat{\Gamma}_j^{\text{in}}} \left(g_{ij}(\beta_j) - g^*_{ij}\right)^2 \ . \end{aligned} \tag{7.17}$$

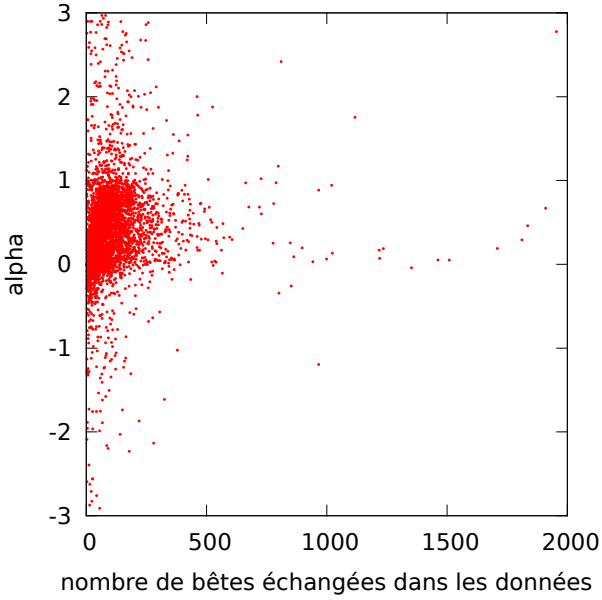
We can then compare the estimated $\hat{\alpha}$ and $\hat{\beta}$ with agents' flow polarity and flow share (Fig. 7.5). We recall flow polarity quantifies agents' positions along the supply chain while flow share quantifies their relative market importance (chapter 5). First of all, we remark most values of $\hat{\alpha}$ and $\hat{\beta}$ are positive, which indicates agents trade preferentially with agents with large flow share. Second, we notice that while independent from flow polarity, $\hat{\alpha}$ and $\hat{\beta}$ tend to increase with flow share, which suggest agents with the largest flow shares preferentially trade with agents with the largest flow shares. Third, we notice an asymmetry in the behaviour $\hat{\alpha}$ and $\hat{\beta}$ with respect to small values of flow share : more values of $\hat{\beta}$ appear to be negative compared with $\hat{\alpha}$. We conclude a non-negligible fraction of small demanders prefer interacting with small suppliers.

Note that we do not have yet produced the data to compare $f_{ij}(\hat{\alpha}_i)$ and $g_{ij}(\hat{\beta}_j)$ with f^*_{ij} and g^*_{ij}. We also need to simulate trade dynamics with $f_{ij}(\hat{\alpha}_i)$ and $g_{ij}(\hat{\beta}_j)$ under the q_{ij}^{new} model. Nevertheless,

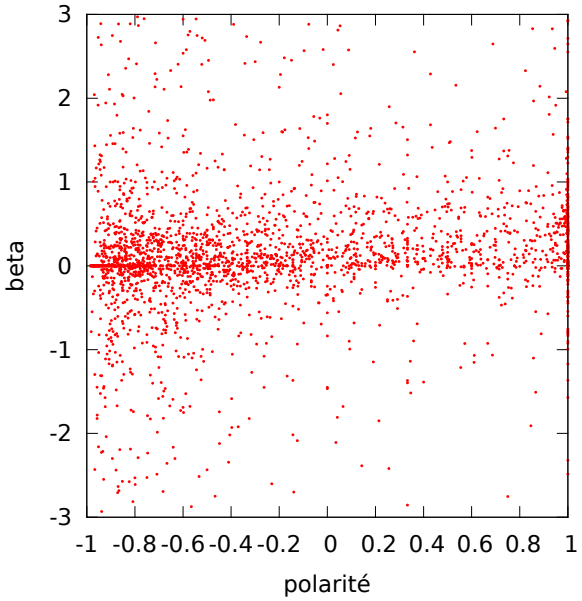
preliminary data based on the original q_{ij} model suggests (see subsection F.6.1 in appendix) : *a*) approach (7.17) has the potential to closely match observed data in partitioning and trade flow at the partnership-level, *b*) the quality of the matching to data can be drastically altered by changing the values of α_i and β_j.



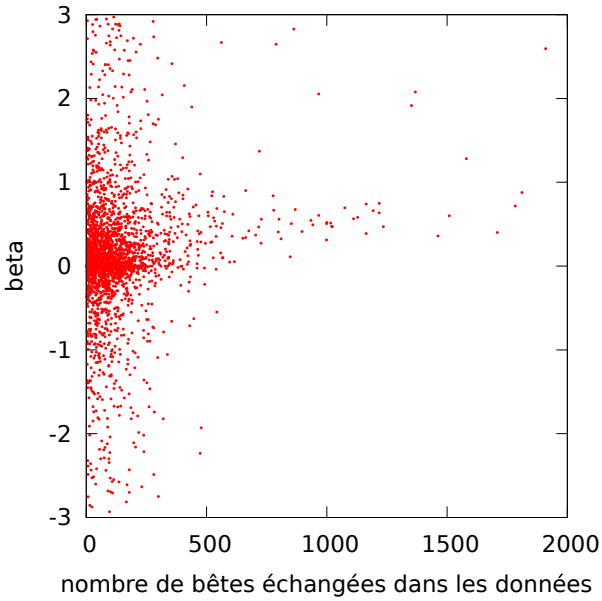
(a) Partitioning preference α_i as function of flow polarity fp_i



(b) Partitioning preference α_i as function of flow share multiplied by total agents $\text{fs}_i \times N$



(c) Partitioning preference β_j as function of flow polarity fp_j



(d) Partitioning preference β_j as function of flow share multiplied by total agents $\text{fs}_j \times N$

FIGURE 7.5 – Agents' partitioning preferences as function of flow polarity and flow share.

Explaining partitioning preferences with an adaptive partitioning model. Inspired by the maximisation partitioning model proposed by [Moslonka-Lefebvre *et al.*, 2012a], we propose to explain the observed partitioning f_{ij}^{*} and g_{ij}^{*} by an adaptive partitioning with memory. We assume agents wish to maximise their exchanges by adjusting their partitioning f_{ij}^{adapt} and g_{ij}^{adapt} every T time units (here every week). More precisely, we assume agent update their partitioning at time $t+T$ in proportion to what they received from their trading partners at time t :

$$\begin{aligned} f_{ij}^{[t+T]} &= \frac{g_{ij}^{[t]}\delta_j}{\sum_k g_{ik}^{[t]}\delta_k} \ , \\ g_{ij}^{[t+T]} &= \frac{f_{ij}^{[t]}\sigma_i}{\sum_k f_{kj}^{[t]}\sigma_k} \ . \end{aligned} \tag{7.18}$$

To close (7.18), we assume agents initially exhibit uniform preferences when interacting with their trading partners :

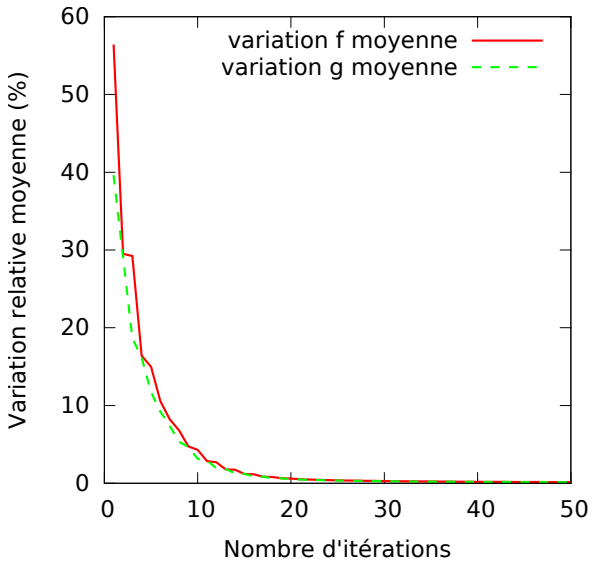
$$\begin{aligned} f_{ij}^{[0]} &= \frac{1}{\hat{N}_i^{\text{out-partners}}} \ , \\ g_{ij}^{[0]} &= \frac{1}{\hat{N}_j^{\text{in-partners}}} \ , \end{aligned} \tag{7.19}$$

where $\hat{N}_i^{\text{out-partners}}$ and $\hat{N}_j^{\text{in-partners}}$ represent the total number of trading partners of agents i and j as observed from data.

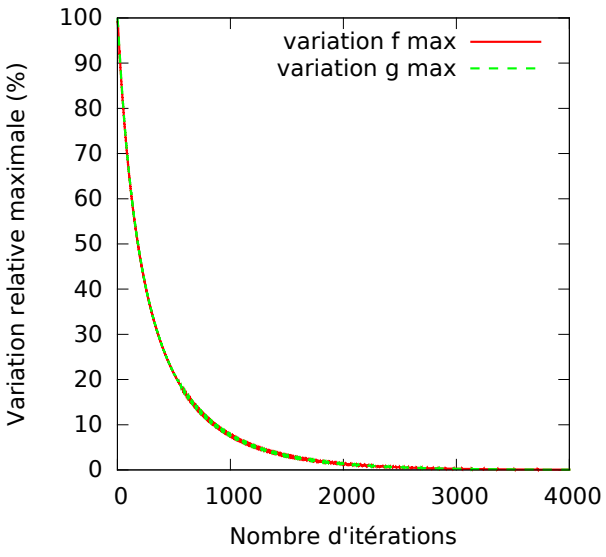
The adaptive partitioning process (7.18–7.19) converges for all pairs of trading agents (Fig. 7.6). We define the adaptive partitioning model by $f_{ij}^{\text{adapt}} = f_{ij}^{[3000]}$ and $g_{ij}^{\text{adapt}} = g_{ij}^{[3000]}$, for which all partitioning functions have reached an equilibrium. Though the fit to observed trade data is far from perfect, the current adaptive partitioning process yields promising results (table 7.2). Our preliminary findings indicate agents may indeed preferentially interact with partners who preferentially interact with them in return. Note that we essentially assume agents have a short memory and only care about the recent past (here the week), which may explain why the fit to data is not excellent. As future work, we could attempt to improve the matching to data by assuming agents adjust their current partitioning based on multiple points of time in the past, with some kind of discounting function to adjust from data. In economics, discounting quantifies how agents values past gains in comparison with present gains ([see e.g. Rowthorn *et al.*, 2009, for more details on discounting]).

TABLE 7.2 – Comparisons of simulations to data assuming the adaptive partitioning f_{ij}^{adapt} and g_{ij}^{adapt}. We carry out linear regressions under the adaptive partitioning model of q_{ij} and q_{ij}^{new} at equilibrium (y) as function of $\hat{q}_{ij}$ (x), where a and b are constants. The values represent the fraction of variance explained (r^2).

	y=ax+b	y=ax
q_{ij} model	0.515	0.536
q_{ij}^{new} model	0.446	0.465



(a) mean relative change in coefficients $f_{ij}^{[t]}$ and $g_{ij}^{[t]}$ as function of time t.



(b) maximal relative change in coefficients $f_{ij}^{[t]}$ and $g_{ij}^{[t]}$ as function of time t.

FIGURE 7.6 – Convergence of the adaptive partitioning process.

7.2 Preventing epidemics by manipulating elementary friction

At the macroscopic-level, friction κ appears to exert a leading influence on joint trade and infection dynamics. In particular, κ can limit and even prevent infection transmission in our ME model, which, among others, assumes homogeneous friction (chapter 6). At the scale of agents, microscopic friction κ_{ij} can be quantified as $\hat{q}_{ij} \geq 1$, the observed batch size transacted from seller i to demander j (section 7.1). In livestock-exchange markets, $\hat{q}_{ij}$ appears highly heterogeneous (Fig. **??**CD), which suggests elementary friction is also heterogeneous among pairs of trading agents. Additionally, when trade flow and the number of business partners are positively correlated at agent level, as e.g. observed in trade of cattle and swine (Fig. D.2AB), we expect heterogeneous contact structures to boost epidemic development in comparison with homogeneous settings [Kamp *et al.*, 2013]. Heterogeneity in agents' features and friction question whether epidemic can also be prevented in actual markets.

Here, we study the influence of heterogeneity in microscopic friction on economic-epidemiological dynamics, and assess to which extent it can also limit disease spread in a mechanistic model of heterogeneous markets. We proceed in three steps. First, we formulate a heterogeneous market-epidemiological (hME) model in a stochastic framework. Second, we conduce percolation experiments on the dynamical trade network simulated using only the economic component of the hME model to assess whether markets with diverse friction can sustain epidemics. Third, we confirm the insights derived from percolation studies by exploring dynamical economic-epidemiological trajectories for contrasted

epidemiological scenarios. Overall, our preliminary findings suggest increasing elementary friction can also prevent epidemics in heterogenous markets.

7.2.1 The heterogeneous market-epidemiological (hME) model

The mutual influence of epidemics and trade dynamics at the microeconomic scale remain largely unknown (chapter 6). Building upon the data-motivated approach adopted in chapter 5 and the proof-of-concept approach exposed in chapter 6, we develop a network-based model of joint market and epidemic dynamics (subsection 2.3.1). We adopt a stochastic standpoint, which is necessary to conduce percolation experiments. We assume the production of supply and demand occur continuously, but describe transactions, infections, removals and re-introduction events as Markovian jump processes.

The market component of the hME model remains very close to the hFTM model. Each agent i remains characterised by its production of supply and demand stocks σ_i and δ_i, its partitioning preferences f_{ij} and g_{ij} and its supply and demand stocks S_i and D_i.

The epidemiological component of the hME model is derived from the stochastic SIRS model. Each agent is in susceptible (X), infected (Y) or removed (Z) state. At a given time point, the current state of an agent is described by the bi-dimensional variable (X_i, Y_i), where X_i and Y_i are binary variables (e.g. $X_i = 1$ and $Y_i = 0$ if i is susceptible). Note that $X_i + Y_i + Z_i = 1$, so we do need to track Z_i.

The coupling processes of the hME model are as follows. We assume an agent is active in trade if it belongs to classes X or Y. Else it is inactive. Removals occurs at rate γ and re-introductions at rate ν. Importantly, infections occur conditionally on transactions. Provided both seller i and demander j are active, elementary transactions occur at rate θ_{ij} :

$$\theta_{ij}(t) = \frac{1}{\kappa_{ij}} \min\left(f_{ij}\sigma_i, g_{ij}\delta_j\right). \tag{7.20}$$

Agents are not passive. Rather, they decide to interact for various motivations, and can modulate their interactions in response to perceived change in their environment, a set of mechanisms which may eventually generate complex feedbacks between exchange and epidemic dynamics. For instance, we can reasonably assume agents only wish to interact with active agents, which yields :

$$f_{ij}(t) = \frac{\mathbb{1}_{j\in X\cup Y} f^*_{ij}}{\sum_k \mathbb{1}_{k\in X\cup Y} f^*_{ik}} \quad \text{and} \quad g_{ij}(t) = \frac{\mathbb{1}_{i\in X\cup Y} g^*_{ij}}{\sum_k \mathbb{1}_{k\in X\cup Y} g^*_{kj}}, \tag{7.21}$$

where f^*_{ij} and g^*_{ij} are the partitioning defined in (7.12). If transaction (i, j) occurs, we consider agents exchange the largest possible batch size q_{ij} :

$$q_{ij}(t) = \min\left(S_i(t), D_j(t)\right). \tag{7.22}$$

If seller i is infected and demander j is susceptible, the transaction (i, j) may lead to an infection event

which occurs with probability $P_{\text{tr}}(q_{ij})$ given by :

$$P_{\text{tr}}(q_{ij}) = 1 - (1-\zeta)^{q_{ij}}, \tag{7.23}$$

where ζ is the probability of infection per good exchanged. Finally, price dynamics is given by :

$$p(t) = p(0)\exp\left(\mu \sum_{k\in X\cup Y}(D_k(t) - S_k(t)) - \mu \sum_{k\in X(0)\cup Y(0)}(D_k(0) - S_k(0))\right). \tag{7.24}$$

To each agent i, we can hence associate a state space $(S_i, D_i, X_i, Y_i, Z_i) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \{0,1\}^3$ where $X_i + Y_i + Z_i = 1$. Random events driving the evolution of state variables are given by :

Event	Agents	Rates	Transitions
Transaction (potential infection)	$i \in X \cup Y$ $j \in X \cup Y$	$\theta_{ij}(t)$	$S_i \leftarrow S_i - q_{ij}$, $D_j \leftarrow D_j - q_{ij}$ and with probability $\mathbb{1}_{j\in X, i\in Y} P_{\text{tr}}(q_{ij})$: $X_j \leftarrow 0$, $Y_j \leftarrow 1$
Removal	$i \in Y$	γ	$Y_i \leftarrow 0$, $Z_i \leftarrow 1$
Re-introduction	$i \in Z$	ν	$Z_i \leftarrow 0$, $X_i \leftarrow 1$ $S_i \leftarrow 0$, $D_i \leftarrow 0$

In addition to point jump processes, stocks of supply and demand continuously grow with production σ_i and δ_i :

$$\dot{S}_i = \sigma_i \quad \text{and} \quad \dot{D}_i = \delta_i \quad \text{outside jumps.} \tag{7.25}$$

We hence define a Markov process in continuous time with space state $E = \mathbb{R}^{+N} \times \mathbb{R}^{+N} \times \{0,1\}^N \times \{0,1\}^N$ where N is the total number of agents. In practice, we implement this stochastic process with an efficient method described in subsection F.3.3.

In the following, the market component of the hME model is parameterised from the same dataset as the hFTM model (subsection 7.1.3). In absence of epidemics, the (stochastic) hME model can be mapped analytically to the (deterministic) hFTM model (see F.4.1 for a proof), which imply both market models, exhibit, on average and as a first approximation, an identical behaviour (see subsection F.4.2 for a comparative numerical exploration).

7.2.2 Preventing trade-driven epidemics by increasing friction : insights from percolation experiments

Our goal is to study the influence of trade dynamics on epidemics at the microscopic scale. As a first step to reach this objective, we resort to *percolation* experiments where we assess whether increasing elementary friction can reduce the potential epidemic risk imputable to trade without having to simulate epidemics. Originally, percolation denotes the capacity of fluids to flow through porous media and is also a well established theory in physics [Albert & Barabasi, 2002; Newman, 2003]. When applied to network epidemiology, percolation usually refers to a set of theoretical tools employed to quantify the ability of an infection process to spread through a network [Meyers *et al.*, 2006]. Infection is hence analogous to a "fluid" transmitted through infection-conducing "pores", namely nodes and links, which are respectively called sites and bounds in physics [Newman, 2003]. Generally speaking, adding nodes and/or links to an existing network increases the epidemic risk by favoring disease spread. Conversely, removing nodes and/or links from a given network, which can be interpreted as disease prevention measures, reduces the epidemic risk [Albert & Barabasi, 2002]. For a given network with the capacity to sustain a given infection, for instance with $R_0 = 10$, removing gradually nodes and/or links will eventually lead to $R_0 \leq 1$, i.e. below the epidemic threshold where the infection will never take-off in the population. In percolation studies, we can assess the effectiveness of such removals on epidemic prevention by calculating the size of the largest strongly connected component (LSCC) [Meyers *et al.*, 2006]. The size of the LSCC is a standard proxy measuring the epidemic threshold and the maximal epidemic size in directed networks (see also chapter 5). Here, we study the capacity of simulated trade networks to conduce epidemics though numerical percolation experiments where we manipulate friction κ_{ij}. We assume an absence of epidemic and that the market is equilibrated. The experimental setting is as follows. *i*) For a given matrix of friction κ_{ij}, we simulate trade with the hME model up the critical time, denoted T^*, where we reach 50 % of the size of the LSCC as observed in the dataset. By analogy with the macroscopic ME model (chapter 6), the critical quantity T^* can be interpreted as a delay in enforcement of regulation that still allows prevention of epidemics. Let 0 and $T \geq T^*$ denotes the initial and final simulation times respectively. In practice, the size of the LSCC is assessed by aggregating all transactions occurring from time 0 to time T as a static network. Provided T is large enough, the LSCC of the trade network aggregated over $[0, T]$ will be identical to the LSCC observed in the dataset. *ii*) Since the hME is stochastic, we iterate step *i*) 200 times to calculate the average T^*. *iii*) We iterate steps *i*) and *ii*) for increased levels of elementary friction κ_{ij}. We explore two contrasted models for increasing elementary friction κ_{ij} : the *dilation model*, where all κ_{ij} are multiplied by a constant ; and the *thresholding model*, where all values of κ_{ij} lower than a critical friction value $\kappa_{\min}$ are replaced with $\kappa_{\min}$. The dilation model (thresholding model) can be interpreted as the introduction of a flat tax on transports (a regulation imposing a minimal shipment size).

Whether by dilatation or by thresholding, gradually increasing friction κ_{ij} increases the critical time T^* (Fig. 7.7 [1]). This suggest increasing elementary friction can eventually prevent trade-driven epidemics.

From a more theoretical perspective, altering microscopic friction corresponds to a particular bond percolation. In fact, the probability of observing the link (i, j) from seller i to demander j over a time period $[0, T]$ is given by :

$$p_{ij}(0, T) = 1 - \exp\left(-\int_0^T \theta_{ij}(t)\mathrm{d}t\right). \tag{7.26}$$

where θ_{ij} is the transaction rate from i to j, which inversely depends on friction κ_{ij}. If, we assume the market is equilibrated from time 0, (7.26) simplifies to :

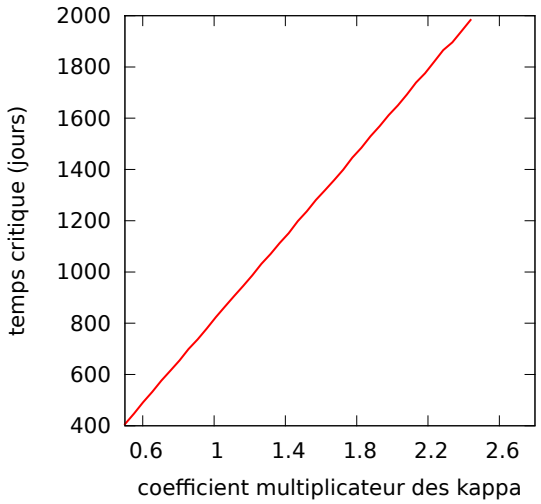
$$p_{ij}(T) = 1 - \exp\left(-T\theta_{ij}^{\mathrm{eq}}\right). \tag{7.27}$$

Let $\theta_{ij}^{\mathrm{ref}}$ and $\kappa_{ij}^{\mathrm{ref}}$ denote the reference values of the transaction rate and friction assuming market equilibrium and without manipulating friction. When changing friction, we obtain at equilibrium the relation : $\theta_{ij}^{\mathrm{eq}} = \theta_{ij}^{\mathrm{ref}} \dfrac{\kappa_{ij}^{\mathrm{ref}}}{\kappa_{ij}}$. The probability of observing the link (i, j) becomes :

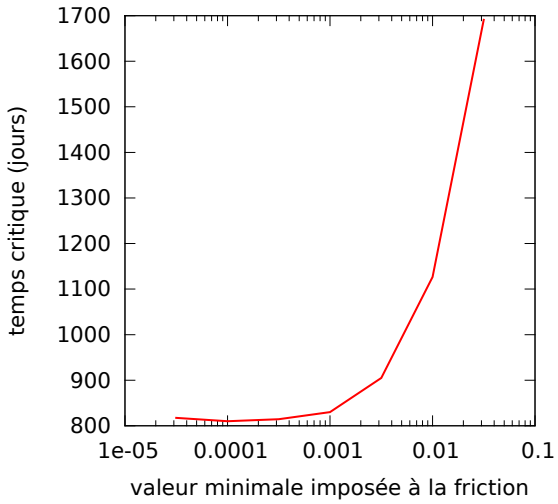
$$p_{ij}(T) = 1 - \exp\left(-T\theta_{ij}^{\mathrm{ref}} \frac{\kappa_{ij}^{\mathrm{ref}}}{\kappa_{ij}}\right). \tag{7.28}$$

To keep $p_{ij}(T)$ constant while increasing κ_{ij}, we need to increase T (7.28). This explain analytically how microscopic friction contributes to increase the critical time T^*.

1. For this particular figure, friction κ_{ij} is *not* quantified by $\hat{q}_{ij} \geq 1$ and can take values lower than 1 (see subsection F.3.4 in appendix for further details on a previous estimation of κ_{ij} we have presently discarded). As future work, we plan to repeat the percolation experiments with $\kappa_{ij} = \hat{q}_{ij}$, which will not affect our findings qualitatively, but will considerably ease the interpretation of $\kappa_{\min}$



(a) Dilatation of all κ_{ij}. The x-axis represents the constant by which we multiply the original κ_{ij}. The y-axis represents the average critical time T^* (in days) out of 200 replicate simulations.



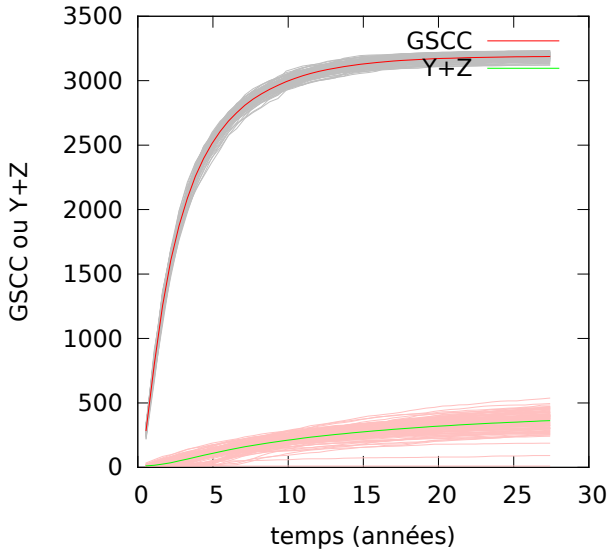
(b) Setting a minimal value for all κ_{ij}. The x-axis represents the minimal value that can be taken by κ_{ij}. The y-axis represents the average critical time T^* (in days) out of 200 replicate simulations.

FIGURE 7.7 – **Evolution of the critical time T^* when we increase friction κ_{ij}.** The critical time T^* is the time for which the temporally aggregated network as generated with the hME model reaches half of its maximal size.

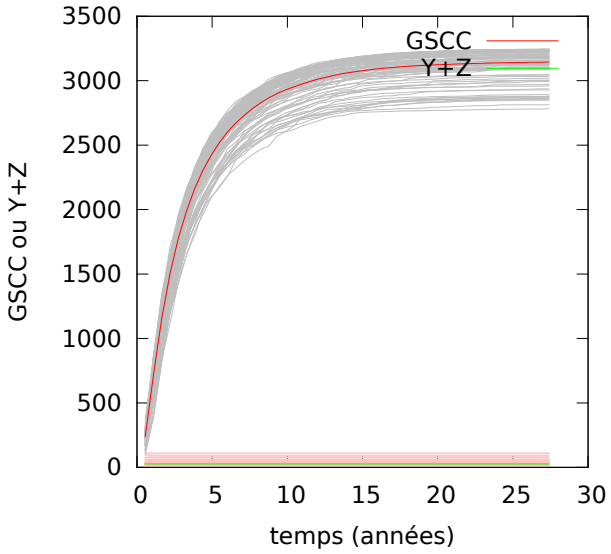
7.2.3 Epidemic trajectories with modified microscopic frictions

Here, we test whether our findings on elementary friction, as derived from percolation experiments, stand when accounting for epidemic feedbacks. With the hME, we compare the economic-epidemiological dynamics arising for various levels of friction ($\kappa_{ij} = \hat{q}_{ij}$, $\kappa_{ij} = 10\ \hat{q}_{ij}$ or $\kappa_{ij} = 0.1\ \hat{q}_{ij}$). We assume agents face contrasted epidemics characterised by differing couples $(\zeta; \frac{1}{\gamma})$, where $\zeta \in \{0.01; 1\}$ is the probability of infection per good exchanged and $\frac{1}{\gamma} \in \{1 \text{ month}; 1 \text{ year}; 10 \text{ years}\}$ is the infectious period. For each experiment corresponding to a unique triplet $(\kappa_{ij}; \zeta; \frac{1}{\gamma})$, we conduce 100 replicate simulations of the hME model where we keep track of two quantities as function of time : the average size of the largest strongly connected component (LSCC) of the temporally aggregated network ; and the average value of $(Y + Z)$. The size of the LSCC somehow represents the overall status of the market (the larger the better), while $(Y + Z)$ is the total number of individuals impacted by the disease (the lower the better). Initial supply and demand stocks are taken from a simulation of the hME model without epidemics after 5000 days. Initial price is equilibrated and we set $\mu = 10^{-6}$. Partitioning functions are given by (7.21), implying agents adjust their preferences when their partners are detected and removed. We assume removed agents are not re-introduced ($\nu = 0$).

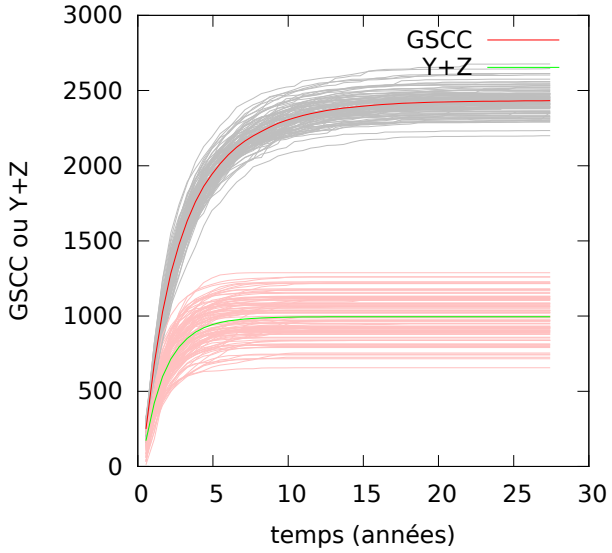
Our preliminary results suggest contagion in markets could be mitigated or even prevented by increasing elementary friction (Fig. 7.8, 7.9 and 7.10).



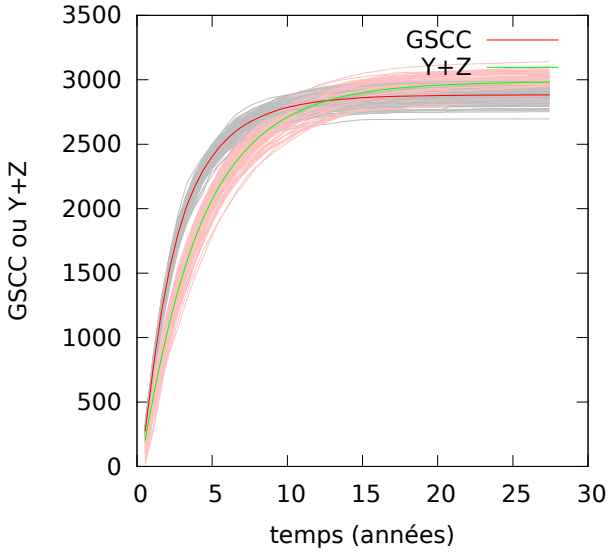
(a) $\zeta = 0.01$, $\frac{1}{\gamma} = 10$ years



(b) $\zeta = 1$, $\frac{1}{\gamma} = 1$ month

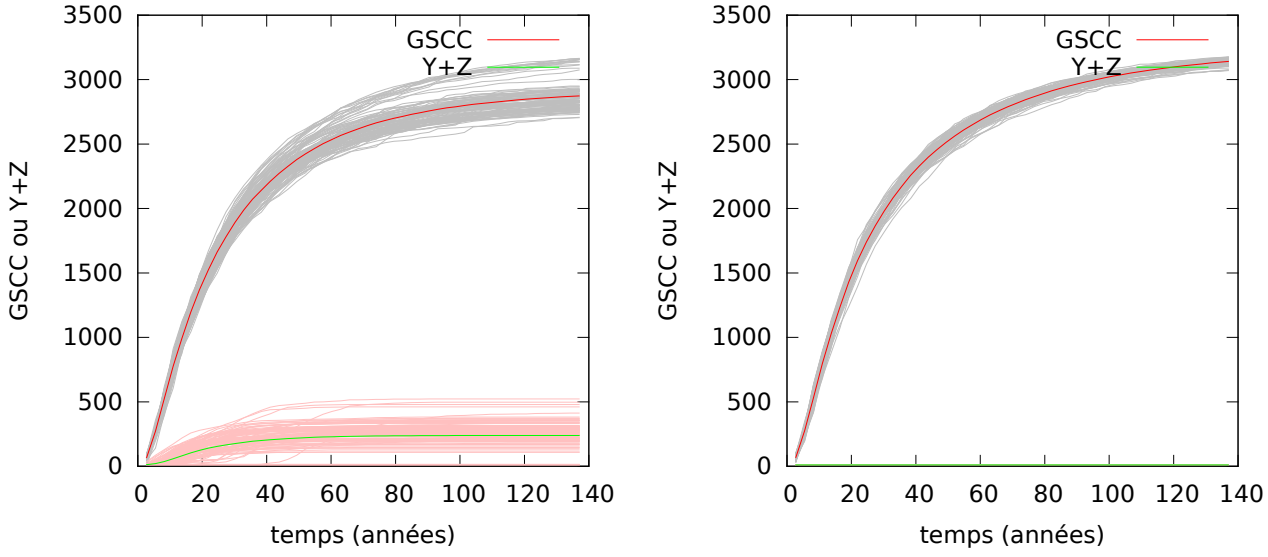


(c) $\zeta = 1$, $\frac{1}{\gamma} = 1$ year



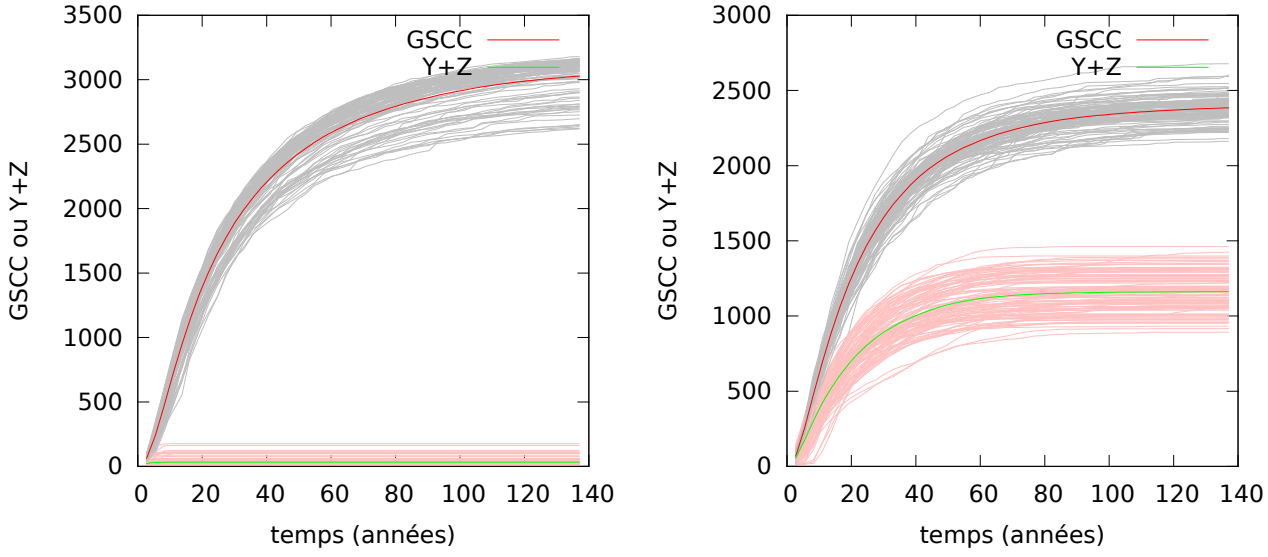
(d) $\zeta = 1$, $\frac{1}{\gamma} = 10$ years

FIGURE 7.8 – Size of the LSCC and $(Y + Z)$ as function of time for estimated friction, i.e. $\kappa_{ij} = \hat{q}_{ij}$.



(a) $\zeta = 0.01$, $\frac{1}{\gamma} = 10$ years

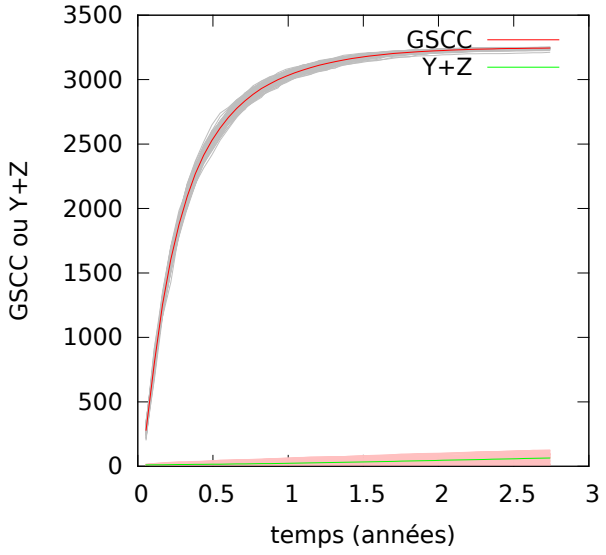
(b) $\zeta = 1$, $\frac{1}{\gamma} = 1$ month



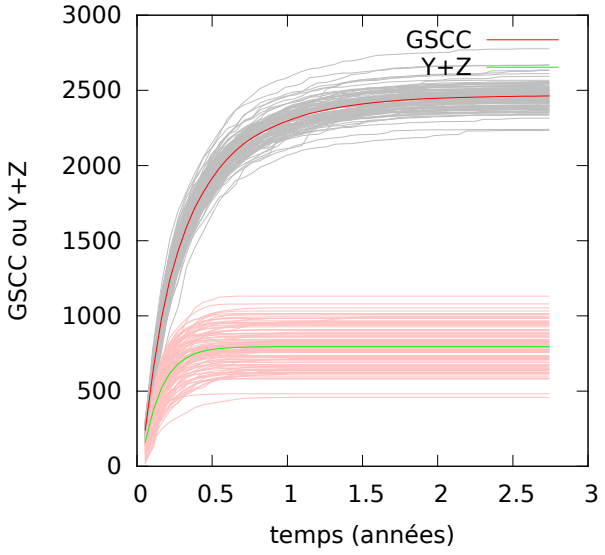
(c) $\zeta = 1$, $\frac{1}{\gamma} = 1$ year

(d) $\zeta = 1$, $\frac{1}{\gamma} = 10$ years

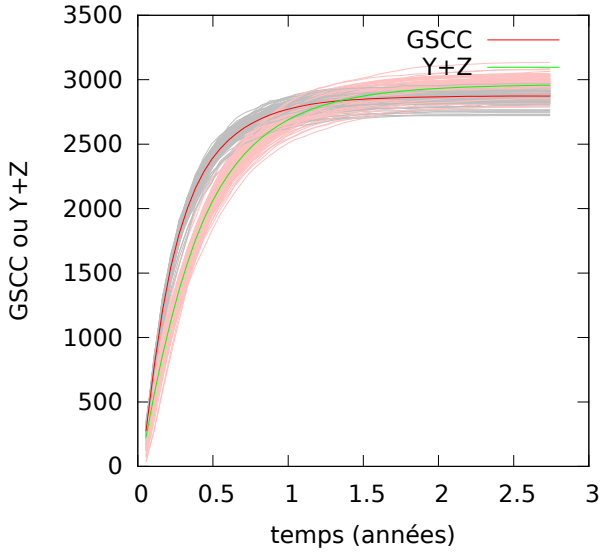
FIGURE 7.9 – Size of the LSCC and $(Y + Z)$ as function of time for increased friction. We take $\kappa_{ij} = 10 \times \hat{q}_{ij}$.



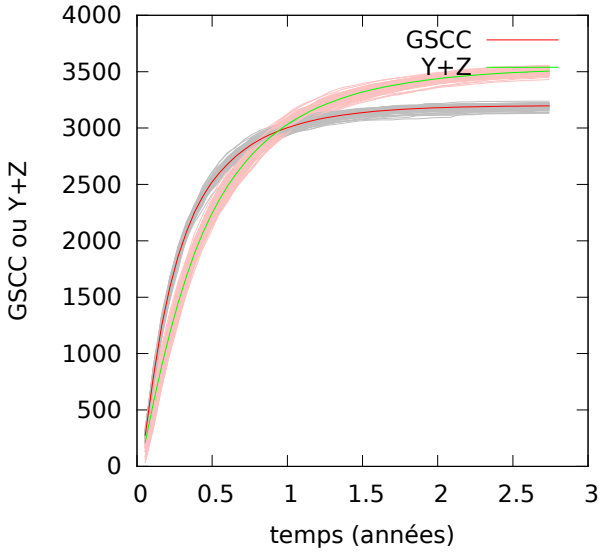
(a) $\zeta = 0.01$, $\frac{1}{\gamma} = 10$ years



(b) $\zeta = 1$, $\frac{1}{\gamma} = 1$ month



(c) $\zeta = 1$, $\frac{1}{\gamma} = 1$ year



(d) $\zeta = 1$, $\frac{1}{\gamma} = 10$ years

FIGURE 7.10 – Size of the LSCC and $(Y + Z)$ as function of time for decreased friction. We take $\kappa_{ij} = 0.1 \times \hat{q}_{ij}$.

Quatrième partie

Discussion

Chapitre 8

A tale of exchanges and epidemics : main insights and prospects

As the saying goes, you can't have your cake and eat it. The same holds true for exchanges among agents, typically individuals or companies, which fulfill various needs such as reproduction and profits, but do not come without efforts and risks (part I). Exchanges are costly to agents, which translates into a series of interaction constraints shaping rich contact structures. Such contact structures constitute the medium along which exchanges occur, but can also support the transmission of infectious diseases (part II). Subsequent outbreaks may result in considerable damages to infected agents and their goods, which in turn can lead to potential spillover effects including on other agents due to subsequent disruptions of exchanges, implementation of control measures and complex behaviourial adaptation (part III).

In an increasingly globalised world, exchanges are more frequent and take place over wider geographical areas. It follows the overall exposure of agents to infectious risk is expected to increase in the forthcoming decades. For the same reasons, and because decisions influencing the transmission of exchange-driven epidemics are taken by agents at various scales, the types and numbers of agents permissive to infection and/or involved in the management of epidemics should also grow.

We start by recalling the main results and messages of this work which contributes to address the growing challenges presented by exchange-driven epidemics (section 8.1). Then, we present some perspectives to extend current limitations of our approaches, and apply our latest modelling frameworks to new systems and questions (section 8.2). We finally conclude this work by a series of general remarks (section 8.3).

8.1 Key results and messages

Tackling exchange-driven epidemics appears highly challenging due to the plethora of potential causative factors and feedback loops dynamically linking exchanges to infections and infections to

exchanges. Blocking all exchanges would surely prevent exchange-driven epidemics, but would also deprive ourselves of tremendous benefits. To hopefully get the best of both worlds, i.e. to suggest control policies preventing or mitigating exchange-driven epidemics while maintaing exchanges as much as possible, it appears crucial to identify the most important mechanisms leading to joint exchange and disease transmission dynamics.

We briefly recal our overall findings which contribute to shed light on exchange-driven epidemics and on potential measures to control them without necessarily disrupting exchanges. Then, we highlight general take-home messages that can be formulated from our work.

8.1.1 Overall findings

Though at first sight HIV appears to bear no resemblance to FMD, both infections are driven by exchanges (sex and trade respectively), which arise in turn as a consequence of agents' decisions to exchange (individuals who wish to engage in sexual intercourses and agricultural holdings which are willing to trade livestock respectively). All exchange-driven epidemics are influenced by agents, namely intelligent entities which pursue particular objectives through exchanges and/or have the ability to take decisions affecting such exchanges. Though diverse in their nature and particular motivations, agents have in common to experience limitations to their interactions, are willing to engage in exchanges and demonstrate responsive capacities, e.g. when facing a crisis such as an epidemic. In other words, all agents face interaction constraints (interactive limitations) and exhibit adaptive host behaviour (willingness to engage in exchanges, which is modulated by perceived modifications in the environment).

By combining a series of mathematical models and data analyses, we shed light on interaction constraints and adaptive host behaviour, and investigate their influence on joint exchange and epidemic dynamics. The resulting mechanistic models of exchange-driven epidemics can prove instrumental in elaborating suitable interventions.

We discuss in two steps the main findings arising from our approaches. First, we summarise how interaction constraints contribute to shape exchange-driven epidemics when assuming passive agents. Second, we highlight the implications of adaptive host behaviour with constrained interactions on joint exchange and disease dynamics.

Interaction constraints shape exchange networks and disease dynamics

If exchanges were costless, all agents could interact in an homogenous-mixing fashion. In practice, empirical epidemiological contacts structures are varied and do not match the homogenous-mixing model. There exist finite exchange budgets arising from limited agents' resources and physical abilities. Such exchange budgets translate into interaction constraints, which limit possible interactions among agents. In this work, we study four types of constraints : the *sparseness constraint*, the *weighting constraint*, the *directional constraint* and the *frictional constraint*. Here, we discuss the main results

arising from our study of interaction constraints when assuming agents are passive. The last constraint we cite, namely the frictional constraint, is rather discussed as part of the analysis of our adaptive behavioural models, because its adequate study requires an explicit modelling of exchange dynamics.

We start with the implications of interaction constraints on the architecture of exchange networks. In contrast with homogenous-mixing, interaction constraints lead to diversified epidemiological contacts structures, as typically observed in data (parts I—III). One central reason explaining this diversity is that even if all agents face constraints, each particular constraint is not affecting all agents identically. First, in systems such as livestock-exchange markets, agents differ in their exchange budgets : a typical trading agent has far more resources for interacting than a typical farm (chapter 5). Second, even in systems where the exchange budget per agent is roughly constant, agents arbitrate — assuming they can decide — differently between individual constraints to spend their exchange budget. For instance, in sexual-contact networks, some individuals have many sexual partners (weak sparseness constraint) but few sex acts per partner (strong weighting constraint), while other agents exhibit the opposite trend (strong sparseness constraint and weak weighting constraint, i.e. few partners and many sex acts per partner). Generally speaking, for a given aggregated exchange level, agents can locally differ in many ways, notably in their relative number of partners (sparseness constraint), their relative exchange weight per partner (weighting constraint) and their relative degree of specialisation (directional constraint, which usually reflect multipartite networks). We can reasonably expect constraints are inter-related by multiple trade-offs because any agent has to arbitrate between various conflicting constraints to reach a given exchange objective at a given exchange budget. In practice, the sparseness constraint is almost always present in empirical contact structures, and additional constraints may apply in particular systems, which further tend to increase the diversity in contact structures.

Then, we discuss the epidemiological consequences of the contact structures implied by interaction constraints. Accounting for constraints will, unsurprisingly, tend to reduce the expected severity of outbreaks by reducing underlying exchanges. We hence concentrate on the various epidemiological responses resulting from the combination of multiple constraints when the overall level of exchange is kept constant. The range of associated epidemiological responses appears extremely rich and arises from the diversity of constrained contact structures. Importantly, this diversity in epidemiological dynamics is explained by differing combinations of constraints at the agent- and interaction-level (parts II and III). We now summarise the influence of diverse sources of agents' diversity on exchange-driven epidemics.

First, almost all agents in empirical exchange systems exhibit a sparseness constraint, where the corresponding degree distribution is generally heterogeneous (e.g. power-law, with a long tail ; or negative binomial, with a shorter tail). For a given number of links, networks with heterogeneous degree

distributions tend to aggravate epidemics in comparison with networks with homogeneous degree distributions (chapter 2). It follows concentrating interventions on the most connected agents, i.e. the 'hubs', is generally far more efficient than targeting agents for control uniformly at random [Albert *et al.*, 2000]. These results are already largely documented in the literature [Pastor-Satorras & Vespignani, 2001], and we only recall them to provide a general picture of the influence of interaction constraints on epidemic processes.

Second, we discuss the consequences of heterogeneity in agents' relative exchange weight per partner. While all agents are rather roughly identical in weights with an heterogeneous degree distribution, as observed for the total number of sex acts per individual per time period, the resulting distribution of weights per partnership is heterogeneous, which eventually leads to drastic reduction in epidemic severity and invasive capacities in comparison with settings where the same aggregated weight is homogeneously distributed at the link-level. Targeting hubs preferentially in systems with heterogeneous weights per link is only moderately more efficient than random targeting, a result suggesting general information campaigns may be more cost-effective than targeted interventions (chapter 3).

Third, agents may differ in their degree of specialisation, as observed in markets of ornamentals and livestock-exchange with suppliers, wholesalers and demanders. The particular specialisation we investigate relates with the relative position of trading agents in the supply chain, and suggest wholesalers are more at risk of contagion and transmission than suppliers and demanders, which imply they should be targeted preferentially for control. This result was obtained both theoretically by keeping the total number of directed links constant for contrasted network architectures (chapter 4) and empirically when controlling for the exchange level per agent in French livestock market (chapter 5).

Adaptive host behavior with constrained interactions leads to complex responses in joint exchange and infection dynamics

Agents are not passive. Rather, they decide to interact for various reasons, and can modulate their interactions in response to perceived change in their environment, a set of mechanisms which may eventually generate complex feedbacks between exchange and epidemic dynamics. While our general motivation lies in the study of exchange-driven epidemics, we concentrate our latest work on a particular type of exchange, namely trade of goods which propagate infections. Trade arises from complex agents' behaviours, and in the case of particular markets such as trade of livestock, can be thoroughly studied thanks to existing datasets recording animal movements among agents. We hence develop a series of mechanistic economic-epidemiological (eco-epi) frameworks where transient and long-term market dynamics are determined by agents' adaptive behaviour and interaction constraints, and can influence disease transmission. Since agents necessarily integrate all constraints they face when deciding to engage in a transaction, interaction constraints are now subsumed into adaptive behaviour. Here, we recall our main findings on the influence of adaptive host behavior on joint trade and infection

dynamics, and suggest appropriate course of action to tackle epidemics in such systems.

First and foremost, our approaches suggest trade pathways propagate epidemics differently from other routes of transmission due to a noteworthy contact structure, which is by essence adaptive, and which directly emerges from agents' motivation to achieve high-levels of profits and well-being (chapter 6). From an epidemiological perspective, this contact structure is dynamical — more precisely transient and adaptive — and consists in potentially recurrent transactions between exchanging agents. Each transaction corresponds to a shipment of a variable volume of goods that may be contaminated and lead to the spread of infection. In practice, the rates at which transactions occur are limited for various reasons. Among others, goods require to be produced prior being shipped, agents need to find suitable trading partners, a process which takes time, and the logistics of stock delivery are costly. In our eco-epi models (chapters 6,7), transaction rates are governed by coefficients of friction κ_{ij}, which aggregate all barriers to transactions, and are defined at the scale of directed pairs of interacting agents (i, j). Under reasonable assumptions, κ_{ij} simply corresponds to the average number of goods exchanged per transaction from agent i to agent j, which can be easily calculated from available data. By controlling the frequency of interaction of the pair (i, j), κ_{ij} quantifies the frictional constraint at a microscopic level (chapter 7). At an aggregated level, the overall level of friction is given by coefficient κ, and quantifies the typical number of goods shipped per transaction (chapter 6).

We now evoke the economic-epidemiological implications of differing overall frequency of interaction κ in a given market when trade is the dominant route of transmission (chapter 6). Our findings show that frictional markets ($\kappa \geq \kappa_{\min} = 1$) are associated with a specific response to infectious diseases contrasting with the response of other complex systems sustaining epidemics and often assumed to be frictionless ($\kappa \to 0^+$). In particular, market friction can be a significantly stronger determinant of epidemics than risk-aversion behaviour. κ is a central parameter governing joint trade and disease dynamics : increasing κ can increase market equilibration time by several orders of magnitude and prevent the invasion of trade-driven epidemics. In our models, except in cases where delays to conclude a transaction may result in stock losses, changing κ do not appear to affect the overall trade volume exchanged per time unit on the long run. As a general prevention measure, our results imply contagion in markets could be minimised by increasing friction, which would result in larger-volume, less-frequent transactions, without necessarily affecting overall trade flow, and therefore, business activity. Our findings hence suggest that while highly frictional markets can still be responsible of disease introductions, trade itself cannot sustain epidemics in such markets. As a preliminary result, this counter-intuitive finding appears to hold at the microscopic level for a particular subsystem of the French cattle-exchange markets (chapter 7). Through percolation experiments, we indeed show increasing all empirically estimated κ_{ij} (by e.g. introducing a flat tax on transports) or only the minimal one $\min_{ij} \kappa_{ij}$ (by e.g. imposing a law to ban transactions lower than a threshold number of goods), eventually results in

effective disease prevention. However, increasing friction may be difficult to achieve in practice due to practical constraints (e.g. trucks have a limited size), additional production costs (e.g. due to increased stocks) or risk of social upheavals (as e.g. observed in France in 2013 when the French government was willing to levy a tax on heavy road transport, and which was eventually revised due to heavy protests from truck drivers and farmers in Brittany).

At a wider scale, our eco-epi frameworks enables the comparison of contrasted trade markets to prioritise the allocation of resources for disease control. As an example of application, while both movements of swine and cattle can transmit infections, our results suggest that in France, the epidemic risk, both in probability of invasion and in size, is substantially larger in cattle than swine due to lower overall friction in cattle, even though overall trade volume in swine is larger (chapter 6). This finding is in line with the descriptive arguments of Rautureau *et al.* [2012], who report lower size of the largest strongly connected component (LSCC) in swine compared with cattle. Note that our theoretical results indeed indicate that systems with higher friction exhibit lower size in their LSCC for a given period of observation (chapter 7). Generally speaking, we believe applying our eco-epi framework at large scales could be instrumental in order to compare epidemiological risks associated with different markets and countries in the EU, which could eventually help in prioritising Community resources accordingly.

From the converse standpoint, our frameworks also prove useful to predict the influence of epidemics on trade dynamics in continuous time. In contrast, economic models often focus on descriptive assessment neglecting temporal dynamics (e.g. [Soliman *et al.*, 2010]) or predictive equilibrium-based discrete-time approaches (e.g. [Zhao *et al.*, 2006]) that can be at odds with the multiple and contrasting time scales encompassed by economic and epidemic processes. It is important to model in continuous time to adequately resolve the multiple processes of infection and the market transactions. In addition, our frameworks depart from existing economic approaches by allowing for disequilibrium between aggregated supply and demand (Appendices of chapter 6). Accounting for such disequilibrium is important, because trade-driven epidemics likely cause asymmetric shocks to markets, in the form of persistent excess of either aggregated supply or aggregated demand, which in turn influences infection dynamics (chapters 6 and 7). For instance, with the exception of wholesalers, transmission occurs from suppliers to demanders. Therefore, suppliers and demanders would generally be differently impacted by disease, and would also be differently targeted by disease-control measures. In practice, the specific distortions to the balance between supply and demand appear to depend on the specific market and epidemic at hand (chapter 5).

8.1.2 Take-home messages

In addition to novel results enlightening particular aspects of exchange-driven epidemics, we believe three general take-home messages can be derived from our work, all of which do not restrict to the

particular mechanisms, model specifics and systems investigated. First, we argue in favour of dual approaches to study exchange-driven epidemics. Second, we advocate coupling of natural and social processes is central to design mechanistic frameworks accounting for exchanges and epidemics. Third, we highlight the insights gained by confronting contrasted agents' standpoints.

Studying exchange-driven epidemics requires a dual methodology

When quantitative epidemiologists attempt to make sense of outbreaks and suggest possible measures for control, they are faced with a dilemma. Either they resort to simple models, typically building upon the standard SIR compartmental model [Keeling & Rohani, 2008], or they attempt to describe the epidemiological system at hand as realistically as possible, integrating all possible transmission pathways, scales, evolutionary and demographic mechanisms, existing data and so on. As an example, the GLEaMviz project (`http://www.gleamviz.org/`) aims at describing epidemics at multiple scales by integrating sociodemographic data and population mobility within spatially embedded networks, namely the airline-transportation network at the global level [Hufnagel *et al.*, 2004] and commuting networks at the local level [Keeling *et al.*, 2010].

For our particular problem, modelling exchange-driven epidemics as realistically as possible appears an impossible task, because exchanges *per se* and agents' behaviours underlying such exchanges are diverse, complex and extremely difficult to understand even without epidemics. In particular, understanding how trade emerges and operates remains a key open and controversial question [Hahn, 1982; Katzner, 2010]. So does sexual behaviour [Foucault, 1984]. We hence organise our work into two progressive steps. We start with the in-depth analysis of interaction constraints and their influence on epidemics, assuming agents exhibit a passive behaviour (part II). Then we allow for adaptive behaviour with constrained interactions and explore their overall influence on joint exchange and disease dynamics (part III). Importantly, these two steps are complementary (table 8.1). Assuming passive agents allows us to gain detailed insights on transmission networks and to approximate overlying epidemic processes with analytical tools. Such insights and tools can be in turn applied to models where adaptive host behaviour is included, typically to explore such models under some approximations, but also to develop intuitions on which model components should be made explicit. Conversely, exploring the influence of adaptive behaviour on disease dynamics can enable us to identify particular cases where assuming passive behaviour is reasonable. These two approaches mutually enrich each other, and enable us to progressively sketch, and improve, by trial and error, a general modelling framework for exchange-driven epidemics.

Though our main application is directed to epidemics driven by trade, we believe this type of dual approach has the potential to shed light on any exchange-driven epidemic, and more generally to any natural system influenced by agents' behaviours and decisions.

TABLE 8.1 – **Pros and cons of neglecting or accounting for adaptive behaviour.**

	Advantages	**Disadvantages**
Passive agents	— easier parametrisation — useful analytical approximations — large demand and acceptability from the epidemiological community	— exchange dynamics, if any, lack mechanistic foundations — agents are assumed passive when submitted to shocks
Adaptive agents	— mechanistic models, with potentially counter-intuitive results — large institutional support — Identification of cases where assuming passive agents is plausible	— agents behaviour may not always be amenable to quantitative approaches, especially on short time scales and in face of major disruptions — exhaustive data are required — cross-disciplinary approaches seem harder to communicate, e.g. in the economic community.

Coupling natural and social processes is central to design mechanistic frameworks and understand exchanges and epidemics

Existing work essentially relies on the structure—function paradigm, where networks of contacts among agents eventually determines infection transmission [Strogatz, 2001; Newman, 2003]. Such networks of contacts, which can be dynamical, are either given by data [e.g. Duncan *et al.*, 2012] or generated with models replicating certain features observed in empirical networks [e.g. Barabasi & Albert, 1999]. Even when dynamical, such contact structures are assumed passive, and the complex socio-economic processes explaining their existence are not taken into account. Though at odds with reality, the structure—function paradigm has the advantage of yielding detailed insights on the potential structural factors affecting infectious disease dynamics, which is precisely why we adopt this point of view when studying interaction constraints (part II)

Some recent epidemiological studies [Gross *et al.*, 2006; Funk *et al.*, 2010; Morin *et al.*, 2013] slightly depart from the structure—fonction paradigm by considering contact adaptation through risk aversion, a particular behaviour where agents take proportionate protective measures as function of the perceived level of epidemic threat, which may change over time. But risk aversion does not explain why interactions exist in the first place, which requires to explicitly consider complex socio-economic processes.

Though we do not claim to deeply understand how exchanges emerge, we depart from existing approaches by explaining how plausible assumptions on agents' behaviours can be incorporated in mathematical models to account for the joint dynamics of exchanges and epidemics. Importantly, the mechanistic nature of such models implies the structure of contacts emerges as a *consequence* of agents' decisions to exchange, which will be affected by perceived threats such as epidemics, but fundamentally determined by socio-economic motivations. We apply this "function—function paradigm" to two types

of fundamental motivations : the willingness to have sex [Moslonka-Lefebvre *et al.*, 2012a], and the willingness to trade (part III). Note that [Moslonka-Lefebvre *et al.*, 2012a], which we evoke in (part II), is actually in between the two paradigms : while the unweighted sexual contact network is extracted from data (structure—function paradigm), we attempt to explain, with contrasted behavioural assumptions, how sex acts are distributed on this particular network (function—function paradigm), with important consequences on disease dynamics. We finally highlight our mechanistic model results in dynamical exchange networks, but all dynamical exchange networks are not generated by mechanistic models of exchange among agents [e.g. Perra *et al.*, 2012].

Considering differing agents' standpoints matters

While only biological entities can be infected (individuals or living 'products' traded by companies such as livestock), various agent types are potentially concerned with epidemics at multiple scales : local (e.g. individuals, small companies, cities), infra-national (e.g. departmental and regional public authorities), national (e.g. States) and international (e.g. multinational corporations, international organisations). As shown by our economic-epidemiological analyses of livestock-exchange data, neglecting one standpoint may drastically affect the recommended course of action to prevent outbreaks (chapter 5). For instance, let us take the case when the cost of epidemic prevention is supported by a public authority responsible for epidemic control, and is essentially proportional to the total number of agricultural holdings protected. From the sole point of view of the public authority, protecting preferentially those holdings with the largest market shares appears the best strategy, which is a standard recommendation in the literature for similar systems [Albert *et al.*, 2000]. In sharp contrast, when the point of view of the market is also taken into account, it may be actually more interesting to protect preferentially economically-minor holdings. Generally speaking, our findings suggest it is important to consider multiple agents' standpoints to adequately assess epidemic risks, damages, and best courses of action for control policies.

8.2 Perspectives

Exchange-driven epidemics are extremely diverse and complex, and our work, in combination with alternative approaches [e.g. Valdano *et al.*, 2014], constitutes only a first step towards an integrative framework to tackle them. Here, we present some research perspectives to overcome current limitations of our approaches and tailor our adaptive frameworks to new applications. We present each perspective in turn, following the reasoning employed so far. First, we place ourselves in a case where agents are passive, which could be particularly interesting to derive analytical tools to uncover the principles underlying epidemic dynamics on directed weighted networks. Second, we assume agents are adaptive. We would like to assess whether certain agents in a market can be considered *not* "too important

to fail" from an economic-epidemiological standpoint. We then discuss two systems which could be studied with our adaptive frameworks, namely the spread of *P. ramorum* in the forestry industry and the natural environment, and the propagation of STI on dynamical sexual-contact networks. We also evoke potential improvements to better account for endemic infections in our adaptive models. Finally, we discuss how existing datasets and model selection techniques could be combined to validate or reject certain putative mechanisms identified as part of our work.

8.2.1 Deriving analytical tools for epidemics on weighted directed networks

Analytical insights are instrumental in understanding the factors determining epidemic invasion and dynamics [Anderson & May, 1991]. For instance, in this work, we derive a series of mathematical tools to shed some light on SIR epidemics propagated on weighted networks (chapter 3). These particular tools are applicable as such to some STI, because the risk of transmission in pairs of sexual partners, is, as first approximation, bidirectional, and key STI such as HIV can be described by SIR processes. But all STI are not perfectly approximated by the SIR process on weighted networks. For most STI, the risk of transmission for a given sexual practice typically differs between donors and recipients, which translates into bipartite directed networks [Gomez-Gardenes *et al.*, 2008]. Though transmission for STI is generally asymmetrical rather than unidirectional, formalising weights leading to infection as undirected is clearly a simplifying assumption. And other epidemics such as Chlamydia are rather described by a SIS epidemic process, because infected individuals recover without developing immunity.

In a more general case, it would be interesting to derive analytical tools to improve our understanding of SIRS epidemics propagated on weighted and directed networks [1]. The contact structures underlying exchange-driven epidemics such as STI are indeed better approximated by weighted directed networks, i.e. networks with asymmetrical weights. This feature would be even more important for trade of livestock, which can be highly asymmetrical and even strictly unidirectional in trade volume per time unit for certain pairs of agents (chapter 5). Besides, the SIRS process is also more general, as allows to describe both endemic states (such as the SIS process, with a persisting infection) and outbreaks which eventually die out (such as the SIR process).

1. To gain analytical insights on SIRS epidemics propagated on weighted directed networks, we could, following the methods described in chapter 3, resort to heterogeneous-mean field assumptions coupled with generating functions of quadri-dimensional distributions specifying the relative frequency of agents' classes, where classes are defined by agents with identical characteristics in [in-weight, in-degree, out-weight, out-degree]. One of the interests and difficulties of this approach lies in considering concurrently weights and directions of links. If successful, this approach could enable us to derive key indicators such as the invasion threshold (R_0) and rate of early epidemic growth (r_0). Such tools could be applied to adaptive frameworks by aggregating the resulting dynamical networks over a certain period of time, which would provide a realistic upper bound for quantities such as R_0 and r_0. The choice of the most relevant period of time could be made depending on the typical response time of adaptive systems, which in our eco-epi models depends on friction (part III). We however foresee two difficulties. First, the re-introduction of susceptible is likely to render analytical results harder to derive, especially to approximate full epidemic trajectories. Second, all possible correlation patterns in the multidimensional distributions may actually be impossible to translate into actual networks due to topological constraints [Serrano *et al.*, 2006].

8.2.2 Investigation of global spillovers : the "too important to fail" phenomenon

All important costs and benefits should be taken into account when comparing policies to bring an infectious disease under control, as neglecting one of the latter can change the recommended course of action [Klein *et al.*, 2007]. While efficient to mitigate epidemics [Albert *et al.*, 2000], targeting prevention measures on central holdings of large size can be costly to the regulator and disrupt the good functioning of markets by creating economic shocks (chapter 5). To the best of our knowledge, the epidemiological implications of such costs remain to be explored with adaptive models of joint trade and infection dynamics, such as the heterogeneous market-epidemiological (hME) model (chapter 7)

Along this line, a knowledge gap that we promptly plan to address with the hME model is the identification of categories of agents that are *not* "too important to fail" (TITF) when submitted to epidemic shocks and subsequent control measures. The concept of TITF stems from economics and has been widely associated with major financial institutions that are both a) locally important (e.g. assessed by receipts or number of employees) and b) systemically important (e.g. assessed by number of business partners or any other centrality measure). A TITF company is so important for the stability of the market as a whole that it cannot be allowed to undergo bankruptcy. The flip side of the coin is that TITF businesses tend to anticipate this tacit rescue by public authorities. TITF regimes hence generate incentives for large businesses to become even bigger than they would normally, thus increasing the risk of global market failure [Kellermann, 2011]. We believe the notion of TITF is peculiarly insightful and relevant for trade-driven epidemics. In empirical markets conducing infections such as livestock markets, there is indeed a small but significant frequency of agents that are both locally and systemically important. These agents can behave as epidemiological super-spreaders, and, if infected, may lead to a market break down by infecting in turn most agents in the market. To this extent, it appears essential to preferentially target disease control measures on those key market players. Conversely, the very same agents play a key role in the functioning of the market, and their removal from the contact structure responsible for infection transmission (by e.g. quarantine measures or excessive prevention) is likely to yield major market imbalances. We hence face a dilemma, as targeting control measures against such TITF agents may result in more harm than good. One possible solution lies in findings other categories of agents to target preferentially. In other words, would like to address the inverse problem and identify agents that are *not* "too important to fail" (nTITF).

In practice, using our market-centric categorisation of agents applied to livestock-exchange data (chapter 5), we can reasonably expect those wholesalers with the largest flow shares to be TITF. And it would be interesting to identify market categories that are nTITF as far as the hME model is concerned. In a sense, we ask exactly the same questions as in chapter 5, but with a dynamical model of adaptive trade which accounts for complex feedbacks loops between epidemics and market dynamics.

8.2.3 Applying the eco-epi frameworks to other contexts

By construction, our eco-epi frameworks are essentially designed to study trade-driven epidemics. However, the ideas underlying such frameworks are general and could be transposed in other contexts. Here, we evoke two alternative applications on which we plan to work : the spread of *P. ramorum* in the forestry industry and in the wild, and the propagation of STI.

Adaptive behavior and the spread of *P. ramorum* in the open

Phytophthora ramorum (*P. ramorum*), the causative agent of Sudden Oak Death (SOD), is a generalist plant pathogen that has emerged in Europe and the USA since the mid-1990s. Until recently, much of the damage in Europe has been limited to nurseries, through shipments of contaminated plants, and relatively confined outbreaks in heathland and woodland in England. However, the recent outbreak in Japanese larch plantations in the UK has shown that some European strains of *P. ramorum* have the ability to spread rapidly and at a large scale in the natural environment [Grunwald *et al.*, 2012]. Given the large losses of nursery stock, and the depletion of oak and tanoak populations in California and Oregon induced by *P. ramorum* during the past decade, the sudden aggravation of the UK epidemic has raised concerns of a 'doom scenario' replicating the severity of the USA epidemic. This is especially worrying for larch timber crops and natural ecosystems, because of large losses from damages to forestry industries and property, and the difficult and expensive control in the wild, which is by definition open, large, and complex.

In such a context, it would be interesting to identify efficient strategies to limit the global spread of *P. ramorum* under adaptive market and budget drivers. Building upon the compartmental eco-epi framework (chapter 6), we could model a bioeconomic system made of three inter-related components : *i*) larch timber crops, *ii*) other timber crops in competition with the larch industry, and *iii*) the natural environment. Two differing types of interactions are expected in such a system. First, epidemiological interactions : while other timber crops are insensitive to the pathogen, larch crops and a subset of the natural environment can be jointly infected by *P. ramorum*. Second, economic interactions : larch crops and other timber crops are in competition to target the same demand for timber. The coupled economic and epidemic dynamics of such a system are probably largely determined by human behaviour such as decisions to harvest, replant and apply control. We could hence explicitly consider two types of agents who exhibit adaptive behaviour with respect to one another's decisions and to trade and disease dynamics : private players who own forest plots and decide the type of forest crop planted, when and how much is harvested and the amount of private control resources invested against SOD ; and public players, who decide when and how much public funds are invested to mitigate SOD. With such a model, we could predict, for various behavioural scenarios, the long-run evolution of the system and quantify biological and economic losses.

STI dynamics with adaptive behaviour

Prostitution is typically a market contributing to the spread of STI [Rocha *et al.*, 2011], but sexual contacts obviously do not restrict to prostitution. One idea would be to extend our models accounting for sexual behaviour (chapter 3) based on the microscopic eco-epi framework (chapter 7), whose components are flexible and could be easily adjusted to capture STI dynamics under contrasted behavioural assumptions.

One interesting study would consists in comparing various assumptions to account for observed STI dynamics. For instance, following Vernon & Keeling [2009], it would be interesting to compare sexual-contact networks that are dynamical or static and/or weighted or unweighted (four possible combinations), and assess which model is more in agreement with data. A properly designed adaptive framework could lead to these four network types.

8.2.4 Tailoring the adaptive frameworks to endemic diseases

With the exception of chapter 4, we assume agents are either entirely susceptible, entirely infectious, or entirely detected and removed. This assumption is sound for quick epidemics such as FMD, but can be inadequate for slow-spreading infections which tend to become endemic such as HIV, with complex feedbacks between intra-agent and inter-agent dynamics [Mideo *et al.*, 2008]. To integrate intra-agent dynamics, a solution would consist in resorting to metapopulation models [Colizza & Vespignani, 2008; Schumm, 2013], and modify our adaptive frameworks accordingly.

From an economic perspective, we can also remark our eco-epi approaches only stand for emerging or re-emerging epidemics since agents adapt their behaviour in response to outbreaks but do not really anticipate them. Further work of interest would transpose the current eco-epi frameworks to endemic diseases that are often anticipated by market agents. Agent anticipation could be implemented using explicit inter-temporal optimisation functions that are widely used in economics [e.g. Zhao *et al.*, 2006].

8.2.5 Selecting models against data to assess mechanistic assumptions

The market models developed as part of this work rely on various mechanistic assumptions on how exchanges operate. Examples of such putative mechanisms include supply-and-demand trade dynamics based on agents' stocks in supply and demand, the existence of trade friction which tends to reduce the frequency of exchanges, a price formation mechanism based on supply and demand creation rates and which will in turn influence the rates at which supply and demand are created and result in actual trade flow. Though probably less convenient to account for epidemic processes, alternative market models and exchange mechanisms are employed in the literature and rely on other types of assumptions than ours (see sections E.1 and E.3.1). Correctly describing such exchange mechanisms is essential since exchanges among agents form the contact structure responsible for the propagation of exchange-driven

epidemics. This raise the question of which models perform better to account for market and integrated market - epidemiological dynamics as observed in real systems.

When compared against existing datasets, model selection techniques [e.g. Courcoul *et al.*, 2010; Sunnaker *et al.*, 2013] can be used to infer which model mechanism(s) and model(s) prove better than others. We could use such techniques to validate or reject our market models and exchange assumptions against data and further comparison with alternative models. Since our market-based models rely on the min function which generates identifiability issues, we rely on Approximate Bayesian Computation (ABC) as model selection techniques [Sunnaker *et al.*, 2013]. In practice, we could apply this inference process in two steps. First, we could focus on isolated market systems compared against economic data such as microeconomic price and exchange data since they are available for livestock systems in France. This approach would enable us to compare contrasted market models for price dynamics and (dis)-equilibrium. We could test whether the law of supply-and-demand applies to livestock markets, and if so, based on which signal (excess of supply stock or rate for instance). We could assess whether agents actually act by optimising their choices [e.g. Gohin *et al.*, 2013], in which way (shape and components of utility functions) and over which time scale (length of the optimisation window). We could also estimate friction without assuming steady-state as we do here, and test whether assuming steady-state for livestock markets is plausible on short timescales. Other questions of economic interest include whether the "best" market model change depending on the system at hand (e.g. cattle versus swine), the timescale considered (short versus long), the granularity of the model (e.g. microscopic versus macroscopic scale) Second, we could focus on coupled market – epidemiological systems compared against economic and epidemiological data to asses whether our coupling mechanisms are plausible. Besides, a comparative model selection approach (market versus market – epidemiological models) would enable us to assess to which extent emerging epidemics and endemic diseases tend to change the usual behaviour of agents such as farmers and managers of markets, assembling centers and slaughterhouses.

8.3 Conclusion

A subtle understanding of exchanges and subsequent infection events requires to assess and integrate contrasted processes. In this work, we investigate two particular mechanisms exerting a key influence on joint exchange and disease dynamics, namely interaction constraints and adaptive host behaviour. Both mechanisms allow us to develop integrative frameworks, which, to the best of our knowledge, constitute the first models of exchange-driven epidemics explicitly accounting for motivations underlying exchanges. We apply our models and ideas to shed light on STI dynamics in sexual-contact networks and trade-driven epidemics as frequently observed in livestock-exchange markets, and suggest in turn relevant disease control policies. Pretending to exhaustively capture complex human behaviours underlying exchanges with mathematical models is probably — and hopefully — infeasible. But exclusively studying exchange-driven epidemics by neglecting such behaviours constitutes an over-simplifying assumption. Rather, we believe our capacity to manage exchange-driven epidemics relies on our ability to integrate contrasted processes at the crossroad of epidemiology and social sciences, even when the behavioural mechanisms underlying exchanges are not precisely understood and should be instead explored theoretically, e.g. by comparing contrasted scenarios.

Besides epidemiology, we believe our approaches can be transposed to any agent-based system where agents exhibit adaptive behaviour and face interaction constraints. We could investigate ecological and evolutionary questions — for instance, study the dynamics of foraging and the evolution/adjustment of foraging behaviour [Sokolowski *et al.*, 1997] — or explore the propagation of financial shocks among banks [Arinaminpathy *et al.*, 2012]. From a more intellectual perspective, our work encompasses many fields, systems, theoretical tools and applications, and illustrates as such the typical insights gained by fostering cross-disciplinary approaches in science.

Bibliographie

Acemoglu, D., 2008 *Introduction to Modern Economic Growth*. Princeton University Press.

Albert, R. & Barabasi, A.-L., 2002 Statistical mechanics of complex networks. *Reviews of Modern Physics* **74**, 47.

Albert, R., Jeong, H. & Barabasi, A.-L., 2000 Error and attack tolerance of complex networks. *Nature* **406**, 378–382.

Anderson, C. M., Plott, C. R., Shimomura, K.-I. & Granat, S., 2004 Global instability in experimental general equilibrium : the scarf example. *Journal of Economic Theory* **115**, 209–249.

Anderson, R. & May, R., 1991 *Infectious diseases of humans : dynamics and control*. Oxford University Press, New-York, USA.

Anderson, R. M. & May, R. M., 1992 Understanding the AIDS pandemic. *Scientific American* **266**, 58–61, 64–66.

Aparicio, J. P. & Pascual, M., 2007 Building epidemiological models from R0 : an implicit treatment of transmission in networks. *Proceedings of the Royal Society B : Biological Sciences* **274**, 505–512.

Arinaminpathy, N., Kapadia, S. & May, R. M., 2012 Size and complexity in model financial systems. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18338–18343.

Atalay, E., Hortaçsu, A., Roberts, J. & Syverson, C., 2011 Network structure of production. *Proceedings of the National Academy of Sciences of the USA* **108**, 5199–5202.

Bajardi, P., Barrat, A., Savini, L. & Colizza, V., 2012 Optimizing surveillance for livestock disease spreading through animal movements. *Journal of the Royal Society interface* **9**, 2814–2825.

Bansal, S., Grenfell, B. T. & Meyers, L. A., 2007 When individual behaviour matters : homogeneous and network models in epidemiology. *Journal of the Royal Society Interface* **4**, 879–891.

Barabasi, A. & Albert, R., 1999 Emergence of scaling in random networks. *Science* **286**, 509.

Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A., 2004a The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3747–3752.

Barrat, A., Barthelemy, M. & Vespignani, A., 2004b Modeling the evolution of weighted networks. *Physical Review E* **70**, 066149.

Barro, R. J. & Sala-i Martin, X., 2003 *Economic growth (2nd edition)*. The MIT Press.

Barthelemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A., 2004 Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Physical Review Letters* **92**, 178701. PRL.

Bartlett, M., 1949 Some evolutionary stochastic processes. *Journal of the Royal Statistical Society : Series B.* **11**, 211–229.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P. & Caldarelli, G., 2012 Debtrank : Too central to fail ? financial networks, the fed and systemic risk. *Sci. Rep.* **2**. 10.1038/srep00541.

Blower, S. M. & Boe, C., 1993 Sex acts, sex partners, and sex budgets : Implications for risk factor analysis and estimation of HIV transmission probabilities. *Journal of Acquired Immune Deficiency Syndromes* **6**, 1347–1352.

Boguna, M., Pastor-Satorras, R. & Vespignani, A., 2003 Absence of epidemic threshold in scale-free networks with degree correlations. *Physical Review Letters* **90**, 028701.

Brandes, U., 2001 A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology* **25**, 163–177.

Britton, T., Deijfen, M. & Liljeros, F., 2011 A weighted configuration model and inhomogeneous epidemics. *Journal of Statistical Physics* **145**, 1368–1384.

Britton, T., Nordvik, M. K. & Liljeros, F., 2007 Modelling sexually transmitted infections : the effect of partnership activity and number of partners on R_0. *Theoretical Population Biology* **72**, 389–399.

Burdett, K., Shi, S. & Wright, R., 2001 Pricing and matching with frictions. *Journal of Political Economy* **109**, 1060–1085.

Buttner, K., Krieter, J., Traulsen, A. & Traulsen, I., 2013a Efficient interruption of infection chains by targeted removal of central holdings in an animal trade network. *PLoS ONE* **8**, e74292.

Buttner, K., Krieter, J., Traulsen, A. & Traulsen, I., 2013b Static network analysis of a pork supply chain in northern germany : characterisation of the potential spread of infectious diseases via animal movements. *Preventive Veterinary Medicine* **110**, 418–428.

Callon, M., 1998 Actor-network theory - the market test. *The Sociological Review* **46**, 181–195.

Campolongo, F., Cariboni, J. & Saltelli, A., 2007 An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software* **22**, 1509–1518.

Caputo, M. R., 2005 *Foundations of dynamic economic analysis : optimal control theory and applications*. Cambridge University Press.

Ceddia, M. G., Bardsley, N. O., Goodwin, R., Holloway, G. J., Nocella, G. & Stasi, A., 2013 A complex system perspective on the emergence and spread of infectious diseases : Integrating economic and ecological aspects. *Ecological Economics* **90**, 124–131.

Chu, X., Guan, J., Zhang, Z. & Zhou, S., 2009 Epidemic spreading in weighted scale-free networks with community structure. *Journal of Statistical Mechanics : Theory and Experiment* **2009**, P07043.

Chu, X., Zhang, Z., Guan, J. & Zhou, S., 2011 Epidemic spreading with nonlinear infectivity in weighted scale-free networks. *Physica A : Statistical Mechanics and its Applications* **390**, 471–481.

Clark, C., 1976 *Mathematical bioeconomics : the optimal control of renewable resources*. John Wiley and Sons, Inc.

Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A., 2006 The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the USA* **103**, 2015–2020.

Colizza, V. & Vespignani, A., 2008 Epidemic modeling in metapopulation systems with heterogeneous coupling pattern : Theory and simulations. *Journal of Theoretical Biology* **251**, 450–467.

Courcoul, A., Vergu, E., Denis, J.-B. & Beaudeau, F., 2010 Spread of q fever within dairy cattle herds : key parameters inferred using a bayesian approach. *Proceedings of the Royal Society B : Biological Sciences* **277**, 2857–2865.

Daley, D. J., Gani, J. & Gani, J. M., 2001 *Epidemic modelling : an introduction*, vol. 15. Cambridge University Press.

Danon, L., Ford, A. P., House, T., Jewell, C. P., Keeling, M. J., Roberts, G. O., Ross, J. V. & Vernon, M. C., 2011 Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases* **2011**, 28.

Danon, L., House, T. A., Read, J. M. & Keeling, M. J., 2012 Social encounter networks : collective properties and disease transmission. *Journal of the Royal Society Interface* .

Deijfen, M., 2011 Epidemics and vaccination on weighted graphs. *Mathematical Biosciences* **232**, 57–65.

Diekmann, O., Heesterbeek, J. A. P. & Roberts, M. G., 2010 The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society Interface* **7**, 873–885.

Dube, C., Ribble, C., Kelton, D. & McNab, B., 2008 Comparing network analysis measures to determine potential epidemic size of highly contagious exotic diseases in fragmented monthly networks of dairy cattle movements in Ontario, Canada. *Transboundary and Emerging Diseases* **55**, 382–392.

Duncan, A. J., Gunn, G. J., Lewis, F. I., Umstatter, C. & Humphry, R. W., 2012 The influence of empirical contact networks on modelling diseases in cattle. *Epidemics* **4**, 117–123.

Durham, D. P. & Casman, E. A., 2012 Incorporating individual health-protective decisions into disease transmission models : a mathematical framework. *Journal of the Royal Society Interface* **9**, 562–570.

Durrett, R., 2007 *Random graph dynamics*. Cambridge University Press.

Durrett, R. & Levin, S., 1994 The importance of being discrete (and spatial). *Theoretical Population Biology* **46**, 363–394.

Eames, K. T. & Keeling, M. J., 2003 Contact tracing and disease control. *Proceedings of the Royal Society B : Biological Sciences* **270**, 2565–71.

Eames, K. T. D., Read, J. M. & Edmunds, W. J., 2009 Epidemic prediction and control in weighted networks. *Epidemics* **1**, 70–76. Doi : DOI : 10.1016/j.epidem.2008.12.001.

Economic Sciences Prize Committee, 2010 Markets with search frictions. *Tech. rep.*, The Royal Swedish Academy of Sciences.

Edmunds, W., Kafatos, G., Wallinga, J. & Mossong, J., 2006 Mixing patterns and the spread of close-contact infectious diseases. *Emerging Themes in Epidemiology* **3**, 10.

Fair, R. C. & Jaffee, D. M., 1972 Methods of estimation for markets in disequilibrium. *Econometrica* **40**, 497–514.

Fenichel, E. P., Castillo-Chavez, C., Ceddia, M. G., Chowell, G., Gonzalez Parra, P. A., Hickling, G. J., Holloway, G., Horan, R., Morin, B., Perrings, C., Springborn, M., Velazquez, L. & Villalobos, C., 2011 Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences of the USA* **108**, 6306–6311.

Ferguson, N. M., Donnelly, C. A. & Anderson, R. M., 2001a The Foot-and-Mouth epidemic in Great Britain : pattern of spread and impact of interventions. *Science* **292**, 1155–1160.

Ferguson, N. M., Donnelly, C. A. & Anderson, R. M., 2001b Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature* **413**, 542–548.

Foucault, M., 1984 *Histoire de la sexualité*. Gallimard, Paris, France.

Funk, S., Bansal, S., Bauch, C. T., Eames, K. T. D., Edmunds, W. J., Galvani, A. P. & Klepac, P., 2014 Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics* **In press**.

Funk, S., Gilad, E., Watkins, C. & Jansen, V. A. A., 2009 The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences of the USA* **106**, 6872–6877.

Funk, S., Salathe, M. & Jansen, V. A. A., 2010 Modelling the influence of human behaviour on the spread of infectious diseases : a review. *Journal of the Royal Society Interface* **7**, 1247–1256.

Gog, J. R., Pellis, L., Wood, J. L. N., McLean, A. R., Arinaminpathy, N. & Lloyd-Smith, J. O., 2014 Seven challenges in modeling pathogen dynamics within-host and across scales. *Epidemics* **In press**.

Gohin, A., Cordier, J., Krebs, S. & Rault, A., 2013 Dynamic impacts of a catastrophic production event : The foot-and-mouth disease case. *Risk Analysis* **33**, 480–492.

Gomez-Gardenes, J., Latora, V., Moreno, Y. & Profumo, E., 2008 Spreading of sexually transmitted diseases in heterosexual populations. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1399–1404.

Goyal, S., 2009 *Connections : an introduction to the economics of networks*. Princeton University Press.

Grenfell, B. T. & Dobson, A. P., 1995 *Ecology of infectious diseases in natural populations*. Publications of the Newton Institute 7. Cambridge ; New York : Cambridge University Press.

Gross, T., D'Lima, C. J. D. & Blasius, B., 2006 Epidemic dynamics on an adaptive network. *Physical Review Letters* **96**, 208701. PRL.

Grunwald, N. J., Garbelotto, M., Goss, E. M., Heungens, K. & Prospero, S., 2012 Emergence of the sudden oak death pathogen *Phytophthora ramorum*. *Trends in Microbiology* **20**, 131–138.

Guesnerie, R., 1996 *L'économie de marché*. Flammarion.

Hahn, F., 1982 *Stability*, vol. 2, chap. 16, pp. 745–793. Elsevier.

Hahn, F. H. & Negishi, T., 1962 A theorem on non-tatonnement stability. *Econometrica* pp. 463–469.

Haldane, A. G. & May, R. M., 2011 Systemic risk in banking ecosystems. *Nature* **469**, 351–355.

Hamilton, D. T., Handcock, M. S. & Morris, M., 2008 Degree distributions in sexual networks : a framework for evaluating evidence. *Sexually Transmitted Diseases* **35**, 30–40.

Henderson, V. & Hobson, D. G., 2002 Real options with constant relative risk aversion. *Journal of Economic Dynamics and Control* **27**, 329–355.

Hirsch, M. W., Smale, S. & Devaney, R. L., 2004 *Differential equations, dynamical systems and an introduction to chaos*, vol. 60 of *Pure and applied mathematics*. Elsevier Academic Press.

Hufnagel, L., Brockmann, D. & Geisel, T., 2004 Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15124–15129.

Jeger, M. J., Pautasso, M., Holdenrieder, O. & Shaw, M. W., 2007 Modelling disease spread and control in networks : implications for plant sciences. *New Phytologist* **174**, 279–97.

Johnson, A. M., Mercer, C. H., Erens, B., Copas, A. J., McManus, S., Wellings, K., Fenton, K. A., Korovessis, C., Macdowall, W., Nanchahal, K., Purdon, S. & Field, J., 2001 Sexual behaviour in Britain : partnerships, practices, and HIV risk behaviours. *The Lancet* **358**, 1835–1842.

Jones, B. A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M. Y., McKeever, D., Mutua, F., Young, J., McDermott, J. & Pfeiffer, D. U., 2013 Zoonosis emergence linked to agricultural intensification and environmental change. *Proceedings of the National Academy of Sciences* **110**, 8399–8404.

Joo, J. & Lebowitz, J. L., 2004 Behavior of susceptible-infected-susceptible epidemics on heterogeneous networks with saturation. *Physical Review E* **69**, 066105. 15244665.

Kamp, C., 2010 Untangling the interplay between epidemic spread and transmission network dynamics. *PLoS Computational Biology* **6**, e1000984.

Kamp, C., Moslonka-Lefebvre, M. & Alizon, S., 2013 Epidemic spread on weighted networks. *PLoS Computational Biology* **9**, e1003352.

Katzner, D., 2010 The current non-status of general equilibrium theory. *Review of Economic Design* **14**, 203–219.

Keeling, M. J., Danon, L., Vernon, M. C. & House, T. A., 2010 Individual identity and movement networks for disease metapopulations. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8866–8870.

Keeling, M. J. & Eames, K. T., 2005 Networks and epidemic models. *Journal of the Royal Society Interface* **2**, 295–307.

Keeling, M. J. & Rohani, P., 2008 *Modeling infectious diseases in humans and animals*. Princeton : Princeton University Press.

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., Cornell, S. J., Kappey, J., Wilesmith, J. & Grenfell, B. T., 2001 Dynamics of the 2001 uk foot and mouth epidemic : Stochastic dispersal in a heterogeneous landscape. *Science* **294**, 813–817.

Kellermann, K., 2011 Too big to fail : a thorn in the side of free markets. *Empirica* pp. 1–19.

Kermack, W. & McKendrick, A., 1927 A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences* **115**, 700–21.

Khan, A. & Thomas, J., 2003 Inventories and the business cycle : an equilibrium analysis of (S, s) policies. *Tech. rep.*, National Bureau of Economic Research.

Kiss, I. Z., Green, D. M. & Kao, R. R., 2006 The network of sheep movements within Great Britain : network properties and their implications for infectious disease spread. *Journal of the Royal Society Interface* **3**, 669–677.

Kitti, M., 2010 Convergence of iterative tatonnement without price normalization. *Journal of Economic Dynamics and Control* **34**, 1077–1091.

Klaes, M., 2008 *History of transaction costs*. Palgrave Macmillan.

Klein, E., Laxminarayan, R., Smith, D. L. & Gilligan, C. A., 2007 Economic incentives and mathematical models of disease. *Environment and Development Economics* **12**, 707–732.

Klepac, P., Metcalf, C. J. E., McLean, A. R. & Hampson, K., 2013 Towards the endgame and beyond : complexities and challenges for the elimination of infectious diseases. *Philosophical Transactions of the Royal Society B : Biological Sciences* **368**.

Klos, T. B. & Nooteboom, B., 2001 Agent-based computational transaction cost economics. *Journal of Economic Dynamics and Control* **25**, 503–526.

Konishi, H. & Fishburn, P., 1996 Quasi-linear utility in a discrete choice model. *Economics Letters* **51**, 197–200.

Kotler, P. & Armstrong, G., 2013 *Principles of Marketing 15th Global Edition*. Pearson.

Lagos, R., 2000 An alternative approach to search frictions. *Journal of Political Economy* **108**, 851–873.

Lal Dutta, B., Ezanno, P. & Vergu, E., 2014 Characteristics of the spatio-temporal networks of cattle movements in france over a 5-year period. *Preventive Veterinary Medicine* **117**, 79–94.

Lee, C.-F., Tsai, C.-M. & Lee, A. C., 2011 Asset pricing with disequilibrium price adjustment : theory and empirical evidence. *Quantitative Finance* **13**, 227–239.

Lentz, H. H. K., Konschake, M., Teske, K., Kasper, M., Rother, B., Carmanns, R., Petersen, B., Conraths, F. J. & Selhorst, T., 2011 Trade communities and their spatial patterns in the German pork production network. *Preventive Veterinary Medicine* **98**, 176–181.

Li, C., Van de Bovenkamp, R. & Van Mieghem, P., 2012 Susceptible-infected-susceptible model : A comparison of N-intertwined and heterogeneous mean-field approximations. *Physical Review E* **86**, 026116.

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. & Aberg, Y., 2001 The web of human sexual contacts. *Nature* **411**, 907–908.

Lloyd, A. L. & May, R. M., 2001 How viruses spread among computers and people. *Science* **292**, 1316–1317.

MAAF, 2013 Maladies animales : la tuberculose bovine (http ://agriculture.gouv.fr/maladies-animales-la-tuberculose-bovine).

Madar, N., Kalisky, T., Cohen, R., Ben-Avraham, D. & Havlin, S., 2004 Immunization and epidemic dynamics in complex networks. *The European Physical Journal B* **38**, 269–276.

Mas-Colell, A., Whinston, M. D. & Green, J. R., 1995 *Microeconomic theory*. New York : Oxford University Press.

Matthews, L., Reeve, R., Gally, D. L., Low, J. C., Woolhouse, M. E. J., McAteer, S. P., Locking, M. E., Chase-Topping, M. E., Haydon, D. T., Allison, L. J., Hanson, M. F., Gunn, G. J. & Reid, S. W. J., 2013 Predicting the public health benefit of vaccinating cattle against escherichia coli o157. *Proceedings of the National Academy of Sciences* **110**, 16265–16270.

May, R. M., 1977 Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* **269**, 471–477.

May, R. M. & Anderson, R. M., 1987 Transmission dynamics of HIV infection. *Nature* **326**, 137–42.

May, R. M. & Arinaminpathy, N., 2010 Systemic risk : the dynamics of model banking systems. *Journal of the Royal Society Interface* **7**, 823–838.

May, R. M., Levin, S. A. & Sugihara, G., 2008 Complex systems : ecology for bankers. *Nature* **451**, 893–895.

McCauley, J. L., 2009 *Dynamics of markets : the new financial economics*. Cambridge University Press.

Meyers, L. A., Newman, M. E. J. & Pourbohloul, B., 2006 Predicting epidemics on directed contact networks. *Journal of Theoretical Biology* **240**, 400–418.

Mideo, N., Alizon, S. & Day, T., 2008 Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends in Ecology and Evolution* **23**, 511–517.

Molloy, M. & Reed, B., 1995 A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms* **6**, 161–180.

Moreno, Y., Pastor-Satorras, R. & Vespignani, A., 2002 Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B - Condensed Matter and Complex Systems* **26**, 521–529.

Morin, B. R., Fenichel, E. P. & Castillo-Chavez, C., 2013 SIR dynamics with economically driven contact rates. *Natural Resource Modeling* **In press**.

Mortensen, D. T. & Pissarides, C. A., 1994 Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies* **61**, 397–415.

Moslonka-Lefebvre, M., 2014 Modelling epidemic dynamics in exchange networks with constrained interactions and adaptive behaviour. *PhD thesis, AgroParisTech ABIES* .

Moslonka-Lefebvre, M., Bonhoeffer, S. & Alizon, S., 2012a Weighting for sex acts to understand the spread of STI on networks. *Journal of Theoretical Biology* **311**, 46–53.

Moslonka-Lefebvre, M., Finley, A., Dorigatti, I., Dehnen-Schmutz, K., Harwood, T., Jeger, M. J., Xiangming, X., Holdenrieder, O. & Pautasso, M., 2011 Networks in plant epidemiology : from genes to landscapes, countries and continents. *Phytopathology* **101**, 392–403.

Moslonka-Lefebvre, M., Harwood, T., Jeger, M. J. & Pautasso, M., 2012b SIS along a continuum (SIS_c) epidemiological modelling and control of diseases on directed trade networks. *Mathematical Biosciences* **236**, 44–52.

Moslonka-Lefebvre, M., Monod, H., Gilligan, C., Vergu, E. & Filipe, J. A. N., 2013 Epidemics in markets with trade friction and imperfect transactions. *Preprint : arXiv :1310.6320* .

Moslonka-Lefebvre, M., Pautasso, M. & Jeger, M. J., 2009 Disease spread in small-size directed networks : epidemic threshold, correlation between links to and from nodes, and clustering. *Journal of Theoretical Biology* **260**, 402–411.

Mourits, M. C. M., van Asseldonk, M. A. P. M. & Huirne, R. B. M., 2010 Multi criteria decision making to evaluate control strategies of contagious animal diseases. *Preventive Veterinary Medicine* **96**, 201–210.

National Audit Office, 2002 The 2001 outbreak of Foot and Mouth Disease – Report by the Comptroller and Auditor General. *HC 939 Session 2001-2002* .

Newman, M. E. J., 2002 Spread of epidemic disease on networks. *Physical Review E* **66**, 016128.

Newman, M. E. J., 2003 The structure and function of complex networks. *SIAM Review* **45**, 167–256.

Newman, M. E. J., Strogatz, S. H. & Watts, D. J., 2001 Random graphs with arbitrary degree distributions and their applications. *Physical Review E* **64**, 026118.

Nicolaides, C., Cueto-Felgueroso, L. & Juanes, R., 2013 The price of anarchy in mobility-driven contagion dynamics. *Journal of the Royal Society Interface* **10**.

Nordvik, M. K. & Liljeros, F., 2006 Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. *Sexually Transmitted Diseases* **33**, 342–349.

Noremark, M., Hakansson, N., Lewerin, S. S., Lindberg, A. & Jonsson, A., 2011 Network analysis of cattle and pig movements in sweden : measures relevant for disease control and risk based surveillance. *Preventive Veterinary Medicine* **99**, 78–90.

Noremark, M. & Widgren, S., 2014 EpiContactTrace : an R-package for contact tracing during livestock disease outbreaks and for risk-based surveillance. *BMC Veterinary Research* **10**, 71.

Nowak, M. A., 2006 *Evolutionary dynamics*. Harvard University Press.

Orzack, S. H. & Sober, E., 1993 A critical-assessment of Levins's *The strategy of model-building in population biology* (1966). *Quarterly Review of Biology* **68**, 533–546.

Park, S. M. & Kim, B. J., 2006 Dynamic behaviors in directed networks. *Physical Review E* **74**, 026114.

Pastor-Satorras, R. & Vespignani, A., 2001 Epidemic spreading in scale-free networks. *Physical Review Letters* **86**, 3200.

Paton, D. J., Sumption, K. J. & Charleston, B., 2009 Options for control of foot-and-mouth disease : knowledge, capability and policy. *Philosophical Transactions of the Royal Society B : Biological Sciences* **364**, 2657–2667.

Pautasso, M. & Jeger, M. J., 2008 Epidemic threshold and network structure : the interplay of probability of transmission and of persistence in small-size directed networks. *Ecological Complexity* **5**, 1–8.

Pautasso, M., Xu, X., Jeger, M. J., Harwood, T. D., Moslonka-Lefebvre, M. & Pellis, L., 2010 Disease spread in small-size directed trade networks : the role of hierarchical categories. *Journal of Applied Ecology* **47**, 1300–1309.

Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V. & Trapman, P., 2014 Eight challenges for network epidemic models. *Epidemics* **In press**.

Perra, N., Goncalves, B., Pastor-Satorras, R. & Vespignani, A., 2012 Activity driven modeling of time varying networks. *Sci. Rep.* **2**.

Perrings, C., Castillo-Chavez, C., Chowell, G., Daszak, P., Fenichel, E. P., Finnoff, D., Horan, R. D., Kilpatrick, A. M., Kinzig, A., Kuminoff, N. V., Levin, S., Morin, B., Smith, K. F. & Springborn, M., 2014 Merging economics and epidemiology to improve the prediction and management of infectious disease. *EcoHealth* pp. 1–12.

Pindyck, R. S., 1982 Adjustment costs, uncertainty, and the behavior of the firm. *The American Economic Review* **72**, 415–427.

Pissarides, C., 2011 Equilibrium in the labor market with search frictions. *American Economic Review* **101**, 1092–1105.

Pissarides, C. A., 1985 Short-run equilibrium dynamics of unemployment, vacancies, and real wages. *The American Economic Review* **75**, 676–690.

Pissarides, C. A., 2001 *Economics of Search*, pp. 13760–13768. Oxford : Pergamon.

Quandt, R. E., 1988 *The econometrics of disequilibrium*. Blackwell Oxford.

Rattana, P., Blyuss, K. B., Eames, K. T. B. & Kiss, I. Z., 2013 A class of pairwise models for epidemic dynamics on weighted networks. *Bulletin of Mathematical Biology* **75**, 466–490.

Rautureau, S., Dufour, B. & Durand, B., 2011 Vulnerability of animal trade networks to the spread of infectious diseases : a methodological approach applied to evaluation and emergency control strategies in cattle, france, 2005. *Transboundary and Emerging Diseases* **58**, 110–120.

Rautureau, S., Dufour, B. & Durand, B., 2012 Structural vulnerability of the french swine industry trade network to the spread of infectious diseases. *Animal* **6**, 1152–1162.

Restrepo, J. G., Ott, E. & Hunt, B. R., 2007 Approximating the largest eigenvalue of network adjacency matrices. *Physical Review E* **76**, 056119.

Rocha, L. E. C., Liljeros, F. & Holme, P., 2011 Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Computational Biology* **7**, e1001109.

Rosenbaum, E. F., 2000 What is a market ? On the methodology of a contested concept. *Review of Social Economy* **58**, 455 – 482.

Rowthorn, R. E., Laxminarayan, R. & Gilligan, C. A., 2009 Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface* **6**, 1135–1144.

Schley, D., Whittle, S., Taylor, M. & Kiss, I. Z., 2012 Models to capture the potential for disease transmission in domestic sheep flocks. *Preventive Veterinary Medicine* **106**, 174–184.

Schneeberger, A., Mercer, C. H., Gregson, S. A. J., Ferguson, N. M., Nyamukapa, C. A., Anderson, R. M., Johnson, A. M. & Garnett, G. P., 2004 Scale-free networks and sexually transmitted diseases : a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sexually Transmitted Diseases* **31**, 380–387.

Schumm, P., 2013 Characterizing epidemics in metapopulation cattle systems through analytic models and estimation methods for data-driven model inputs. Ph.D. thesis.

Serrano, M. A., Boguna, M. & Pastor-Satorras, R., 2006 Correlations in weighted networks. *Physical Review E* **74**, 055101. PRE.

Silvestre, J., 2008 *Fixprice models*. Palgrave Macmillan.

Singh, N. & Vives, X., 1984 Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics* pp. 546–554.

Sissors, J. Z., 1966 What is a market ? *The Journal of Marketing* **30**, 17–21.

Smith, T. W., 1992 Discrepancies between men and women in reporting number of sexual partners : A summary from four countries. *Biodemography and Social Biology* **39**, 203–211.

Sokolowski, M. B., Pereira, H. S. & Hughes, K., 1997 Evolution of foraging behavior in drosophila by density-dependent selection. *Proceedings of the National Academy of Sciences* **94**, 7373–7377.

Soliman, T., Mourits, M. C. M., Oude Lansink, A. G. J. M. & van der Werf, W., 2010 Economic impact assessment in pest risk analysis. *Crop Protection* **29**, 517–524.

Stehle, J., Voirin, N., Barrat, A., Cattuto, C., Colizza, V., Isella, L., Regis, C., Pinton, J.-F., Khanafer, N., Van den Broeck, W. & Vanhems, P., 2011 Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine* **9**, 87.

Stevens, M., 2007 New microfoundations for the aggregate matching function. *International Economic Review* **48**, 847–868.

Strogatz, S., 2001 Exploring complex networks. *Nature* **410**, 268–276.

Sunnaker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M. & Dessimoz, C., 2013 Approximate bayesian computation. *PLoS Comput Biol* **9**, e1002803.

Tatem, A. J., Rogers, D. J. & Hay, S. I., 2006 *Global Transport Networks and Infectious Disease Spread*, vol. Volume 62, pp. 293–343. Academic Press.

Tesfatsion, L., 2001 Introduction to the special issue on agent-based computational economics. *Journal of Economic Dynamics and Control* **25**, 281–293.

Valdano, E., Poletto, C., Giovannini, A., Palma, D., Savini, L. & Colizza, V., 2014 Predicting epidemic risk from past temporal contact data. *arXiv :1406.1449* .

Vernon, M., 2011 Demographics of cattle movements in the united kingdom. *BMC Veterinary Research* **7**, 31.

Vernon, M. C. & Keeling, M. J., 2009 Representing the uk's cattle herd as static and dynamic networks. *Proceedings of the Royal Society B : Biological Sciences* **276**, 469–476.

Volz, E., 2008 SIR dynamics in random networks with heterogeneous connectivity. *Journal of Mathematical Biology* **56**, 293–310.

Volz, E. & Meyers, L., 2009 Epidemic thresholds in dynamic contact networks. *Journal of the Royal Society Interface* **6**, 233–241.

Volz, E. & Meyers, L. A., 2007 Susceptible-infected-recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society B : Biological Sciences* **274**, 2925–2933.

Wallinga, J., Teunis, P. & Kretzschmar, M., 2006 Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology* **164**, 936–944.

Wang, H. Z., Liu, Z. R. & Xu, H. H., 2007 Epidemic spreading on uncorrelated heterogenous networks with non-uniform transmission. *Physica A* **382**, 715–721.

Williamson, O. E. & Masten, S. E., 1999 *The economics of transaction costs*. Edward Elgar Publishing.

Yang, Z. & Zhou, T., 2012 Epidemic spreading in weighted networks : an edge-based mean-field solution. *Physical Review E* **85**, 056106.

Zhao, Z., Wahl, T. I. & Marsh, T. L., 2006 Invasive species management : Foot-and-Mouth Disease in the US beef industry. *Agricultural and Resource Economics Review* **35**, 98–115.

Annexes

Annexe A

Research paper : *Weighting for sex acts to understand the spread of STI on networks*

Journal of Theoretical Biology 311 (2012) 46–53

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Weighting for sex acts to understand the spread of STI on networks

Mathieu Moslonka-Lefebvre [a,b,*], Sebastian Bonhoeffer [a], Samuel Alizon [c]

[a] Institute for Integrative Biology, ETH, Universitätstrasse 16, 8092 Zürich, Switzerland
[b] INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France
[c] Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2), 911 Avenue Agropolis, 34394 Montpellier Cedex 5, France

HIGHLIGHTS

- Human sexual networks exhibit a heterogeneous structure.
- Expected network-based properties are at odds with epidemiological data for STI.
- Most network models typically assume a constant transmission risk per partnership.
- We develop a framework to weight sexual networks based on biological assumptions.
- Our weighting model re-conciliate network theory and epidemiological data for STI.

ARTICLE INFO

Article history:
Received 20 February 2012
Received in revised form
21 June 2012
Accepted 25 June 2012
Available online 3 July 2012

Keywords:
Weighted networks
Epidemiology
STI
Heterogeneity
Sex acts

ABSTRACT

Human sexual networks exhibit a heterogeneous structure where few individuals have many partners and many individuals have few partners. Network theory predicts that the spread of sexually transmitted infections (STI) on such networks should exhibit striking properties (e.g. rapid spread). However, these properties cannot be found in epidemiological data. Current network models typically assume a constant STI transmission risk per partnership, which is unrealistic because it implies that sexual activity is proportional to the number of partners and that individuals have the same activity with each partner. We develop a framework that allows us to weight any sexual network based on biological assumptions. Our results indicate that STI spreading on the resulting weighted networks do not have heterogeneous-related properties, which is consistent with data and earlier studies.

1. Introduction

Networks reflecting human sexual contacts over several years in a population have been shown to exhibit a high variance in the number of sexual partners per individual (Liljeros et al., 2001; Hamilton et al., 2008). The exact nature of the distribution of partners per individual is debated: some argue that it tends to follow a power-law (PL; Liljeros et al., 2001; Schneeberger et al., 2004; Hamilton et al., 2008), others argue that it tends to follow a negative-binomial (NB; Handcock and Jones, 2004; Hamilton et al., 2008). At any rate, both these distributions are the appanage of highly heterogenous networks in which a few nodes in a network (i.e. individuals) are highly connected (i.e. have many partners), while the others nodes have few connections. Theory predicts that epidemics spreading on heterogeneous networks (either PL or NB) should exhibit specific properties such as a very low epidemic threshold, a rapid spread or in the case of PL, a doubling time that decreases with population size (Newman, 2003; Keeling and Eames, 2005). There is an ongoing debate in the literature concerning the relevance of these predictions for biological systems (especially the absence of epidemic threshold; May, 2006; Hamilton et al., 2008). For instance, specific epidemiological properties of highly heterogeneous networks are not observed generally in the data (Liljeros, 2004; Handcock and Jones, 2006; Britton et al., 2007; Hamilton et al., 2008).

Most network-based epidemiological studies consider unweighted networks. In the context of sexually transmitted infections (STI), it means that the risk of infection is the same on all the edges (i.e. for all the interactions) in the network (but see Eames et al., 2009). While these assumptions may be justifiable for diseases such as flu or tuberculosis that only require short-term interactions between hosts, they are unrealistic for STI. Here, we develop a new framework to add sex acts explicitly on sexual contact networks and show that such a biologically relevant weighting has major epidemiological consequences.

* Corresponding author at: INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France. Tel.: +33 1 34 65 22 47.
E-mail address: mathieu.moslonka-lefebvre@jouy.inra.fr (M. Moslonka-Lefebvre).

0022-5193/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtbi.2012.06.031

Modelling the spread of STI on unweighted networks makes two implicit assumptions. First, the total number of sex acts of an individual per unit of time is assumed to be strictly proportional to his/her number of sexual partners. Second, these sex acts are assumed to be partitioned equally among all the partners. That individuals with five sexual partners generally have five times more sex acts per unit of time than individuals with one partner is unrealistic and contradicts empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006; Britton et al., 2007).

Several studies point out the importance of weighted networks in general (see e.g. Barrat et al., 2004b's seminal work). Some studies consider the spread of STIs on weighted networks of infinite size. Newman (2002) already showed that assuming that the transmission rate from one infected individual to another is not constant but rather a function of node degree (i.e. number of partners) can affect disease spread. More precisely, he shows that for epidemics to spread rates of transmission need to fall off slower than inversely with node degree (independently of the network structure). Wang et al. (2007) approximate an infinite size so-called 'scale free' network with mean field equations and show that assuming that the transmission from one node to another is proportional to the node degree affects epidemic spreading. Joo and Lebowitz (2004) use a similar approach but consider a more elaborate transmission function, which they allow to saturate with increasing node degree. They find that such a saturation further decreases the speed of disease spread. Recently, Britton et al. (2011) study infinite size weighted sexual contact networks, where weights are drawn in a distribution that is allowed to depend on node degree. They show that if nodes (vertices) with high degree tend to have a low weight, then it is harder for an epidemics to take off. The two main limitations of these models are that they consider infinite size networks and/or that they do not consider any biological basis for the weighting they use. Note also that in most of these formalisms, the transmission rate from an individual A to an individual B differs from the transmission rate from B to A if they have different number of partners. In Britton et al. (2011) the assumption is slightly different because a node can only be linked to a node that has the same weight (which means the weighting constraints the shape of the network).

More biologically oriented approaches develop individual-based models to understand how sexual contact networks emerge (see e.g. Althaus et al., 2010) or pair-approximation models but without taking into account sex acts (Ferguson and Garnett, 2000). Among these studies, two consider the effect of sexual activity on the spread of an STI more explicitly. Röttingen and Garnett (2002) study the association between STI risk and number of sex acts without a network-based approach by fitting the relationship between the HIV transmission probability per-partnership and the number of sex acts. However, their model ignores epidemiological feedbacks, thus assuming for instance that all sexual partners of one individual have the same probability to be infected (whereas this probability should at least depend on the number of partner these partners have). More recently, Britton et al. (2007) model the spread of an STI in a heterogeneous population (in terms of number of sexual partners) where individual of each class can have both 'steady' or 'casual' partnerships. The latter partnerships have higher transmission probabilities than the former and the ratio of each type of partnership is inferred from empirical data on the number of sex acts per partner. They show that the basic reproduction number (R_0) can be over-estimated when different types of partnerships are not considered.

For sexual contact networks, the weight of an edge between two nodes corresponds to the number of sex acts that are actually realised between two individuals per unit of time (here weeks). In the following, we refer to these as realised sex acts (RSA). We derive the number of RSA between any interacting pair (i.e. all the weights of the network edges) from the network topology based on two simple biological assumptions. First, each individual is allocated a number of potential sex acts (PSA). PSA can be seen as a quantification of the sexual activity of an individual. Second, each individual partitions his/her PSA among his/her sexual partners. The number of RSA between two individuals i and j depends on the number of PSA that i attributed to j and vice versa (see the Methods section).

We consider three models that describe the link between the number of PSA allocated to an individual and the number of partners of this individual: a linear relationship (the linear allocation model A_{lin}), a saturating relationship (the saturating allocation model A_{sat}), or a constant relationship where each individual receive the same number of PSA (the constant allocation model, A_{cst}). Further details about the model are available in the Material and Methods section. Classical models implicitly make the strong assumption that PSA are allocated linearly with the number of sex partners, which is at odds with empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006).

In our study, PSA can then be partitioned by an individual among his/her partners in three ways: (i) equally (P_{equ}), where all the partners get the same share, (ii) randomly (P_{rand}), where the fraction each partner gets is random and (iii) to maximise his/her number of RSA (P_{max}). Classical models implicitly assume an equal or a maximising partitioning, but the random partitioning arguably better reflects current sexuality, where the majority does not act in terms of maximising sex budgets (Foucault, 1976).

Our framework stands out compared to earlier studies because we use a finite size network inferred from contact tracing data, contrary to earlier models that are based on infinite 'scale-free' networks (Joo and Lebowitz, 2004; Wang et al., 2007) or that use degree distributions without explicit network topology (Röttingen and Garnett, 2002; Britton et al., 2007). This allows us to model explicitly how individual sexual behaviours affect the network weighted topology and, hence disease spread. Earlier studies that account for sex acts either do not include epidemiological feedbacks (Röttingen and Garnett, 2002) or only model a heterogenous population with two types of partnerships and therefore cannot keep track of individual behaviours (Britton et al., 2007). Our framework also allows us to model sex acts explicitly on any network type. Overall, we show that minimal deviations from implicit assumptions regarding sexual behaviour made in earlier models can strongly affect diseases spread.

2. The model

2.1. The network

A sexual network is a representation of individuals (the network nodes) and sexual relations between these individuals (the network edges). Each individual i has k_i sexual partners. Sexual networks evolve through time by creation and removal of sexual relationships among partners. However, the networks we usually have access to summarise sexual contacts over long periods of time because, in most studies, participants are asked to list their sexual partners over several years (see e.g. Potterat et al., 1985; Liljeros, 2004). Our method of weighting allows the modelling of epidemics occurring on a short time scale (days), even with a contact network inferred over a longer time scale (years).

For simplicity, we focus on disease spread on static networks of men having sex with men (MSM), even if our methods are also applicable to dynamical and bipartite (heterosexual) networks. Such networks can be described using an adjacency matrix ($\mathcal{A}$),

the elements of which ($\mathcal{A}(i,j)$) are equal to 1 if there is a link connecting node i to node j and 0 otherwise.

Except when stated otherwise, the results we present are obtained on an empirical network of 381 nodes built from a contact tracing study in Colorado Springs involving a cohort of individuals infected with gonorrhoea (Potterat et al., 1985). The average degree of the network (i.e. number of sexual partners) is 2. As shown in the Supplementary Methods, the degree distribution of this network is best described by a power law of coefficient $\gamma = 1.98$.

We also generated theoretical power-law (PL) and negative-binomial (NB) networks of different sizes with the Barabasi–Albert algorithm with a γ coefficient of 3 (Barabasi and Albert, 1999) and the so-called 'configuration model' (Catanzaro et al., 2005) respectively. For both PL and NB, we set the average degree to 4 (see Supplementary Methods). This allowed us to check that our results were not due to the specific topology of the empirical network.

2.2. Weighting the network

Each individual i is allocated a number of potential sex acts (PSA) per unit of time (denoted b_i). This individual then partitions his/her PSA among his/her k_i sexual partners. It is thus possible to define the availability of an individual i for interacting with j, which is the number of PSA allocated by i to j (denoted $[i \to j]$).

2.2.1. Allocation of sex acts

We consider three models as to how the number of PSA (b_i) is allocated to an individual i:

- The linear allocation model (A_{lin}), which assumes that b_i is strictly proportional to k_i, the number of partners.
- The saturating allocation model (A_{sat}), which assumes that b_i is proportional to $\sqrt{k_i}$.
- The constant allocation model (A_{cst}), which assumes that b_i is the same for any node of the network.

2.2.2. Partitioning of sex acts

PSA are then partitioned by an individual i among his/her partners according to three models:

- The equal partitioning model (P_{equ}), where all the partners of an individual i get an equal share $[i \to j] = b_i/k_i$.
- The random partitioning model (P_{rand}), where all the partners of an individual i get a share $[i \to j] = b_i\, \rho_j/R$, where ρ_j is drawn in a uniform distribution and $R = \sum_{j=1}^{k_i} \rho_j$.
- The maximisation partitioning model ($P_{\max}$), in which an individual uses information about the topology to maximise their number of realised sex acts by giving to his/her partners in proportion to what theses partners are giving to him (see the Supplementary Methods for further details).

STI models with an unweighted network implicitly assume a linear allocation model with an equal or a maximising partitioning.

2.2.3. Realised sex acts

The number of sex acts actually realised (RSA) between individuals i and j (denoted $W(i,j)$) is then assumed to be

$$W(i,j) = \min([i \to j],[j \to i]) \tag{1}$$

because an individual i cannot have more sex acts with a partner j than the number of sex acts that j is willing to have with i and vice versa. From this, we obviously have $W(i,j) = W(j,i)$.

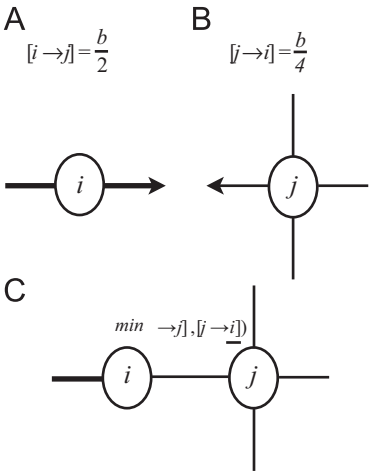


Fig. 1. Illustration of our weighting model. In this example, individuals i (A) and j (B) are allocated a constant sex budget (set to $b=4$ sex acts per week) that they partition equally among their partners. The number of sex acts actually realised (RSA) between individuals i and j is the minimum of the two contributions (C).

Our weighting framework (Fig. 1) thus allows us to transform a classical adjacency matrix ($\mathcal{A}$) into a weighted adjacency matrix (W), the elements of which are the realised number of sex acts (as defined in Eq. (1)).

2.2.4. Loss of sex acts due to the weighting

Care is needed because, even though our definition on how sex acts are realised is biologically sensible, it introduces a potential confounding factor because the more $[i \to j]$ and $[j \to i]$ differ, the more PSA will be lost. We quantify this loss of sex acts on the whole network with a parameter δ defined as

$$\delta = \frac{\sum_i \sum_j W(i,j)}{\sum_i \sum_j s^* \mathcal{A}(i,j)} \tag{2}$$

where the $W(i,j)$ are the elements of the weighted adjacency matrix, $\mathcal{A}(i,j)$ are the elements of the adjacency matrix and s^* is the average number of potential sex acts between two individuals.

We want to compare the spread of a pathogen on different weighted networks and this spread is obviously a function of the total number of sex acts realised on the network. In order to isolate the effect of the network itself, we define the weights on the network such that the average number of sex acts realised per individual per unit of time on a network is equal to the average degree of this network ($s^* = \overline{k} = 4$). We hence calculate a corrected weighted adjacency matrix $\tilde{W} = W/\delta$. Note that for the genuine GonoCocci network, denoted the GC network throughout, $\overline{k}_{GC} = 2$ so in order to get values comparable to the other networks we corrected the W for GC by $2\overline{k}_{GC}$. In summary, the total number of realised sex acts is constant for any network structure, allocation model and partitioning model.

2.3. The epidemiological model

Since our goal is to investigate the importance of weighting the network, we deliberately consider a simple epidemiological SI model, where hosts can be either susceptible or infected (Anderson and May, 1991). When an infected individual has sex with a susceptible one, the disease is assumed to be transmitted at a constant rate per sex act. We further assume that there is no recovery and that death does not occur over the duration of the simulation.

This simple model is a reasonably good approximation to the transmission of HIV over short periods of time (Liljeros, 2004), as there is no significant removal of individuals due to death on timescales close to the average duration of an HIV infection if untreated ($\approx$10 year; Anderson and May, 1991). Furthermore, transmission rate per sex acts was set to $\beta = 0.01$, which is consistent with HIV data (Vittinghoff et al., 1999; Boily et al., 2009). The speed of disease spread is estimated through the median doubling time t_d. The doubling time is an appropriate summary of disease dynamics for a pathogen such as HIV at the beginning of an outbreak (May et al., 2001). Some results can be inferred directly from the network topology but this is beyond the scope of this study.

In the case of PL, the basic reproduction number, denoted R_0, increases with network size (Farkas et al., 2001). As R_0 and doubling time are negatively correlated in the early stages of an epidemic (May et al., 2001), doubling time is expected to decrease with population size (Newman, 2003; Keeling and Eames, 2005).

Further details are available in the Supplementary Methods of the Electronic Supplementary Material (ESM).

3. Results

We show that using biological assumptions to weight sexual contact networks strongly affects the epidemiology of STI. As the exact structure of these networks is the subject of an intense debate (Hamilton et al., 2008), we first present the results obtained on an empirical network (Potterat et al., 1985). However, our results hold when we consider theoretical finite size networks with PL and NB distributions.

3.1. Resulting realised number of sex acts

Our framework requires two biological assumptions to weight the network. The first assumption has to do with the allocation of the potential number of sex acts (one could think of this as the 'libido' of the individuals). The second has to do with the variance in the number of sex acts individuals have with each of their partners. For both these assumptions, existing data are very limited. However, one can evaluate the relevance of these assumptions by looking at the realised number of sex acts.

We find that for all allocations and partitioning, the total number of RSA and the total number of partners (degree) are positively correlated, as shown in Fig. 2. While the assumptions used in most models (A_{lin} with either $P_{\max}$ or P_{rand}) yields a linear relationship between RSA and degree, the constant case (A_{cst} and P_{rand}) is characterised by a saturating slope for large number of partners. The increasing but saturating relationship in realised sex acts (RSA) as a function of degree observed in the case with constant allocation of sex acts is apparently counterintuitive. This paradox is explained by the fact that high degree nodes are less likely to loose potential sex acts (PSA) than low degrees nodes: even though the PSA are constant, a high-degree node is more likely to find partners to satisfy all its PSA whereas a low-degree node is less likely to find partners to use them. Note that the correction of the weighting by delta to keep the average number of RSA constant and equal to the average number of PSA is applied with the same intensity to each node of the corrected network: it is hence not changing the shape of the relationship between RSA (prior or after correction) and degree.

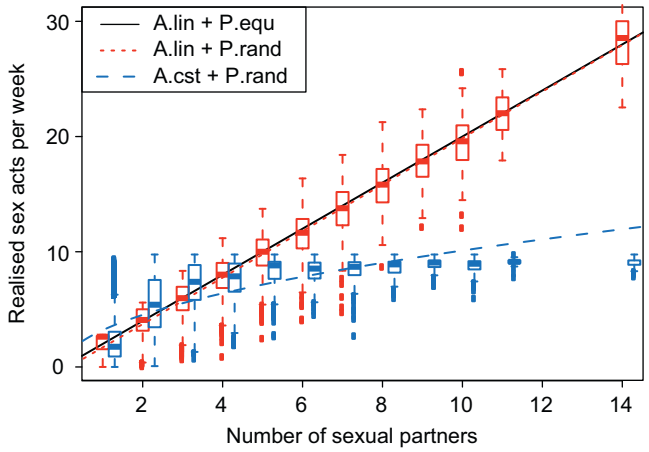


Fig. 2. Relationships between realised sex acts and number of partners for the empirical network. These curves show median values for 1000 stochastic SI simulations on the empirical network, where $\beta = 0.01$ and individuals have on average four sex acts per week. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Empirical studies based on wide surveys tend to find that number of sex acts tend to increase with the number of partners, but that this increase saturates and is clearly not strictly proportional to the number of partners (Blower and Boe, 1993; Nordvik and Liljeros, 2006; Britton et al., 2007). This suggests that saturating or constant allocation models for PSA might be more realistic than the linear allocation, which is often implicitly assumed.

As explained in the Model section, any discrepancy in the total number of sex acts is corrected for in the following analyses. More precisely, we use δ to normalise the edge weights and thus keep the average number of RSA per individual constant for *all* the networks that we consider. Therefore, any change we observe in disease spread are only due to the weighted topology.

3.2. Disease spread on the empirical network

Fig. 3 shows the effect of biological assumptions on the topology of the network and how this affects disease spread. The 'classical' approach assumes the linear (A_{lin}) model for PSA allocation with equal (P_{equ}) or maximising ($P_{\max}$) partitioning. As a consequence, all the interactions have the same weight and the number of RSA increases proportionally with the number of sexual partners (Fig. 3A). Assuming that the partitioning of the PSA among sexual partners is random (P_{rand}) generates variance in the weights of the edges (Fig. 3B). If we further assume that all the individuals are allocated the same number of potential sex acts (A_{cst}), we observe drastic changes. First, the variance in edge weights increases even more. Second, there is a complete shift in terms of sexual activities: nodes with few neighbours, whose activity was negligible with the classical approach, become important at the expenses of super-connected nodes (Fig. 3C). Fig. 3D illustrates the consequences of these assumptions on the spread of an STI (see the Methods for a description of the epidemiological model used). Assuming a P_{rand} partitioning model instead of a P_{equ} model slows the spread of the disease by more than 50% and assuming an A_{cst} allocation model instead of an A_{lin} model slows it by an extra 50%. Note that the longer the epidemics runs, the less realistic some of our model assumptions become (e.g. in the case of HIV, infected individuals are going to die).

To get a better understanding of these results, we analyse key variables on the same empirical network (Table 1). One of these variables is the median doubling time (t_d) of an STI. t_d estimates the speed of spread of a disease and we define it as the time required in stochastic simulations for the disease prevalence to go from 1% to 2% of the host population. We find that the median doubling time is significantly higher with a constant allocation model or with a random partitioning, when compared to the

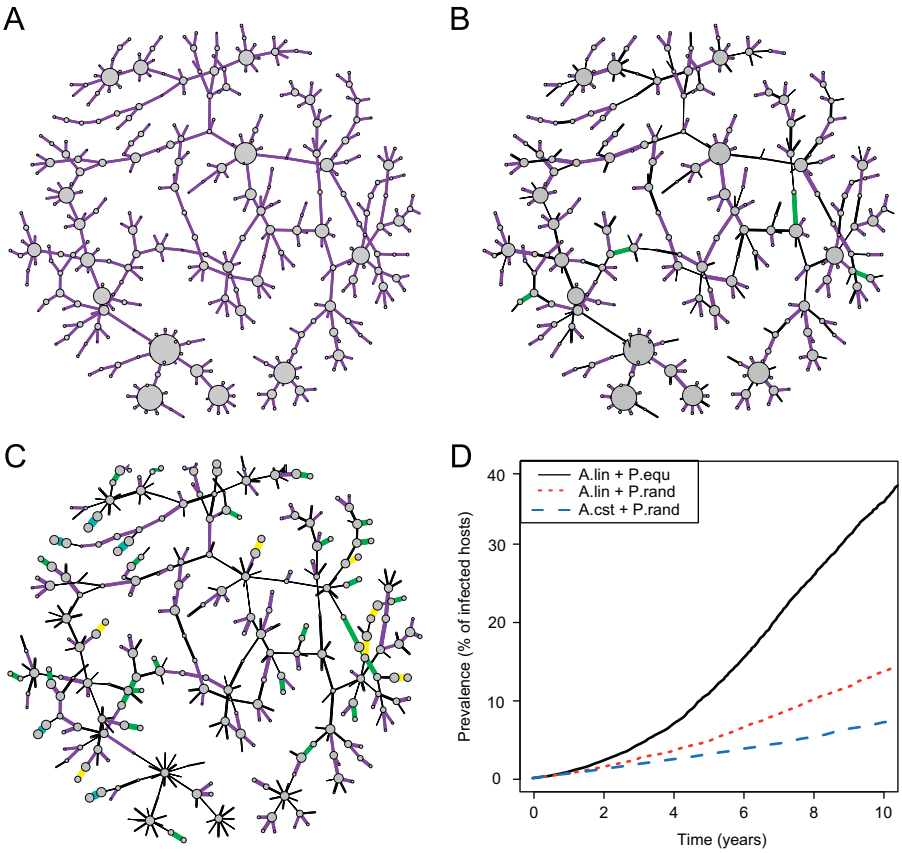


Fig. 3. Topological and epidemiological consequences of weighting the empirical network with realised sex acts. (A) Empirical contact network weighted with the number of realised sex acts implicitly assumed by classical models (A_{lin} and P_{equ}), (B) same as A but with a random partitioning, (C) same as A but with a constant allocation and a random partitioning. In panels A, B and C node diameter (edge width respectively) represents the number of realised sex acts per time step by an individual (the number of realised sex acts per time step by a couple respectively). Edge colour indicates less than 2 (in black), between 2 and 4 (purple), between 4 and 6 (green), between 6 and 8 (yellow) and more than 8 (cyan) sex acts per week. (D) Median STI prevalence as a function of time with the classical approach (in black), with a random partitioning (in red) and with a constant PSA allocation and a random partitioning (in blue). Notations and methods are as in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Effect of weighting the empirical network on disease spread. Each individual is allocated a number of PSA (lines) which can be partitioned in three different ways (columns). For each combination, the table shows the median doubling time (t_d, in weeks) with first and third quartile and δ, which is the ratio between the total number of realised sex acts (RSA) and the total number of potential sex acts (PSA) on the network. Edge weights are normalised by δ so that all the t_d are calculated on weighted networks with the same total number of RSA. We assume $\beta = 0.01$ and four RSA per week per individual on average. There were 1000 repetitions for each random partitioning. Further details are available in the Material and Methods and in the Supplementary Results.

Allocation	Estimator	**Partitioning**		
		P_{equ}	P_{max}	P_{rand}
A_{lin}	t_d	33 (19–57)	33 (19–57)	39 (22–79)
	δ	1.0	1.0	0.74 ± 0.01
A_{sat}	t_d	36 (23–61)	67 (37–130)	48 (28–92)
	δ	0.71	0.81	0.64 ± 0.01
A_{cst}	t_d	45 (27–76)	149 (72–404)	59 (34–109)
	δ	0.47	0.61	0.44 ± 0.01

classical approach (A_{lin} and P_{equ}). We also find an interaction between A_{cst} and P_{max}. Results obtained on 100 theoretical PL networks also show a significant effect of A_{sat} (see Supplementary Results).

As soon as we deviate from the classical approach (A_{lin} with P_{equ} or with P_{max}), the number of RSA over the whole network decreases. This is captured by δ, which is the ratio between the total number of RSA and the total number of PSA over the whole network prior correction of the weighting. This decrease comes from the fact that super-connected individuals cannot match the availability of all their partners. Note that we use δ to normalise the edge weights and thus keep the average number of RSA per individual constant for *all* the networks that we consider.

3.3. Results on theoretical networks and HIV data

For the sake of generality, we tested our results on theoretical heterogeneous network of different sizes assuming two types of partner distributions: power-law (PL) and negative-binomial (NB). Further details about the theoretical networks are available in the Supplementary Methods. In Fig. 4, we show the results for 100 PL and 100 NB networks (see Supplementary results for a negative control using homogeneous small-world networks). In all cases, we find that diseases spread faster on networks weighted with the 'classical' A_{lin} PSA allocation model than on networks weighted using the A_{sat} model, which themselves lead to a faster spread than networks weighted using the A_{cst} model.

One of the consequences of the weighting is that the doubling times estimated at the beginning of an epidemic (Table 1) and the number of sex acts per week needed to reproduce early HIV prevalence observed throughout the world are better explained by weighted than unweighted PL networks (both on empirical or theoretical networks). The doubling times we obtain from the simulated evolution of epidemic prevalence on realistically weighted networks are in the same order of magnitude as that

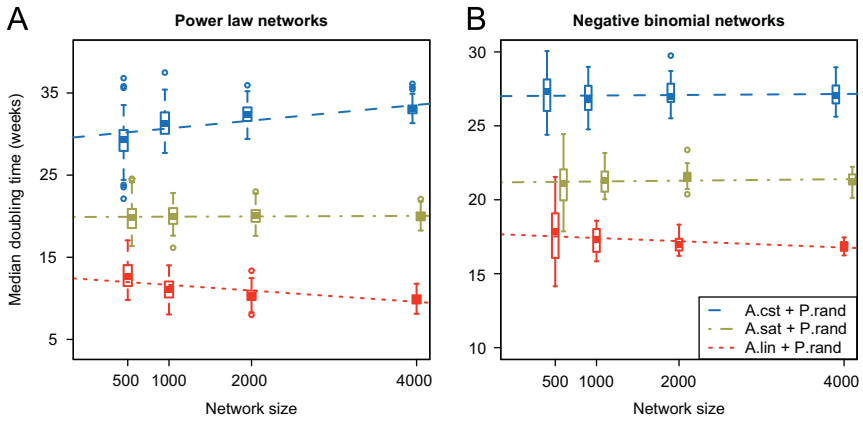


Fig. 4. Effect of weighting theoretical PL (A) and NB (B) networks of different sizes on disease spread. Disease spread is estimated through the median doubling time (t_d) computed over 100 runs of the simulation per replicate network. The box plot shows the median values over 100 different networks, the three quartiles and the outliers. Dashed lines are the output of a generalised linear model testing the effect of network size on t_d (see the main text). Parameters are as in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observed early in an HIV epidemic, i.e. approximately 1 year for men having sex with men (May and Anderson, 1987; Anderson and May, 1991). Note that the exact value of the doubling time is scaled by the product of the transmission rate per sex act (here set to $\beta = 0.01$; Vittinghoff et al., 1999; Boily et al., 2009) and the average number of realised sex acts per week per individual (here set to 4, see Supplementary Methods).

In the case of PL networks (Fig. 4A), we also observe that, with the A_{lin} PSA allocation model, disease doubling time decreases significantly with network size (in red, slope $= -6.9e{-}4$, p-value < 0.001), as expected for a network with a PL distribution (Farkas et al., 2001). For A_{sat}, however, the speed of disease spread is not affected by network size (in brown, slope $= 3.2e{-}5$, p-value $= 0.51$). Finally, for A_{cst}, the spread is even slower on larger networks (in blue, slope $= 9.5e{-}4$, p-value < 0.001). This result is interesting because, in the case of HIV, the doubling time of the disease has been shown to be constant whatever the geographic scale (city, country or continent) considered for the sexual network (May and Anderson, 1987, 1988; Anderson and May, 1991). The current explanation for this is that some sexual contact networks that seem to be PL are actually NB (Hamilton et al., 2008; Handcock and Jones, 2006). Our results suggest a new explanation, which is that even if the network node degree distribution follows a power law, edge weights can lead to the loss of decrease in doubling time with network size. Put differently, the mismatch between epidemiological data and the network structure could be due to the fact that using an unweighted network implicitly implies making unrealistic biological assumptions regarding the number of sex acts.

3.4. Effect of node removal on disease spread

Finally, we consider the effect on the median doubling time of removing a node on the empirical network (Fig. 5A). Node removal could be interpreted as any measure that blocks potential transmission. The difference between a prevention strategy and a treatment strategy is that node removal is independent from the host's infectious status. In other words, the goal is to modify the shape of the network in order to slow the speed of spread of an STI that may emerge in the population. Note that we are looking at doubling times so the higher the value, the more slowly the disease spreads. With the A_{lin} model, removing individuals with the highest number of partners (i.e. a targeted strategy) has the strongest effect on disease spread. This is captured by the slope between the degree of the node removed and the increase in doubling time: a zero slope means that there is no effect of targeting individuals with an increasing degree. Here, we find a strongly increasing slope. This can be explained intuitively by the fact that super-connected individuals (or hubs) are also super-spreaders of the disease. For the A_{cst} model, targeting hubs has less impact than in the A_{lin} model (the slope of the regression between the change in median doubling time and the degree of the node removed drops from approximately 1/3 compared to the slope of the linear case). This suggests that other criteria might be more appropriate than node degree to detect key spreaders but none of the parameters we tested (number of sex acts realised, average degree of the neighbouring nodes, generalisation of this value to weighted networks; Eguiluz and Klemm, 2002; Barrat et al., 2004a, 2004b) performed better than the degree.

In a second procedure, we allow for multiple node removal (Fig. 5B and C). In this case, we compare the efficiency of a prevention policy aimed at individuals with many partners (in green) to a non-specific policy (in black). Again our goal is to see which modification of the network leads to a slower spread of a future emerging STI. We find that several individuals need to be removed for there to be a significant difference between these policies. If the PSA allocation is constant, prevention policies have less effect on the disease doubling time and more individuals need to be removed for there to be a difference between specific and non-specific targeting (see also Supplementary Results). This corroborates the results obtained in Fig. 4A by showing that in biologically realistic networks, aiming prevention policies at super-connected individuals might not yield a decrease in the spread of emerging STI as strong as one might expect. These results hold for theoretical networks (see Supplementary Results).

4. Discussion

Over the last decade, a passionate debate has emerged because sexual contact networks have been shown to have a highly heterogeneous structure (Liljeros et al., 2001; Hamilton et al., 2008; Schneeberger et al., 2004) but the spread of STI does not have the properties one would expect to find given such network structures (Hamilton et al., 2008; Handcock and Jones, 2006). Our results fit into this debate because they show that epidemiological trends in the data are consistent with the assumption that not all the edges of the network have the same weight. In other words, even though sexual networks do have a highly heterogeneous structure, they do not have associated properties because they are weighted and this weighting is not captured by studies that use contact-tracing or other methods to infer the shape of the network.

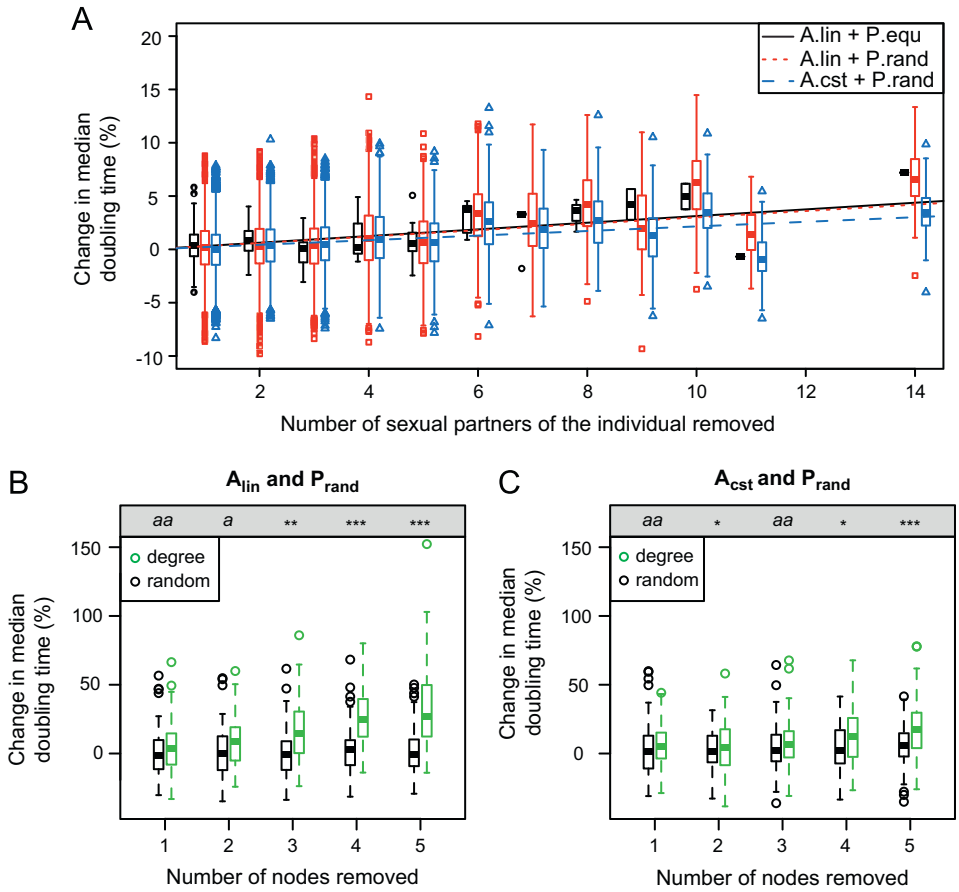


Fig. 5. Effect of prevention policies on disease spread on the empirical network. (A) Effect of single node removal on disease median doubling time as a function of node degree in the scenario A_{lin} with P_{equ} (the 'classical' case, in black), A_{lin} with P_{rand} (in red) or A_{cst} with P_{rand} (in blue). Dashed lines indicate the correlation found using a generalised linear model. Parameters are as in Fig. 2. Removing certain nodes increases disease spread (especially for a random partitioning) because we consider a prevention policy, which is done before the epidemic begins. (B) Effect of multiple node removal on disease median doubling time with the A_{lin} with P_{rand} scenario. In black, individuals are targeted at random (random) whereas in green the individuals with highest number of sexual partners (degree) are removed first. The grey box on the top indicates the *p*-value of a Wilcoxon signed-rank test comparing removal strategies. The significancy code is '***' for ≤ 0.001, '**' for ≤ 0.01, '*' for ≤ 0.05 and *a* for non-significant but ≤ 0.1 and *aa* for non-significant. (C) As panel B in the scenario with A_{cst} with P_{rand}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We developed a framework to weight sexual contact networks using biological assumptions. This allows us to show that the underlying assumptions made when modelling disease spread on unweighted networks lead to correlation patterns between number of partners and number of sex acts that are at odds with empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006). We also show that adopting a more realistic (weighted) network affects epidemiological dynamics in a way that is consistent with epidemiological data on HIV. Furthermore, these results hold for theoretical power-law networks and negative-binomial networks.

Our goal in this study is not to yield the most realistic model (we actually doubt that a decent level of realism can be achieved without adding a very large number of assumptions). Rather, our aim is to develop a general framework that is flexible enough to be applied to different epidemiological and sociological scenarios, while remaining biologically relevant. To the best of our knowledge, our allocation-partitioning model is first to present a biologically plausible relationship between total number of sex acts and total number of partners derived from explanatory behavioural hypotheses. In earlier studies (e.g., Newman, 2002; Joo and Lebowitz, 2004), this relationship is imposed without any mechanistic justification.

Interestingly the ratio of realised over potential number of sex acts (δ in Table 1) can yield insights into associated human behaviours (allocation and partitioning). This suggests that combining our model with epidemiological data could allow us to investigate some sexual behaviours.

Optimal public health policies depend on the way diseases spread (Anderson and May, 1991). Even if weighted networks lose their heterogenous-related properties, removing the most connected nodes still has the strongest effect on the STI spread. However, if the prevention policy covers a too small fraction of the population, targeting super-spreader might not be a significantly better strategy than random targeting. This effect is accentuated by using a more realistic allocation model for sex acts. Of course, these results on the effect of public health should be taken with care, for instance because the epidemiological model we use is extremely simplified.

Other studies have addressed the problem of the discrepancy between network theory and HIV epidemiological data (see e.g. Röttingen and Garnett, 2002; Hamilton et al., 2008). In particular, a widespread method is to model 'concurrent partnerships' by considering that not all the edges of the network are active at the same time (Morris and Kretzschmar, 1997; Bauch and Rand, 2000; Eames and Keeling, 2004). Such models have increased biological realism but they require many additional assumptions that are difficult to test empirically. Moreover, they too make strong assumptions concerning the number of sex acts. In a future study, we will investigate the consequence of adding dynamical partnerships to a weighted network in order to disentangle the effect of the two on disease spread.

This study can be extended in other ways. First, other network topologies could be investigated (see also our Supplementary Results). We could consider more complicated epidemiological models and/or other STI. Also, it is very likely that people modify their behaviour once they know that they are infected with an STI (Funk et al., 2010). Finally, we focused on STI but our method can be applied to any type of networks with time-consuming interactions, e.g. in economics or in ecology (Newman, 2003).

Finally, this study calls for further studies on the link between the number of partners and the number of sex acts. Such data can be acquired without knowing the exact contact structure detail and it could greatly help to check whether the allocation and partitioning hypotheses of the model make sense. Furthermore, this study also underlines the importance of deriving new estimators to predict the speed of disease spread on weighted finite-size networks.

Acknowledgements

Many thanks to P. Bild, J. Cox, J. Heijne, M. Konschake, R. Kouyos, H. Lentz, G. Leventhal, S. Lion, G. Paul, M. Pautasso, M. Salathé and T. Stadler for helpful discussions. We are very grateful to J.J. Potterat and S.Q. Muth for sharing their contact-network data. We thank the ETH Zürich, the ENS Cachan, the Île-de-France Regional Council under the MIDEM project in the framework DIM Malinf, and the French Ministries in charge of Agriculture and Environment for financial support. MML and SA performed the simulations and analysed the data, all three authors designed the study and wrote the paper. Electronic supplementary material (ESM) contains Supplementary Methods and Supplementary Results. SA is funded by an ATIP Avenir from CNRS and INSERM.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jtbi.2012.06.031.

References

Althaus, C.L., Heijne, J.C.M., Roellin, A., Low, N., 2010. Transmission dynamics of *Chlamydia trachomatis* affect the impact of screening programmes. Epidemics 2 (3), 123–131.

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, New York, USA.

Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512, http://dxdoi.org/10.1126/science.286.5439.509.

Barrat, A., Barthelemy, M., Vespignani, A., 2004a. Modeling the evolution of weighted networks. Phys. Rev. E 70 (6), 66149, http://dx.doi.org/10.1103/PhysRevE.70.066149.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A., 2004b. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. U.S.A. 101 (11), 3747–3752, http://dx.doi.org/10.1073/pnas.0400087101.

Bauch, C., Rand, D.A., 2000. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. Proc. R. Soc. Lond. B 267 (1456), 2019–2027, http://dx.doi.org/10.1098/rspb.2000.1244.

Blower, S.M., Boe, C., 1993. Sex acts, sex partners, and sex budgets: implications for risk factor analysis and estimation of HIV transmission probabilities. J. Acquir. Immune Defic. Syndr. 6 (12), 1347–1352.

Boily, M.C., Baggaley, R.F., Wang, L., Masse, B., White, R.G., Hayes, R.J., Alary, M., 2009. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect. Dis. 9 (2), 118–129, http://dxdoi.org/10.1016/S1473-3099(09)70021-0.

Britton, T., Nordvik, M.K., Liljeros, F., 2007. Modelling sexually transmitted infections: the effect of partnership activity and number of partners on R_0. Theor. Popul. Biol. 72 (3), 389–399 , http://dx.doi.org/10.1016/j.tpb.2007.06.006.

Britton, T., Deijfen, M., Liljeros, F., 2011. A weighted configuration model and inhomogeneous epidemics. J. Stat. Phys. 145, 1368–1384, http://dx.doi.org/10.1007/s10955-011-0343-3.

Catanzaro, M., Boguñá, M., Pastor-Satorras, R., 2005. Generation of uncorrelated random scale-free networks. Phys. Rev. E 71 (2), 027103, http://dx.doi.org/10.1103/PhysRevE.71.027103.

Eames, K.T.D., Keeling, M.J., 2004. Monogamous networks and the spread of sexually transmitted diseases. Math. Biosci. 189, 115–130, http://dx.doi.org/10.1016/j.mbs.2004.02.003.

Eames, K.T.D., Read, J.M., Edmunds, W.J., 2009. Epidemic prediction and control in weighted networks. Epidemics 1 (1), 70–76, http://dx.doi.org/10.1016/j.epidem.2008.12.001.

Eguiluz, V.M., Klemm, K., 2002. Epidemic threshold in structured scale-free networks. Phys. Rev. Lett. 89 (10), 108701, http://dx.doi.org/10.1103/PhysRevLett.89.108701.

Farkas, I.J., Derenyi, I., Barabasi, A.L., Vicsek, T., 2001. Spectra of 'real-world' graphs: beyond the semicircle law. Phys. Rev. E 6402 (2), 026704, http://dx.doi.org/10.1103/PhysRevE.64.026704.

Ferguson, N.M., Garnett, G.P., 2000. More realistic models of sexually transmitted disease transmission dynamics: sexual partnership networks, pair models, and moment closure. Sex. Transm. Dis. 27 (10), 600–609.

Foucault, M., 1976. Histoire de la sexualité. Gallimard, Paris, France.

Funk, S., Salathé, M., Jansen, V.A.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J. R. Soc. Interface 7 (50), 1247–1256 , http://dx.doi.org/10.1098/rsif.2010.0142.

Hamilton, D.T., Handcock, M.S., Morris, M., 2008. Degree distributions in sexual networks: a framework for evaluating evidence. Sex. Transm. Dis. 35 (1), 30–40 , http://dx.doi.org/10.1097/OLQ.0b013e3181453a84.

Handcock, M.S., Jones, J.H., 2004. Likelihood-based inference for stochastic models of sexual network formation. Theor. Popul. Biol. 65 (4), 413–422, ISSN 0040-5809, http://dx.doi.org/10.1016/j.tpb.2003.09.006.

Handcock, M.S., Jones, J.H., 2006. Interval estimates for epidemic thresholds in two-sex network models. Theor. Popul. Biol. 70 (2), 125–134 , http://dx.doi.org/10.1016/j.tpb.2006.02.004.

Joo, J., Lebowitz, J.L., 2004. Behavior of susceptible-infected-susceptible epidemics on heterogeneous networks with saturation. Phys. Rev. E 69 (6), 66105 , http://dx.doi.org/10.1103/PhysRevE.69.066105.

Keeling, M.J., Eames, K.T., 2005. Networks and epidemic models. J. R. Soc. Interface 2 (4), 295–307, http://dx.doi.org/10.1098/rsif.2005.0051.

Liljeros, F., 2004. Sexual networks in contemporary Western societies. Physica A 338 (1–2), 238–245, ISSN 0378-4371, http://dx.doi.org/10.1016/j.physa.2004.02.046.

Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Aberg, Y., 2001. The web of human sexual contacts. Nature 411 (6840), 907–908 , http://dx.doi.org/10.1038/35082140.

May, R.M., 2006. Network structure and the biology of populations. Trends Ecol. Evol. 21 (7), 394–399 , http://dx.doi.org/10.1016/j.tree.2006.03.013.

May, R.M., Anderson, R.M., 1987. Transmission dynamics of HIV infection. Nature 326 (6109), 137–142 , http://dx.doi.org/10.1038/326137a0.

May, R.M., Anderson, R.M., 1988. The transmission dynamics of human immunodeficiency virus (HIV). Philos. Trans. R. Soc. B 321 (1207), 565–607, http://dx.doi.org/10.1098/rstb.1988.0108.

May, R.M., Gupta, S., McLean, A.R., 2001. Infectious disease dynamics: what characterizes a successful invader? Philos. Trans. R. Soc. B 356 (1410), 901–910.

Morris, M., Kretzschmar, M., 1997. Concurrent partnerships and the spread of HIV. AIDS 11 (5), 641–648, ISSN 0269-9370.

Newman, M.E.J., 2002. Spread of epidemic disease on networks. Phys. Rev. E 66 (1), 16128 , http://dx.doi.org/10.1103/PhysRevE.66.016128.

Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Rev. 45 (2), 167–256.

Nordvik, M.K., Liljeros, F., 2006. Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. Sex. Transm. Dis. 33 (6), 342–349, http://dx.doi.org/10.1097/01.olq.0000194601.25488.b8.

Potterat, J.J., Rothenberg, R.B., Woodhouse, D.E., Muth, J.B., Pratts, C.I., Fogle, J.S., 1985. Gonorrhea as a social disease. Sex. Transm. Dis. 12 (1), 25–32.

Röttingen, J.-A., Garnett, G.P., 2002. The epidemiological and control implications of HIV transmission probabilities within partnerships. Sex. Transm. Dis. 29 (12), 818–827.

Schneeberger, A., Mercer, C.H., Gregson, S.A.J., Ferguson, N.M., Nyamukapa, C.A., Anderson, R.M., Johnson, A.M., Garnett, G.P., 2004. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. Sex. Transm. Dis. 31 (6), 380–387.

Vittinghoff, E., Douglas, J., Judson, F., McKirnan, D., MacQueen, K., Buchbinder, S.P., 1999. Per-contact risk of human immunodeficiency virus transmission between male sexual partners. Am. J. Epidemiol. 150 (3), 306–311.

Wang, H.Z., Liu, Z.R., Xu, H.H., 2007. Epidemic spreading on uncorrelated heterogenous networks with non-uniform transmission. Physica A 382 (2), 715–721, http://dx.doi.org/10.1016/j.physa.2007.04.034.

Annexe B

Supporting information for : *Epidemic spread on weighted networks*

B.1 Equations for the epidemic model on weighted networks

We consider an epidemic of a disease that is transmitted with probability π per interaction event and from which infected individuals recover at a rate γ. Susceptible individuals with k contacts and l interaction events per time interval are infected at a rate proportional to π, l and the probability that a susceptible individual's contact is made with an infected individual p_{SI}. This leads to the following equations for the evolution of the number of susceptible and infected individuals with k (infectious) contacts and l interaction events per time

$$\dot{S}_{kl} = -\pi p_{SI} l S_{kl} \tag{B.1a}$$

$$\dot{I}_{kl} = +\pi p_{SI} l S_{kl} - \gamma I_{kl} \tag{B.1b}$$

$$\dot{R}_{kl} = \gamma I_{kl}. \tag{B.1c}$$

A detailed overview of the model's notation and parameters is given in Table 3.1.

Adding up the contributions for all k and l introduces the average number of interaction events per time and susceptible individual $\langle l \rangle_S = \sum_l l P_{Skl} = \sum_l l \frac{S_{kl}}{S}$ into the equations. This average number can also be expressed in terms of probability generating function $G_S(x, y, t) = \sum_{kl} P_{Skl}(t) x^k y^l$ of the joint probability distribution to find k contacts and l interaction events per time among susceptible individuals P_{Skl} : $\langle l \rangle_S = \sum_{k,l} l \frac{S_{kl}}{S} = G_S^{(0,1)}(1, 1, t)$. The $(0, 1)$ exponent of G_S indicates the orders of the partial derivatives with respect to the first and second argument of G_S (see Table 3.1).

Summation of S_{kl} and I_{kl} over k and l results in equations for the total number of susceptible and

infected hosts :

$$\dot{S} = -\pi p_{SI} S G_S^{(0,1)}(1,1,t) \tag{B.2a}$$

$$\dot{I} = \pi p_{SI} S G_S^{(0,1)}(1,1,t) - \gamma I \tag{B.2b}$$

$$\dot{R} = \gamma I \tag{B.2c}$$

To close this set of equations we also need to derive equations for p_{SI}, as well as for the probability generating function (PGF) $G_S(x,y,t)$.

We begin by deriving the temporal dynamics of the probabilities for a link starting from a randomly selected susceptible individual to point to a susceptible or infected individual, p_{SS} and p_{SI}, respectively. Following the argument in [Volz, 2008] we write $p_{SS} = M_{SS}/M_S$ and $p_{SI} = M_{SI}/M_S$ to express these probabilities in terms of the total number of links/contacts that connect susceptible and infected hosts (M_{SS}, M_{SI}) and total number of links/contacts of susceptible hosts (M_S) in the network. From this, we get :

$$\dot{p}_{SS} = \frac{\dot{M}_{SS}}{M_S} - \frac{\dot{M}_S}{M_S} p_{SS} \tag{B.3a}$$

$$\dot{p}_{SI} = \frac{\dot{M}_{SI}}{M_S} - \frac{\dot{M}_S}{M_S} p_{SI} \tag{B.3b}$$

for which expressions are derived in the following paragraph. From the definition of M_S, we can write the following equation :

$$\dot{M}_S = \sum_{k,l} k \dot{S}_{kl} \tag{B.4}$$

Substitution of $\dot{S}_{kl}$ from equation (B.1a) results in

$$\dot{M}_S = -\pi p_{SI} S G_S^{(1,1)}(1,1,t) \tag{B.5}$$

We then follow the arguments made in an earlier study [Volz, 2008], which rely on the assumption that the number of contacts from susceptible hosts to susceptible, infected and recovered hosts is multinomially distributed with probabilities p_{SI}, p_{SS} and $p_{SR} = 1 - p_{SS} - p_{SI}$. We also assume that the same applies to the number of interaction events/sex acts per time interval. If a node with k contacts has j contacts with susceptible individuals and i contacts with infected individuals its interaction events with susceptible, infected and recovered individuals n_{SS}, n_{SI} and n_{SR}, respectively, are distributed according to

$$\frac{l!}{n_{SS}! n_{SI}! n_{SR}!} \left(\frac{j}{k}\right)^{n_{SS}} \left(\frac{i}{k}\right)^{n_{SI}} \left(\frac{k-j-i}{k}\right)^{n_{SR}} \tag{B.6}$$

with averages $\langle n_{SS}\rangle = \frac{jl}{k}$,$\langle n_{SI}\rangle = \frac{il}{k}$ and $\langle n_{SR}\rangle = \frac{(k-j-i)l}{k}$. Note that $l = n_{SS} + n_{SI} + n_{SR}$.

The probability that a susceptible node with l interaction events and j, i and $k-i-j$ contacts to susceptible, infected and recovered individuals, respectively, is reached from an infected node, i.e. chosen with a probability proportional to the average number of interaction events with infected nodes ($\langle n_{SI}\rangle = \frac{li}{k}$) is then

$$\frac{P_{kl}\frac{k!}{i!j!k-i-j!}p_{SS}^j p_{SI}^i p_{SR}^{k-j-i}\langle n_{SI}\rangle}{p_{SI}G^{(0,1)}(1,1)} \quad \overset{\langle n\rangle_{SI}=\frac{li}{k}}{=} \quad \frac{lP_{kl}\frac{(k-1)!}{(i-1)!j!(k-1-(i-1)-j)!}p_{SS}^j p_{SI}^{i-1} p_{SR}^{k-j-i}}{G^{(0,1)}(1,1)} \tag{B.7}$$

Note that the denominator can be derived using the observation that

$$\begin{aligned}
&\sum_{k,l} P_{kl} l p_{SI} \sum_{i,j\le k} \frac{(k-1)!}{(i-1)!j!(k-1-(i-1)-j)!} p_{SS}^j p_{SI}^{i-1} p_{SR}^{k-j-i} \\
&= \sum_{k,l} P_{kl} l p_{SI} (p_{SS}+p_{SI}+p_{SR})^{k-1} && \text{(B.8a)}\\
&= p_{SI}\sum_{k,l} l P_{kl} = p_{SI} G^{(0,1)}(1,1) && \text{(B.8b)}
\end{aligned}$$

Therefore, the probability distribution for the contacts and potential interaction events of a node which was chosen proportional to $\langle n_{SI}\rangle = \frac{li}{k}$ is generated by

$$\begin{aligned}
&\frac{\sum_{kl} l P_{kl} y^l \sum_{i,j\le k} \frac{(k-1)!}{(i-1)!j!(k-1-(i-1)-j)!}(x_S p_{SS})^j (x_I p_{SI})^i (x_R p_{SR})^{k-j-i}}{p_{SI} G^{(0,1)}(1,1)} \\
&= \frac{\sum_{kl} l P_{kl} x_I y^l (x_S p_{SS} + x_I p_{SI} + x_R p_{SR})^{k-1}}{G^{(0,1)}(1,1)} && \text{(B.9a)}\\
&= \frac{x_I y}{x_S p_{SS} + x_I p_{SI} + x_R p_{SR}} \frac{G^{(0,1)}(x_S p_{SS} + x_I p_{SI} + x_R p_{SR}, y)}{G^{(0,1)}(1,1)} && \text{(B.9b)}
\end{aligned}$$

Choosing a node proportional to its average number of interaction events per time ($\langle n_{SI}\rangle = il/k$) instead of the actual number of interaction events (n_{SI}) implies the assumption of a time scale at which $l \gg k$, i.e. a case in which fluctuations around $\langle n_{SI}\rangle$ can be expected to be small. Taking all this together, the average degrees of a susceptible node that was reached from an infected node to susceptible or infected nodes are

$$\begin{aligned}
\delta_{SI}(S) &= \frac{\partial}{\partial x_S} \frac{x_I y}{x_S p_{SS} + x_I p_{SI} + x_R p_{SR}} \frac{G_S^{(0,1)}(x_S p_{SS} + x_I p_{SI} + x_R p_{SR}, y, t)}{G_S^{(0,1)}(1,1)}\Bigg|_{x_S=x_I=x_R=y=1} \\
&= p_{SS}\left(\frac{G_S^{(1,1)}(1,1,t)}{G_S^{(0,1)}(1,1,t)} - 1\right) && \text{(B.10a)}\\
\delta_{SI}(I) &= \frac{\partial}{\partial x_I} \frac{x_I y}{x_S p_{SS} + x_I p_{SI} + x_R p_{SR}} \frac{G_S^{(0,1)}(x_S p_{SS} + x_I p_{SI} + x_R p_{SR}, y, t)}{G_S^{(0,1)}(1,1)}\Bigg|_{x_S=x_I=x_R=y=1} \\
&= p_{SI}\left(\frac{G_S^{(1,1)}(1,1,t)}{G_S^{(0,1)}(1,1,t)} - 1\right) + 1 && \text{(B.10b)}
\end{aligned}$$

The average number of contacts to susceptible and infected nodes needs to be discounted by one for the number of contacts with infected nodes to take the contact to the source of infection into account (directly considering this in the PGF gives the same result), i.e. the total excess degree of a node that was chosen proportional to its average number of interaction events with infected individuals $\langle n_{SI} \rangle = il/k$ is $\frac{G_S^{(1,1)}(1,1,t)}{G_S^{(0,1)}(1,1,t)} - 1$.

Bookkeeping of the changes in the numbers of links among susceptible and infected hosts due to the epidemic process results in :

Changes due to epidemic spread

$$\dot{M}_{SI} = -\dot{S}(p_{SS} - p_{SI})\left(\frac{G_S^{(1,1)}(1,1,t)}{G_S^{(0,1)}(1,1,t)} - 1\right)$$
change in the number of susceptible nodes $\dot{S} = -\pi p_{SI} S G_S^{(0,1)}(1,1,t)$ (cf. equation B.2a) due to the epidemic multiplied by their average excess contacts to susceptible and infected nodes (cf. equations B.10a and B.10b)

$$-\pi \frac{G_S^{(0,1)}(1,1,t)}{G_S^{(1,0)}(1,1,t)} M_{SI}$$
discount for link along which the infection spread

$$-\gamma M_{SI}$$
link loss due to recovery of infected

$$\dot{M}_{SS} = \dot{S} 2 p_{SS}\left(\frac{G_S^{(1,1)}(1,1,t)}{G_S^{(0,1)}(1,1,t)} - 1\right)$$
change in the number of susceptible nodes $\dot{S} = -\pi p_{SI} S G_S^{(0,1)}(1,1,t)$ due to the epidemic multiplied by their average excess contacts to infected nodes (bi-directional)

In summary this results in

$$\dot{M}_{SI} = \pi p_{SI} S\left(G_S^{(0,1)}(1,1,t) - G_S^{(1,1)}(1,1,t)\right)(p_{SI} - p_{SS}) - \pi \frac{G_S^{(0,1)}(1,1,t)}{G_S^{(1,0)}(1,1,t)} M_{SI} - \gamma M_{SI} \tag{B.11a}$$

$$\dot{M}_{SS} = -2\pi p_{SS} p_{SI} S\left(G_S^{(1,1)}(1,1,t) - G_S^{(0,1)}(1,1,t)\right) \tag{B.11b}$$

which finally leads to

$$\dot{p}_{SI} = \pi p_{SI} p_{SS} \frac{G_S^{(1,1)}(1,1,t) - G_S^{(0,1)}(1,1,t)}{G_S^{(1,0)}(1,1,t)} - \pi p_{SI}(1 - p_{SI}) \frac{G_S^{(0,1)}(1,1,t)}{G_S^{(1,0)}(1,1,t)} - \gamma p_{SI} \tag{B.12a}$$

$$\dot{p}_{SS} = -\pi p_{SI} p_{SS} \frac{G_S^{(1,1)}(1,1,t) - 2G_S^{(0,1)}(1,1,t)}{G_S^{(1,0)}(1,1,t)}. \tag{B.12b}$$

To close the set of equations we also need to derive an equation for the probability generating function (PGF) $G_S(x,y,t)$, which corresponds to the probability to find individuals with k contacts and l inter-

action events (e.g. sex acts) per time interval among susceptible hosts, i.e. P_{Skl}. From the definitions of the PGF and $P_{Skl} = \frac{S_{kl}}{S}$, we obtain

$$\dot{G}_S(x,y,t) = \sum_{k,l}\left(\frac{\dot{S}_{k,l}}{S} - \frac{\dot{S}}{S}P_{Skl}\right)x^k y^l, \tag{B.13}$$

which results with equations B.1a and B.2a in

$$\dot{G}_S(x,y,t) = \pi p_{SI}\left(G_S^{(0,1)}(1,1,t)G_S(x,y,t) - yG_S^{(0,1)}(x,y,t)\right) \tag{B.14}$$

The probability generating functions $G_I(x,y,t)$ and $G_R(x,y,t)$ can be derived analogously, though they are not needed to close the system of equations :

$$\dot{G}_I(x,y,t) = -\pi p_{SI}\frac{S}{I}\left(G_S^{(0,1)}(1,1,t)G_I(x,y,t) - yG_S^{(0,1)}(x,y,t)\right) \tag{B.15a}$$

$$\dot{G}_R(x,y,t) = \pi\frac{I}{R}\left(G_I(x,y,t) - G_R(x,y,t)\right). \tag{B.15b}$$

B.2 Conditional probabilities and risk groups

Note that $P_{kl} = P(k,l)$ and that for $A \in \{S,I,R\}$ $P_{Akl} = P(k,l|A)$ are the conditional probabilities to have k contacts and l transmission events given that you have status A. This allows for a direct derivation of the average number of contacts $\langle k\rangle_A$ and transmission events $\langle l\rangle_A$ given that you have status A.

$$\langle k\rangle_A = G_A^{(1,0)}(1,1,t) \tag{B.16a}$$

$$\langle l\rangle_A = G_A^{(0,1)}(1,1,t). \tag{B.16b}$$

Bayes' theorem together with equations B.14-B.15b allows to derive the conditional probabilites for individuals with a certain number of contacts k and interaction events l to be susceptible, infected recovered, i.e. to identify risk groups :

$$P(A|k,l) = \frac{P(A)}{P(k,l)}P(k,l|A) = \frac{P(A)}{P(k,l)}\frac{1}{k!l!}\frac{\partial^{k+l}}{\partial x^k \partial y^l}G_A(x,y,t)_{|x=y=0} \tag{B.17a}$$

$$P(A|k) = \frac{P(A)}{P(k)}P(k|A) = \frac{P(A)}{P(k)}\frac{1}{k!}\frac{\partial^k}{\partial x^k}G_A(x,y,t)_{|x=0,y=1} \tag{B.17b}$$

$$P(A|l) = \frac{P(A)}{P(l)}P(l|A) = \frac{P(A)}{P(l)}\frac{1}{l!}\frac{\partial^l}{\partial y^l}G_A(x,y,t)_{|x=1,y=0}. \tag{B.17c}$$

B.3 The basic reproductive ratio R_0

The basic reproductive ratio R_0 of a SIR epidemic with transmission rate π and recovery rate γ on a classical (unweighted) network can be derived as [Durrett, 2007]

$$R_0 = \frac{g^{(2)}(1)}{g^{(1)}(1)} \int_0^\infty \left(1 - e^{-\pi t}\right) \gamma e^{-\gamma t} dt \tag{B.18a}$$

$$= \frac{\pi}{\pi + \gamma} \left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right) \tag{B.18b}$$

where g is the probability generating function of the network's degree distribution. It is the product of the average excess degree of a node which was reached according to its degree and the transmissibility, i.e. the probability that an infection is spread along a link before recovery (here $\pi/(\pi+\gamma)$). These terms do not factorize in the case where there are l interaction events per node defined through the joint probability distribution P_{kl}. To derive R_0 for this case we first derive the excess degree distribution Q_{kl} of a node that was reached with probability proportional to its activity l and that has k excess contacts and l interaction events per time interval

$$Q_{kl} = \frac{l P_{(k+1)l}}{G^{(0,1)}(1,1)}, \tag{B.19}$$

and get, for l interaction events multinomially distributed among the $k+1$ contacts (with probabilities $p_1 = ... = p_{k+1} = \frac{1}{k+1}$, $m_1 + ... + m_{k+1} = l$),

$$R_0 = \sum_{k,l} \int_0^\infty \gamma e^{-\gamma t} \sum_{m_1,...,m_{k+1}} \frac{l!}{m_1!...m_{k+1}!} p_1^{m_1} ... p_{k+1}^{m_{k+1}} \sum_{j=1}^{k} (1 - e^{-m_j \pi t}) Q_{kl} dt \tag{B.20a}$$

$$= \sum_{k,l} \sum_{m_1,...,m_{k+1}} \frac{l!}{m_1!...m_{k+1}!} p_1^{m_1} ... p_{k+1}^{m_{k+1}} \sum_{j=1}^{k} \frac{m_j \pi}{m_j \pi + \gamma} Q_{kl} \tag{B.20b}$$

The SI model is trivially included in the SIR network models in the limit $\gamma \to 0$ for which R_0 can be derived from equation (B.20b) :

$$R_0 = \sum_{k,l} \sum_{m_1,...,m_{k+1}} \frac{l!}{m_1!...m_{k+1}!} p_1^{m_1} ... p_{k+1}^{m_{k+1}} k Q_{kl} \tag{B.21a}$$

$$= \sum_{k,l} kl \frac{P_{k+1,l}}{G^{(0,1)}(1,1)} \tag{B.21b}$$

$$= \frac{G^{(1,1)}(1,1) - G^{(0,1)}(1,1)}{G^{(0,1)}(1,1)} \tag{B.21c}$$

$$= \frac{\langle kl \rangle - \langle l \rangle}{\langle l \rangle} = \frac{\langle kl \rangle}{\langle l \rangle} - 1 \tag{B.21d}$$

As long as infected individuals stay indefinitely infected, R_0 is not affected by the transmission rate and it measures whether there is a giant connected component in classical random networks. The effect of weighting the contact network using the number of interaction events l is here also noticeable : there are not only contacts required for an infection to spread beyond an individual but also sufficient interaction events.

The basic reproductive ration R_0 for a SIR model on a weighted network can be approximated by

$$\begin{aligned} R_0 &= \sum_{k,l} \sum_{m_1,\ldots,m_{k+1}} \frac{l!}{m_1!\ldots m_{k+1}!} p_1^{m_1} \ldots p_{k+1}^{m_{k+1}} \int_0^\infty \gamma e^{-\gamma t} \sum_{j=1}^{k} (1 - e^{-m_j \pi t}) Q_{kl} dt && \text{(B.22a)} \\ &\approx \sum_{k,l} \sum_{m_1,\ldots,m_{k+1}} \frac{l!}{m_1!\ldots m_{k+1}!} p_1^{m_1} \ldots p_{k+1}^{m_{k+1}} \int_0^\infty \gamma e^{-\gamma t} k (1 - e^{-\langle \frac{l}{k} \rangle \pi t}) \frac{l P_{k+1 l}}{G^{(0,1)}(1,1)} dt \\ &= \sum_{k,l} kl \frac{P_{k+1 l}}{G^{(0,1)}(1,1)} \int_0^\infty \gamma e^{-\gamma t} (1 - e^{-\langle \frac{l}{k} \rangle \pi t}) dt \\ &= \frac{\langle \frac{l}{k} \rangle \pi}{\langle \frac{l}{k} \rangle \pi + \gamma} \frac{G^{(1,1)}(1,1) - G^{(0,1)}(1,1)}{G^{(0,1)}(1,1)}. && \text{(B.22b)} \end{aligned}$$

Note that for the linear case with $P_{l|k} = \delta_{lk}$, we obain $G(x,y) = \sum_{k,l} x^k y^l \tilde{P}_k \delta_{kl} = \sum_k (xy)^k \tilde{P}_k = \bar{G}(xy)$ and $R_0 = \frac{\pi}{\pi + \gamma} \frac{\tilde{G}^{(2)}(1)}{\tilde{G}^{(1)}(1)}$, which is consistent with earlier findings.

B.4 The recovery of the classical equations in the linear case $P_{kl} = P_k \delta_{kl}$

The set of equations for the weighted networks (equations B.2a-B.2c,B.12a-B.12b,B.14) includes the case of classical network epidemic models, i.e. the linear case where $k = l$ or $P_{kl} = P_k \delta_{kl}$. Focusing on the degree distribution among susceptible hosts P_{Sk} with probability generating function $g_S(x)$, the PGF of P_{Skl} is given by $G_S(x,y) = g_S(xy)$. Substitution of $G_S(x,y)$ by $g_S(xy)$ results in

$$\begin{aligned} G_S^{(1,0)}(1,1,t) &= g_S'(1,t) && \text{(B.23a)} \\ G_S^{(0,1)}(1,1,t) &= g_S'(1,t) && \text{(B.23b)} \\ G_S^{(1,1)}(1,1,t) &= g_S''(1,t) + g_S'(1,t) && \text{(B.23c)} \end{aligned}$$

and for the time evolution of $G_S(x,y)$:

$$\dot{g}_S(xy,t) = \pi p_{SI} \left(g_S'(1) g_S(xy,t) - xy g_S'(xy,t) \right). \qquad \text{(B.24)}$$

Together, this leads to the set of equations for SIR dynamics on a classical configuration type network defined by the degree distribution P_k [Volz, 2008; Kamp, 2010]

$$\dot{S} = -\pi\, p_{SI}\, S\, g_S'(1,t) \tag{B.25a}$$

$$\dot{I} = \pi\, p_{SI}\, S\, g_S'(1,t) - \gamma\, I \tag{B.25b}$$

$$\dot{R} = \gamma\, I \tag{B.25c}$$

$$\dot{p}_{SI} = \pi\, p_{SI}\, p_{SS}\, \frac{g_S''(1,t)}{g_S'(1,t)} - \pi\, p_{SI}\,(1-p_{SI}) - \gamma\, p_{SI} \tag{B.25d}$$

$$\dot{p}_{SS} = -\pi\, p_{SI}\, p_{SS} \left(\frac{g_S''(1,t)}{g_S'(1,t)} - 1 \right) \tag{B.25e}$$

$$\dot{g}_S(x,t) = \pi p_{SI} \left(g_S'(1,t) g_S(x,t) - x g_S'(x,t) \right). \tag{B.25f}$$

B.5 Network segregation and the limiting case $P_{kl} = P_k \delta_{\langle l \rangle l}$ (constant case)

The analytical approximation assumes that an individual distributes his/her interaction events l multinomially among his/her k contacts and is infected at a rate proportional to his/her average number of potential transmission events with i infected contacts. This averaging implies the choice of a time scale such that $\langle l \rangle > \langle k \rangle$. This leads to an unrealistic network segregation in some artificial networks, specifically for $\langle l \rangle \gg \langle k \rangle$, as the weights of an individual's contacts level at about l/k which enforces contacts only between individuals with (almost) identical l/k. This network segregation affects epidemiological dynamics. As the analytical approach is node-centric it does not consider the constraints on half-contacts to match half-contacts of similar weight. In consequence, the change in epidemic dynamics due to networks segregation cannot be seen in the analytical approach.

The effect is particularly pronounced if we have a network with a heterogeneous degree distribution (which corresponds to a case where many individuals have only one contact) combined with a constant number of interaction events per individual. Degree one nodes have only one contact to assign their interaction events to, which leaves their contacts on average already with twice the weight seen in individuals with two contacts (i.e. $\langle l \rangle$ for $k = 1$ vs. $\frac{\langle l \rangle}{2}$ for $k = 2$). This weight separation leads to a situation that almost only allows contacts among individuals with a single contact, i.e. monogamous couples (contact or degree assortativity). Therefore, individuals with one contact can only be infected if their partner is initially infected but not later on through the epidemic process because they are not connected to the giant component of the network. In the case of a constant number of interaction events per individual $\langle l \rangle$, the analytical approach breaks into independent equations for all k classes

with $\langle l \rangle_S = \langle l \rangle$ in which epidemic prevalence grows at the same rate :

$$\dot{S}_{k\langle l \rangle} = -\pi p_{SI} \langle l \rangle S_{k\langle l \rangle} \tag{B.26a}$$

$$\dot{I}_{k\langle l \rangle} = +\pi p_{SI} \langle l \rangle S_{k\langle l \rangle} - \gamma I_{k\langle l \rangle} \tag{B.26b}$$

$$\dot{R}_{k\langle l \rangle} = \gamma I_{k\langle l \rangle}. \tag{B.26c}$$

Due to the network segregation, epidemic prevalence is reduced in these networks at least by a factor proportional to the fraction of nodes with a single contact as compared to the standard result of the analytical approach. Again, this is because in these type of networks the single-contact nodes do not participate in the epidemic process.

B.6 Agreement between approximations and simulations

We evaluate the agreement between simulations and approximations based on the evolution of infected from time $t = 0$ to time $t = 1500$ expressed in arbitrary units (see main text for parameter values and details). For each network type, we compare analytical approximations (denoted $y_{raw}(t)$, $y_{rem}(t)$ and $y_{emp}(t)$) to the mean of 2000 simulation replicates at each point of time t (denoted $x(t)$). $y_{raw}(t)$, $y_{rem}(t)$ and $y_{emp}(t)$ correspond respectively to raw approximations, approximations when nodes with one partner are removed and approximations based on the empirical distributions.

The level of agreement between simulations and approximations is assessed with cross-correlation analysis and linear regression :

- *Cross correlation analysis.* We calculate $\tau^* = \max_{\tau} \mathcal{C}(x, y, \tau)$ and $\mathcal{C}^* = \mathcal{C}(x, y, \tau^*)$, where $\mathcal{C}(x, y, \tau)$ is the sample cross correlation coefficient between time series x (set as the reference) and y for a time lag set to τ (function *ccf* in R). τ takes values in $[-1500, 1500]$ since time has set in $[0, 1500]$. τ^* is the value of τ that maximizes the value of $\mathcal{C}(x, y, \tau)$, denoted $\mathcal{C}^*$. In other words, τ^* quantifies the point in time where times series x and y are optimally correlated. A perfect agreement between simulations and data implies $\tau^* = 0$ and $\mathcal{C}^* = 1$.
- *Linear regression.* We fit of the model $y(t) = ax(t)$ (function *lm* in R). The match between simulations and approximations is optimal when $a = 1$, p-value< 0.001 and $R^2 = 1$.

TABLE B.1 – Agreement between approximations and simulations

	type	P_k Poisson P_l Poisson	P_k Poisson P_l power	P_k Poisson $P_l = \delta_{kl}$	P_k power P_l Poisson	P_k power P_l power	P_k power $P_l = \delta_{kl}$
τ^*	y_{raw}	+21	-39	+13	+7	0	+9
	y_{rem}	+21	-39	+13	+7	0	+9
	y_{emp}	+12	-26	+19	-29	0	0
$\mathcal{C}^*$	y_{raw}	0.99	0.97	0.99	0.99	1.00	1.00
	y_{rem}	0.99	0.97	0.99	0.99	1.00	1.00
	y_{emp}	0.99	0.98	0.99	0.98	1.00	1.00
a	y_{raw}	1.06	1.02	1.01	1.78	1.47	1.05
	y_{rem}	0.99	0.95	0.94	0.96	0.80	0.57
	y_{emp}	1.07	1.01	1.02	1.65	1.26	1.02
p value	y_{raw}	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	y_{rem}	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	y_{emp}	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
R^2	y_{raw}	0.97	0.94	0.98	0.99	0.99	0.99
	y_{rem}	0.97	0.94	0.98	0.99	0.99	0.99
	y_{emp}	0.98	0.97	0.97	0.96	0.99	1.00

τ^* is the time lag that maximizes the cross-correlation coefficient between simulations and approximations ($\mathcal{C}$; the maximum of which is denoted $\mathcal{C}^*$). a, p-value and R^2 represent respectively the slope, p-value and percentage of variance explained associated with a linear regression of approximations against simulations. Based on our summaries, the match between simulations and approximations is maximal when $\tau^* = 0$, $\mathcal{C}^* = 1$, $a = 1$, p-value< 0.001 and $R^2 = 1$

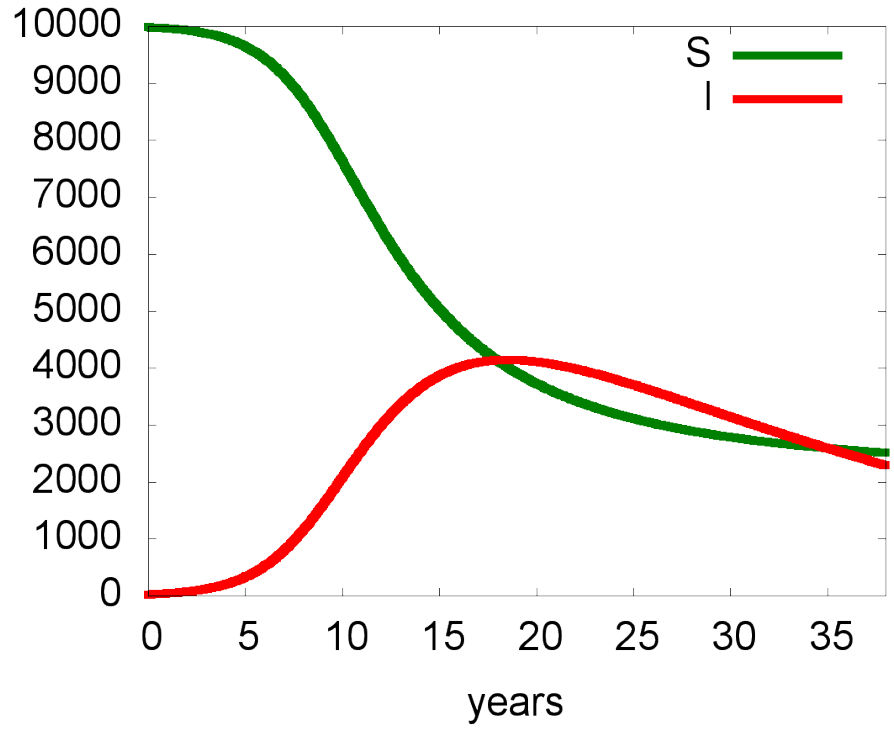


FIGURE B.1 – Epidemic SIR dynamics on the network as presented in Fig. 3.3. Transmission probability per sex act is also $\pi = 0.01$ but recovery can occur at a rate $\gamma = 0.004$ per 4 weeks, i.e. parameters corresponding to Fig. 3.2. Different from the SI dynamics shown in Fig. 3.3, hosts may recover and do not spread infection indefinitely.

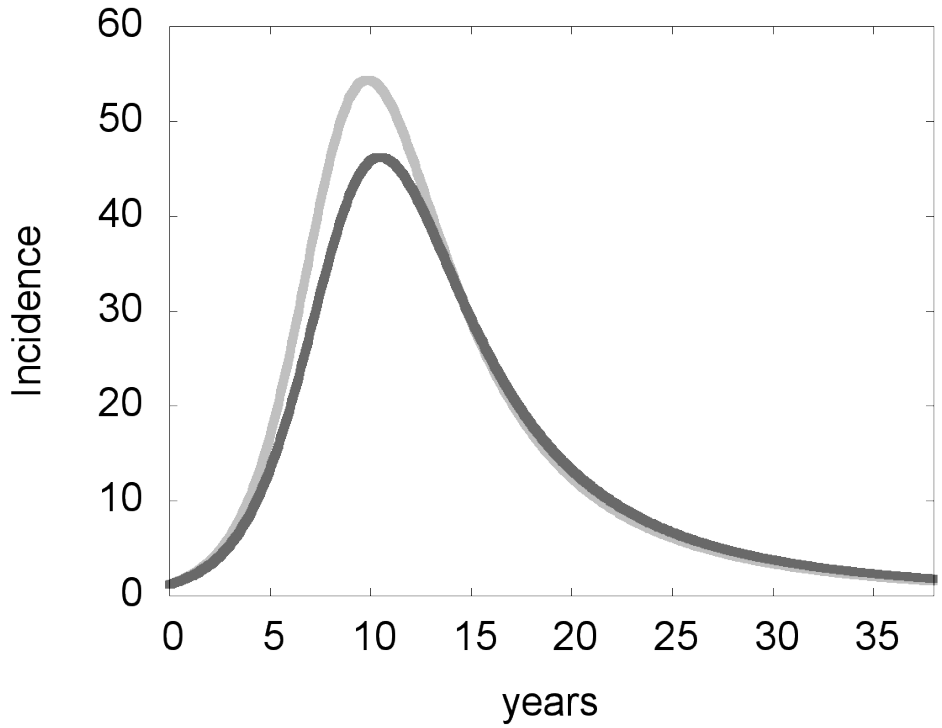


FIGURE B.2 – Epidemic incidence or rate of infection $\pi p_{SI} S G_S^{(0,1)}(1,1,t) = \pi p_{SI} S \langle I \rangle_S$ (cf. equation B.2b) for SI dynamics (grey line) and SIR dynamics (dark grey line) on the network as presented in Fig. 3.3.

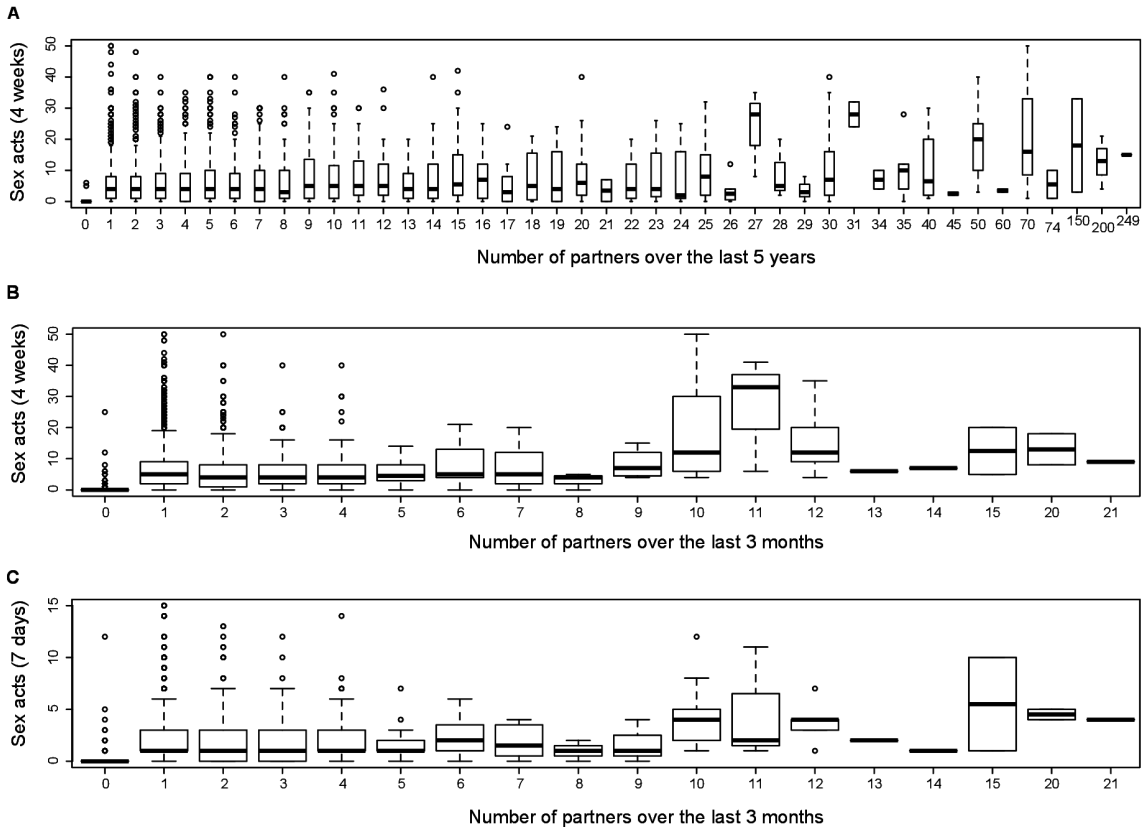


FIGURE B.3 – Relationship between a person's total number of sex acts and number of partners derived from the NATSAL data. In Panel A, we plot the self-reported number of sex acts over the last 4 weeks *vs.* the self-reported number of sexual partners over the last 4 years. In Panel B, we plot the self-reported number of sex acts over the last 4 weeks *vs.* the self-reported number of sexual partners over the last 3 months. In Panel C, we plot the self-reported number of sex acts over the last 7 days *vs.* the self-reported number of sexual partners over the last 3 months. In all three cases, the data do not support a linear relationship (the number of sex acts per partner decreases with the number of partners/contacts).

Annexe C

Research paper : *SIS along a continuum (SIS_C) epidemiological modelling and control of diseases on directed trade networks*

Mathematical Biosciences 236 (2012) 44–52

Contents lists available at SciVerse ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

SIS along a continuum (SIS_c) epidemiological modelling and control of diseases on directed trade networks

Mathieu Moslonka-Lefebvre [a,b,c,*], Tom Harwood [d], Mike J. Jeger [c], Marco Pautasso [e]

[a] INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France
[b] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK
[c] Division of Biology, Imperial College London, Silwood Park, Ascot, SL5 7PY, UK
[d] CSIRO Ecosystem Sciences, GPO Box 1700, Canberra, ACT 2601, Australia
[e] Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), CNRS, 34293 Montpellier Cedex 5, France

ARTICLE INFO

Article history:
Received 24 February 2011
Received in revised form 11 January 2012
Accepted 13 January 2012
Available online 28 January 2012

Keywords:
Directed trade networks
Epidemiology
Hierarchical categories
Leading eigenvalue
Metapopulation
SIS

ABSTRACT

Network theory has been applied to many aspects of biosciences, including epidemiology. Most epidemiological models in networks, however, have used the standard assumption of either susceptible or infected individuals. In some cases (e.g. the spread of *Phytophthora ramorum* in plant trade networks), a continuum in the infection status of nodes can better capture the reality of epidemics in networks. In this paper, a Susceptible-Infected-Susceptible model along a continuum in the infection status (SIS_c) is presented, using as a case study directed networks and two parameters governing the epidemic process (probability of infection persistence (p_p) and of infection transmission (p_t)). The previously empirically reported linear epidemic threshold in a plot of p_p as a function of p_t (Pautasso and Jeger, 2008) [29] is derived analytically. Also the previously observed negative correlation between the epidemic threshold and the correlation between links in and out of nodes (Moslonka-Lefebvre et al., 2009) [30] is justified analytically. A simple algorithm to calculate the threshold conditions is introduced. Additionally, a control strategy based on targeting market hierarchical categories such as producers, wholesalers and retailers is presented and applied to a realistic reconstruction of the UK horticultural trade network. Finally, various applications (e.g., seed exchange networks, food trade, spread of ideas) and potential refinements of the SIS_c model are discussed.

1. Introduction

Network concepts have recently found application in many fields, including conservation biogeography [1], landscape genetics [2], the study of food webs [3] and fire ecology [4]. Among these fields, the use of networks in epidemiology has proven particularly fruitful in understanding disease dynamics and control [5–7]. The advantage of network-based modelling is to represent epidemics more accurately than standard mass-action mixing models [5]. Standard mass-action mixing models describe the dynamics of infection within a population of hosts by implicitly assuming that each infected individual has a given probability of infecting every susceptible individual in the population. Such models have a robust and surprisingly good ability to predict real world epidemics [8,9]. However, there are mechanistic difficulties with the basic assumptions of the mass-action models. Most infectious diseases spread through a network of contacts, so that infection has a much higher probability of spreading to a more limited set of susceptible units. Network theory is thus a powerful tool to investigate the effect of network topology (i.e. the way the units are connected) on epidemiological processes [10–13]. Network epidemiology can help in determining which nodes should be targeted for control depending on the network topology [14–16]. For example, targeting nodes randomly in scale free networks is a less efficient control strategy than targeting super-connected nodes [17]. For plant diseases, network epidemiology is a key tool to model disease spread and to determine relevant control measures against the many plant pathogens spread by commercial transport of plants [17,18]. One example is the recently discovered oomycete *Phytophthora ramorum*, a generalist pathogen causing widespread tree mortality both in Europe and North-America and whose spread is facilitated by the trade of ornamental plants (e.g. *Rhododendron, Camellia, Pieris, Viburnum*) [19,20]. Similarly, network concepts can be applied to invasive ornamental plants [21–24].

Trade flows are well described by directed networks. For example, the plant trade is better approximated by directed networks, given that connections from e.g. growers to wholesalers and retailers are more likely than in the reverse direction [25]. Similarly, seed exchange networks and food trade tend to be

* Corresponding author at: INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France.
E-mail address: mathieu.moslonka-lefebvre@jouy.inra.fr (M. Moslonka-Lefebvre).

0025-5564/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2012.01.004

asymmetric processes. Most studies have focused on undirected networks where nodes are connected symmetrically. Directed networks, given their *a priori* asymmetric adjacency matrices, have been used relatively rarely to model epidemics [26–28]. An epidemiological model specifically designed for directed networks (e.g. with the analytic determination of the epidemic threshold) would enable a more realistic representation of epidemics spreading on markets. Moreover, such a model could be readily applied to undirected networks, as directed networks are a generalisation of undirected networks. An undirected network is indeed a directed network in which all links are bidirectional. The aim of this paper is therefore to understand the analytical relationship between directed networks and disease spread within an epidemiological modelling framework which we term Susceptible-Infected-Susceptible along a continuum (SIS_c). The SIS_c model describes the infection status of each individual unit of a directed network along a continuum ranging between full susceptibility and full infection. The SIS_c model has many applications such as information, disease and rumour spread among individuals or organisations. In this paper, we apply the SIS_c model to the particular case of disease spreading on horticultural trade networks, such as *P. ramorum* [17]. It is nevertheless important to point out that the SIS_c model can be used in various contexts including spread of ideas, metapopulation dynamics and disease spread among hospitals and related care-centres.

The main strength of the SIS_c model over classical binary models is that it avoids absorbing locks. In binary models where individuals are either infected or not, once nodes are infected, they cannot be re-infected: they are locked in an absorbing state until they recover. On the contrary, in the SIS_c framework, infectious nodes can be cumulatively infected. The SIS_c makes sense for e.g. plant nurseries, which can host a varying proportion of infected plants. If instead nodes represent individual units, a binary model may be more parsimonious than the SIS_c model.

In the following sections, the general framework of the SIS_c model is presented and analyzed in comparison to the existing (binary) SIS models. The previously empirically reported linear relationship at epidemic equilibrium between probability of infection persistence and probability of infection transmission [29] is then derived analytically. A similar analysis supports the previously reported negative correlation between the epidemic threshold and the correlation between links in and out of nodes [30]. Then, an algorithm to compute the equilibrium conditions is introduced. Additionally, hierarchical categories of the UK horticultural trade network such as producers, wholesalers and retailers [25] are targeted for control through various preventive removal strategies. Finally, various applications and potential extensions of the SIS_c model are discussed.

2. SIS along a continuum (SIS_c) model

2.1. Brief presentation of the SIS_c model

A simple way for describing epidemics is to model the infection process with a Susceptible-Infected-Susceptible or SIS model [31]. The original Kermack–McKendrick SIS model equations [32] still form today the basis for much epidemiological work in both continuous and discrete time e.g. [31,33–37]. In a SIS model, each unit is either susceptible or infected at any given time t, where time is either a discrete or a continuous variable. The discrete model is given as:

$$\begin{aligned} S(t+T) &= S(t)(1-\beta T I(t)) + \gamma T I(t) \\ I(t+T) &= I(t)(1-\gamma T) + \beta T S(t) I(t) \end{aligned} \quad (1)$$

where $t = kT$ $(k \in \mathbb{N})$, T is a fixed time interval, $S(t)$ (resp. $I(t)$) is the proportion of susceptible (resp. infected) individuals at discrete time t, and β (resp. γ) is the transmission (resp. recovery) rate.

The equivalent with a continuous time step t (when $T \to dt$ in system (1)) is given as:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t) I(t) + \gamma I(t) \\ \frac{dI(t)}{dt} &= -\gamma I(t) + \beta S(t) I(t) \end{aligned} \quad (2)$$

It is generally assumed in discrete SIS models that a long discrete-time interval T corresponds to the length of the latent period (time from infection to infectiousness) and the infectious period is assumed to have length zero or at least a length very small in comparison with the latent period and vice versa for a continuous model [31,38]. As we aim to describe epidemics in the plant trade, we decide to focus in this paper on a discrete-time model, the time step of the model being small or large depending on the relative weight one wants to put on the latent period (large T for a large latent period compared to the infectious period) or the infectious period (small T for a large infectious period compared to the latent period). We can nonetheless extend some results of this paper to continuous time by decreasing the discrete time step up to the limit 0.

In our model, similarly to the households studied by Hiebeler [39], the infectious state of each node is represented as a continuum of states, ranging from absence of infection (susceptible node) to full infection. In this network-based SIS_c model, an infected node is infected along a continuum of values. This is realistic for epidemics in the plant trade, as plant nurseries and retail centres can be completely devoid of a pathogen, fully infected, or everything in between [29].

Moreover, we deliberately chose not to implement a SIR model, i.e. an epidemiological model where infected nodes are eventually removed [31], because a SIS model is a realistic assumption for the nursery trade. Given that immunization is not possible, if growers carry on trading susceptible species, then they are still at risk of becoming infected even after eradication of a certain pathogen and the adoption of good practices.

We define two probabilities: the probability of infection transmission p_t, i.e. the probability that an infected unit will transmit the infection to a susceptible unit, and the probability of infection persistence p_p, i.e. the probability that an infected unit will still be infected at the next time step of the epidemic. Hence, the variable p_p incorporates the concept of infectious period or length of infectiousness, thus enabling a better modelling of a SIS situation in a discrete time step framework compared to the other SIS discrete models which are assuming implicitly a very small infectious period. Additionally, p_p can be considered as a descriptor of the recovery rate, detection, and control measures [29]. This aggregation of properties in p_p enables a parsimonious description of the system. This choice can be linked to the classical modelling problem of Richard Levins, where there is a trade off in a model between generality, realism, and precision [40]. In our case, we clearly put more weight in favour of generality (p_p is an aggregate) and realism (the relationships between units are explicitly described in a relational network) compared to precision (e.g. we cannot dissociate the effects of recovery rate and length of infectiousness on epidemics). The aggregate p_p makes our model parsimonious but it is e.g. impossible to disentangle the individual contributions of recovery rate and control measures to p_p.

2.2. Formal presentation of the SIS_c model

2.2.1. Discrete SIS_c model

A network of N interacting units (e.g. plants, individuals, species, nurseries, farms, hospitals or airports) is composed of N nodes

or vertices interacting with each other by edges or links, the connections being here purely relational and not necessarily related to the Euclidean distance between two entities [41]. Mathematically, a network can be represented by an $N \times N$ adjacency matrix **A** which is defined in the following way: elements $A(i,j)$ $(i \neq j)$ are equal to $p_t \in]0\,1]$ if there is a link connecting node i to node j and 0 otherwise. In this paper, we are interested in the general situation where $A(i,j)$ is not necessarily equal to $A(j,i)$. In this case, the network is directed, as opposed to undirected networks. In directed networks, a link going from node i to node j is called an arrow or directed link and is symbolized by the couple (i,j). Additionally, to implement the concept of probability of persistence of an infection, we set the elements of the diagonal of that adjacency matrix, that is to say the elements $A(i,i)$, all equal to $p_p \in [0\,1]$.

Moreover, p_p and p_t are intrinsic parameters of the SIS_c model because they are found directly in the adjacency matrix, as opposed to most SIS models where the probability of infection transmission is for instance an external parameter of the system. Systems (1) and (2) give an illustration of the external states of parameters β (transmission rate) and, γ (recovery rate) as opposed to p_t and p_p respectively. p_t represents the intrinsic probability of transmission of a disease between two connected nodes (strength of an arrow or directed edge) and p_p represents the intrinsic persistence of a disease after an infection and is constitutive of a node. It is also important to notice that, in our model, p_t and p_p are the same for a whole network replicate. Our model makes thus an assumption of homogeneity concerning the internal states of both the nodes and their relations. Heterogeneity is nevertheless provided by the fact that two different nodes do not have necessarily the same number of directed links. The direction of the directed edges also adds heterogeneity in the relations between nodes.

At time $t = 0$, we describe the initial infectious state of each node by the initial infection vector $\mathbf{x}(0) = (x_1(0), \ldots, x_i(0), \ldots x_N(0)) \in \mathbb{R}^N$ where $x_i(0)$ is the initial infectious state (total number of infected units) of the node i. As for all $i \in [\![1 : N]\!]$ $x_i(0) \in \mathbb{R}$, the following results will remain true whatever the initial infectious states of the nodes. Similarly, we define the infection vector at time t as $\mathbf{x}(t) = (x_1(t), \ldots, x_i(t), \ldots, x_N(t)) \in \mathbb{R}^N$

In the discrete SIS_c model, the relationship between $\mathbf{x}(t+T)$ and $\mathbf{x}(t)$, is given by the following set of N difference equations:

$$\forall i \in [\![1 : N]\!]\ x_i(t+T) = \sum_{j=1}^{N} B_{ij} x_j(t) \tag{3}$$

with **B** the transpose of **A** ($\mathbf{B} = \mathbf{A}^{Tr}$) and $T \in \mathbb{R}^+$ the discrete time step of the system (e.g. for a given set of simulations and parameters: a month ($T = 1$), a fortnight ($T = 1/2$), a week ($T = 1/4$); for another given set of simulations: a year ($T = 1$), a month ($T = 1/12$), etc.).

Relationship Eq. (3) implies that the infectious state of any node $x_i(t)$ is part of the set $\mathbb{R} \cup \{+\infty\}$.

At the non-trivial (non-zero) equilibrium, and for a large enough value of t, any $x_i(t)$ is by definition a constant nonnegative real number, with at least one $x_i(t)$ being strictly positive. The non-trivial equilibrium is simply termed the equilibrium throughout. Under equilibrium, and with a large enough t, any $x_i(t)$ is equal to 0, which correspond to the trivial equilibrium. Eventually, above equilibrium, some $x_i(t)$ diverge to $+\infty$ when $t \to +\infty$, which mean that the maximal infectious state of a node has a priori no upper limit above equilibrium. While unrealistic, this feature of our model is unproblematic as: (i) there is no saturation effect at the beginning of an epidemic, (ii) we are solely interested in the threshold conditions, i.e. values of p_t and p_p for which the system is at equilibrium.

Nevertheless, to solve simply that problem, we could also make sure that $x_i(t)$ does not diverge. For the latter condition to be satisfied, a carrying capacity $q_i \in \mathbb{R}^+$ can be associated with any node i through an alternative formulation of Eq. (3):

$$\forall i \in [\![1 : N]\!]\ x_i(t+T) = p_p x_i(t) + \left[\sum_{j=1}^{N} (1-\delta_{ij}) B_{ij} x_j(t)\right]\left[1 - \frac{x_i(t)}{q_i}\right] \tag{3q}$$

where δ_{ij} is the Kronecker delta defined as: $\begin{cases} \delta_{ij} = 1 \text{ when } i = j \\ \delta_{ij} = 0 \text{ when } i \neq j \end{cases}$

As there is no saturation effect at the beginning of an epidemic, we have $x_i(t) \ll q_i(t)$ and we can approximate (3q) by (3). Hence our key results presented in the following stand. Such a refinement would lead to a less analytically tractable model and we therefore only use Eq. (3) in the following.

Eq. (3) can be written in a more compact form:

$$\forall k \in \mathbb{N}\ \mathbf{x}^{Tr}(kT + T) = \mathbf{B} \times \mathbf{x}^{Tr}(kT) \tag{4}$$

Eq. (4) leads recursively to:

$$\forall k \in \mathbb{N}\ \mathbf{x}^{Tr}(k.T) = \mathbf{B}^{k} \times \mathbf{x}^{Tr}(0) \tag{5}$$

2.2.2. Overview of the continuous SIS_c model

When the discrete time step becomes very small ($T \to dt$), we obtain the N differential equations describing the continuous SIS_c model by using Eq. (3):

$$\forall t \in \mathbb{R}\ \forall i \in [\![1 : N]\!]\ \frac{dx_i(t)}{dt} = -\frac{(1-p_p)}{dt} x_i(t) + \sum_{j=1}^{N} \frac{(1-\delta_{ij})B_{ij}}{dt} x_j(t) \tag{6}$$

Notice that $\frac{(1-p_p)}{dt}$ and $\frac{(1-\delta_{ij})B_{ij}}{dt} = \frac{p_t}{dt}$ can be interpreted as rates of recovery and transmission respectively.

2.3. Equilibrium state of the discrete SIS_c model

What are the threshold values of p_t, and p_p, i.e. what are the paired values of (p_t^*, p_p^*) for which the SIS_c model described by (3) is at equilibrium?

2.3.1. First fundamental relationship at equilibrium

At equilibrium and for t large enough, we have by definition the relationship:

$$\forall i \in [\![1 : N]\!]\ x_i(t+T) = x_i(t) = x_i^* \in \mathbb{R} \tag{7}$$

Since $\forall i \in [\![1 : N]\!]\ B^*(i,i) = p_p^*$, we deduce from Eqs. (3) and (7) that:

$$\forall i \in [\![1 : N]\!]\ \sum_{j=1}^{N} (B^*(i,j) - p_p^* \delta_{ij}) x_j^* = (1 - B^*(i,i)) x_i^* \tag{8}$$

Assuming $p_p^* \neq 1$, Eq. (8) leads to the relationship:

$$\forall i \in [\![1 : N]\!]\ \frac{1}{1-p_p^*} \sum_{j=1}^{N} (B^*(i,j) - p_p^* \delta_{ij}) x_j^* = x_i^* \tag{9}$$

Eq. (9) leads to the more compact equation:

$$\mathbf{C}^* \times \mathbf{x}^{*Tr} = \mathbf{x}^{*Tr} \tag{10}$$

where $\mathbf{x}^* = (x_1^*, \ldots, x_i^*, \ldots, x_N^*) \in \mathbb{R}^N$ and $\mathbf{C}^* = \frac{\mathbf{B}^* - p_p^* \mathbf{I}}{1-p_p^*}$

By setting $\mathbf{D}^* = \frac{(\mathbf{B}^* - p_p^* \mathbf{I})^{Tr}}{p_t^*}$ we have:

$$\mathbf{C}^{*Tr} = \frac{1}{1-p_p^*} (\mathbf{B}^* - p_p^* \mathbf{I})^{Tr} = \frac{p_t^*}{1-p_p^*} \mathbf{D}^* \tag{11}$$

With $p_t^* \neq 0$, Eqs. (9) and (11) leads to the following statement (after transposing):

$$\mathbf{x}^* \times \mathbf{D}^* = \frac{1-p_p^*}{p_t^*} \mathbf{x}^* \tag{12}$$

Moreover,

$$\mathbf{D}^* = \frac{(\mathbf{B}^* - p_p^*\mathbf{I})^{\mathrm{Tr}}}{p_t^*} = \frac{(\mathbf{A}^* - p_p^*\mathbf{I})}{p_t^*} \tag{13}$$

By construction of the adjacency matrix and as p_t and p_p are intrinsic parameters of the SIS_c model, we have additionally the relationship:

$$\forall (p_t, p_p) \in]0,1] \times [0,1] \ \frac{\mathbf{A} - p_p\mathbf{I}}{p_t} = \mathbf{\Psi} \tag{14}$$

with $\mathbf{\Psi}$ being the unique $\mathbf{A}$ matrix with $p_t = 1$ and $p_p = 0$.

Specifically, we have:

$$\forall (p_t, p_p) \in]0,1] \times [0,1] \ \frac{\mathbf{A}^* - p_p^*\mathbf{I}}{p_t^*} = \mathbf{\Psi} \tag{15}$$

From Eqs. (13) and (15), we see that $\mathbf{D}^* = \mathbf{\Psi}$. We therefore deduce from Eq. (12) the first fundamental relationship at equilibrium:

$$\boxed{\mathbf{x}^* \times \mathbf{\Psi} = \tfrac{1-p_p^*}{p_t^*}\mathbf{x}^*} \tag{16}$$

Transposing Eq. (16) shows straightforwardly that $\lambda = \frac{1-p_p^*}{p_t^*}$ is an eigenvalue of $\mathbf{\Psi}^{\mathrm{Tr}}$. Hence, λ is an eigenvalue of $\mathbf{\Psi}$.

This result is fundamental, because it shows that the SIS_c equilibrium can be easily deduced from the "natural" binary adjacency matrix $\mathbf{\Psi}$.

2.3.2. Second fundamental relationship at equilibrium

2.3.2.1. Lemma λ is the leading eigenvalue of $\mathbf{\Psi}$. Let λ be the eigenvalue of $\mathbf{\Psi}$ of largest magnitude. According to Perron's Theorem [42], λ is real and nonnegative and by definition $\lambda \leqslant \Lambda$.

As λ is an eigenvalue of $\mathbf{\Psi}$ we know that

$$\exists z \in \mathbb{R}^N \setminus \{0\} \ \mathbf{\Psi} \times z^{\mathrm{Tr}} = \Lambda z^{\mathrm{Tr}} \tag{17}$$

Eqs. (15) and (17) lead us to

$$\mathbf{A}^* \times \mathbf{z}^{\mathrm{Tr}} = (\Lambda p_t^* + p_p^*)\mathbf{z}^{\mathrm{Tr}} \tag{18}$$

Transposing Eq. (18) leads to

$$\mathbf{z} \times \mathbf{B}^* = (\Lambda p_t^* + p_p^*)\mathbf{z} \tag{19}$$

Transposing Eq. (19) shows immediately that $(\Lambda p_t^* + p_p^*)$ is an eigenvalue of $\mathbf{B}^{*\mathrm{Tr}}$. Hence, $(\Lambda p_t^* + p_p^*)$ is an eigenvalue of $\mathbf{B}^*$.

According to MacCluer (2000), an iteration scheme $\mathbf{x}^{\mathrm{Tr}}(kT) = \mathbf{B}^{\mathrm{k}} \times \mathbf{x}^{\mathrm{Tr}}(0), \mathrm{k} \in \mathbb{N}$ (see (5)) converges for every initial state $\mathbf{x}^{\mathrm{Tr}}(0)$ if and only if the eigenvalues λ of $\mathbf{B}$ distinct from 1 have modulus $|\lambda| < 1$, and, if $\lambda = 1$ is an eigenvalue, its eigenspace is of full rank; i.e., its rank equals the multiplicity of the root $\lambda = 1$ in the characteristic equation $P(\lambda) = \det(\lambda\mathbf{I} - \mathbf{B})$.

As $\mathbf{B}^*$ converges, any eigenvalue λ of $\mathbf{B}^*$ satisfies $|\lambda| \leqslant 1$. Peculiarly, $(\Lambda p_t^* + p_p^*) \leq 1$, which leads to $\Lambda \leq \frac{1-p_p^*}{p_t^*} = \lambda$.

Therefore λ is equal to Λ, the leading eigenvalue of $\mathbf{\Psi}$. As $\lambda = \Lambda \geqslant 0$, $p_p^* \leq 1$ which is always true as by definition p_p^* is part of [0 1].

2.3.2.2. Second fundamental relationship at equilibrium. We can now express p_p^* as a function of p_t^* to obtain the second fundamental relationship at equilibrium:

$$\boxed{p_p^* = 1 - \Lambda p_t^*} \tag{20}$$

For $p_t^* = 0, p_p^* = 1$ and for $p_p^* = 0, p_t^* = \frac{1}{\Lambda}$.

These points determine a line describing the value of p_p^* as a function of p_t^*, thus demonstrating analytically the empirical results of Pautasso and Jeger [29].

2.4. Algorithm for computing the threshold conditions

Based on the analytical results of this paper, the following algorithm will always find the threshold pairings (p_t^*, p_p^*) of any network, i.e. the infinite set of $\mathbf{A}^*$ corresponding to the "natural" binary matrix $\mathbf{\Psi}$ (which can be easily known from real-world data and observations).

1. Describe $\mathbf{\Psi}$ or generate it with any algorithm designed to build a directed network, with zeros in the diagonal (which means that a node is not connected to itself) and, everywhere else, either zeros or ones depending on whether directed links exists between vertices (which means that the links are unweighted). This type of directed network is very general and useful, as stated in Restrepo et al. [43] and references therein.
2. Compute Λ, the leading eigenvalue of $\mathbf{\Psi}$. In relationship $(1 - p_{p^*})/(p_{t^*}) = \Lambda$, set p_p^* arbitrarily to 0. The value of p_t^* is then $\frac{1}{\Lambda}$.
3. Using the second fundamental relationship at equilibrium (Eq. (20)), deduce the infinite set of threshold couples (p_t^*, p_p^*) corresponding to $\mathbf{\Psi}$.
4. With (5), check that a computed (p_t^*, p_p^*) of interest actually corresponds to an equilibrium: $\mathbf{B}^{*k}$ must be equal to $\mathbf{B}^{*k+1}$ for $k \in \mathbb{N}$ large enough. Empirically, there is a non-periodic convergence of $\mathbf{B}^*$ in almost all cases. We need to add this fourth step because we cannot find the conditions on $\mathbf{A}$ for an equilibrium to occur.

3. Applications to disease control

The SIS_c model developed and analysed above shows that the leading eigenvalue of the adjacency matrix of any directed network yields the epidemic threshold. The leading eigenvalue is therefore a key indicator, providing a link between topology and disease spread and can therefore be used, as shown below, to design effective control policies if the topology of the network is known.

3.1. Theoretical considerations

Assuming that directed networks lack an edge-degree correlation (a node with a given degree is connected uniformly at random to nodes with either low or high degrees), leading eigenvalues can be approximated by a mean field approximation (MFA) given by (21), as explained in Restrepo et al. [44]. We assume here that the absence of degree correlation condition necessary for applying the MFA is satisfied.

With the MFA [44], the leading eigenvalue Λ of a network is given by:

$$\Lambda \approx < d^{in} d^{out} > / < d > \tag{21}$$

where $< \cdot >$ denotes an average over nodes, $d^{in,i}$ (resp, $d^{out,i}$) is the in-degree (resp. out-degree) of node i and $<d> = <d^{in}> = <d^{out}>$ (the averages of the in and out-degrees have to be the same). From Eq. (20), we hence have:

$$\Lambda = (1 - p_p^*)/p_t^* \approx < d^{in} d^{out} > / < d > \tag{22}$$

Interestingly, Eq. (22) is highly similar to Eq. (5) in Schwartz et al. [45] where the percolation threshold is given by:

$$q_c = < d > / < d^{in} d^{out} > \tag{23}$$

with q_c being the threshold fraction of nodes retained (i.e. not removed from the network). Thus, equilibrium conditions of the SIS_c model and threshold conditions of percolation processes are closely related under the MFA conditions.

Moreover, Eq. (22) shows that correlations between in and out-degrees ($<d^{in}d^{out}>$) have a positive effect on Λ and consequently a negative effect on the values of p_t^* and p_p^*: the epidemic threshold will decrease if the correlations between links in and out of nodes increase. This theoretical result is in perfect agreement with the empiric results previously reported by Moslonka-Lefebvre et al. [30].

Moreover, (22) suggests that an effective control of disease spread should reduce high correlations between links in and links out by e.g. targeting wholesalers, as shown empirically by Pautasso et al. [25] and in the figures below.

However Eq. (22) stands only under the MFA assumptions that are restrictive and should be, therefore, confronted to empirical results on a case-by-case basis.

3.2. Application to the UK horticultural trade network

3.2.1. Realistic directed trade networks of plant nurseries

To test and compare various control strategies, five large-size realistic directed networks were generated with the model of UK plant nurseries and other key horticultural market units described in Harwood et al. [46]. The analyses described in Table 1 and Fig. 1 show that these networks tend to be highly heterogeneous (non-negligible probability of observing highly connected market players, i.e. hubs, as shown in Fig. 1) and have a highly significant positive correlation between links in and links out (Table 1). While non scale-free (see Table 1), the networks of Harwood et al. [46] are therefore similar to scale-free networks for which there is a high correlation coefficient between in- and out-degree of nodes as described in Moslonka-Lefebvre et al. [30].

In the five networks described above, nodes are classified as producers, wholesalers or retailers as explained in Pautasso et al. [25]. More precisely, the boundaries between these three hierarchical categories can be set according to the following quantity:

$$\Delta_i = \frac{d_i^{out} - d_i^{in}}{(d_i^{out} + d_i^{in})/2}\% \qquad (24)$$

where d_i^{out} is the number of outgoing links from node i and d_i^{in} is the number of incoming links to node i. Given that $-d_i^{in} \leq d_i^{out} - d_i^{in} \leq d_i^{out}$, that $0 \leq \frac{d_i^{in}}{d_i^{in}+d_i^{out}} \leq 1$ and that $0 \leq \frac{d_i^{out}}{d_i^{in}+d_i^{out}} \leq 1$, it is apparent that $-200\% \leqslant \Delta_i \leqslant +200\%$.

Producers were defined as nodes with Δ_i higher than or equal to 20 (60 or 100)%. Retailers were nodes with Δ_i lower than or equal to −20 (−60 or −100)%. Wholesalers were the remaining nodes (Table 2). Producers can be visualized as mainly exporters, retailers as mainly importers, and wholesalers as nodes with no preponderance of either export or import, thus linking the other two categories [25]. The notions of "export" and "import" are related to node topological properties and are not necessarily referring to trade between countries, although the same model could also be applied to trade among countries.

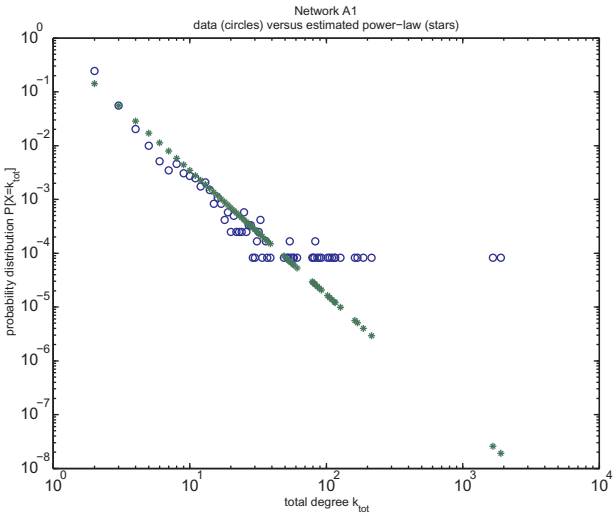


Fig. 1. The total-degree (denoted k_{tot}, which is the total number of incoming and outgoing links of each node) distribution and the associated power-law fit of network A1.

3.2.2. Control strategies

We focus on preventive control strategies where various numbers of nodes are disconnected from the network prior infection. In the real world, such a disconnection can e.g. correspond to a nursery which is put under quarantine by a safety agency. Five strategies are compared:

- A "random strategy" where nodes are targeted randomly.
- A "total-degree strategy" where highly connected nodes are targeted as a matter of priority. The order of targeting is determined by the total-degree of the remaining nodes.
- A "producer strategy" where producers are targeted randomly.
- A "wholesaler strategy" where wholesalers are targeted randomly.
- A "retailer strategy" where retailers are targeted randomly.

3.2.3. Results

As shown by the stronger increases in the slopes of the linear regressions of the probability of transmission at equilibrium p_t^* as of function of the number of nodes removed n_r (Table 3 and Fig. 2), the best control strategy consists in targeting large-size nurseries, i.e. hubs. However targeting only large-size nurseries might lead to market imbalances due to a supply shortages and the resultant price increases for buyers.

Table 1
Key properties of realistic directed networks. Each of the five networks was generated with the model of plant nurseries described in Harwood et al. [46]. N is the number of nodes, l is the total number of directed links, <C> is the mean clustering (which quantifies the likelihood for two neighbour nodes of any given node to be connected with each other), Λ_A is the leading eigenvalue of the adjacency matrix **A**, γ is the coefficient of the power law (as regard to the total-degree distribution), SE is the standard error for γ, PL stands for 'power-law' and states whether the degree distribution of the network considered follows a power-law depending on the associated p-value (derived from a χ^2 goodness-of-fit test performed with function *chi2gof* with the Statistics Toolbox from Matlab), $C_{in/out}$ is the correlation coefficient between links in and links out, i.e. the likelihood for any node to have both many links in and many links out. Parameters were estimated using Matlab.

Name	N	L	<C>	Λ_A	γ	SE	PL (p-value)	$C_{in/out}$ (p-value)
A1	12100	14206	0.001	2.62	2.31	0.013	No ($p = 0.15$)	0.46 ($p < 0.001$)
A2	12046	14169	0.001	2.57	2.31	0.013	No ($p = 0.15$)	0.48 ($p < 0.001$)
A3	12088	14220	0.001	2.60	2.31	0.013	No ($p = 0.15$)	0.47 ($p < 0.001$)
A4	12075	14204	0.001	2.60	2.31	0.013	No ($p = 0.15$)	0.47 ($p < 0.001$)
A5	12067	14193	0.001	2.57	2.31	0.013	No ($p = 0.29$)	0.45 ($p < 0.001$)

Table 2
Number (proportion) of hierarchical categories among the five networks analysed as a function of $|\Delta_i|$.

Name	$\|\Delta_i\|$%	Producers	Wholesalers	Retailers
A1	20	4581 (0.38)	2758 (0.23)	4761 (0.39)
	60	4522 (0.37)	2919 (0.24)	4668 (0.39)
	100	4488 (0.37)	3625 (0.30)	3987 (0.33)
A2	20	4545 (0.38)	2723 (0.23)	4778 (0.39)
	60	4491 (0.37)	2872 (0.24)	4683 (0.39)
	100	4448 (0.37)	3612 (0.30)	3986 (0.33)
A3	20	4594 (0.38)	2718 (0.22)	4776 (0.40)
	60	4537 (0.38)	2864 (0.24)	4687 (0.38)
	100	4492 (0.37)	3622 (0.30)	3974 (0.33)
A4	20	4568 (0.38)	2730 (0.23)	4777 (0.39)
	60	4510 (0.37)	2868 (0.24)	4697 (0.39)
	100	4462 (0.37)	3630 (0.30)	3983 (0.33)
A5	20	4564 (0.38)	2745 (0.23)	4758 (0.39)
	60	4507 (0.37)	2888 (0.24)	4672 (0.39)
	100	4452 (0.37)	3646 (0.30)	3969 (0.33)

Table 3
Targeting strategies on network A1 with $|\Delta_i| = 20,60,100\%$. $C[n_r/p_t^*]$ is the correlation coefficient between the number of nodes removed, denoted by n_r, and the epidemic threshold p_t^*. $C[n_r/p_t^*]$ quantifies how strong n_r affects p_t^*. The slope of the linear regression of p_t^* as a function of n_r (intercept set to p_t^* when $n_r = 0$) with its associated R^2 is also provided. Parameters were estimated using R (functions *cor. test* and *lm*).

$\|\Delta_i\|$	Strategy	$C[n_r/p_t^*]$ (p-value)	Slope (p-value)	R^2
20%	Random	0.45 ($p < 0.001$)	0.00005 ($p < 0.001$)	0.47
	Total-degree	0.35 ($p = 0.005$)	0.00344 ($p < 0.001$)	0.64
	Producers	0.39 ($p < 0.001$)	0.00008 ($p < 0.001$)	0.41
	Wholesalers	0.94 ($p < 0.001$)	0.00011 ($p < 0.001$)	0.95
	Retailers	0.84 ($p < 0.001$)	0.00002 ($p < 0.001$)	0.90
60%	Random	0.50 ($p < 0.001$)	0.00005 ($p < 0.001$)	0.54
	Total-degree	0.38 ($p = 0.002$)	0.00375 ($p < 0.001$)	0.68
	Producers	0.45 ($p < 0.001$)	0.00008 ($p < 0.001$)	0.42
	Wholesalers	0.95 ($p < 0.001$)	0.00011 ($p < 0.001$)	0.95
	Retailers	0.86 ($p < 0.001$)	0.00002 ($p < 0.001$)	0.91
100%	Random	0.56 ($p < 0.001$)	0.00005 ($p < 0.001$)	0.61
	Total-degree	0.35 (p=0.005)	0.00331 ($p < 0.001$)	0.75
	Producers	0.52 ($p < 0.001$)	0.00009 ($p < 0.001$)	0.53
	Wholesalers	0.90 ($p < 0.001$)	0.00014 ($p < 0.001$)	0.90
	Retailers	0.83 ($p < 0.001$)	0.00005 ($p < 0.001$)	0.90

As shown by the slopes and the total variation explained in Table 3, the second best strategy consists in targeting wholesalers randomly. This result confirms previous findings suggesting that in networks without the presence of hubs (local, random and small-world networks), control is better targeted towards wholesalers, i.e. the middle-tier of trade networks [25]. Targeting wholesalers randomly, while still epidemiologically efficient, is less likely to produce market imbalances than the "hub strategy" and could be therefore preferred for implementing control polices.

Of course, the strategies described above are not mutually exclusive and targeting large-size wholesalers could be more efficient than the individual strategies presented here.

4. Discussion and conclusions

Network theory is an important development in mathematical biosciences, as it has a range of key applications in, e.g., epidemiology [17,47]. Most epidemiological models in networks, however, have used the standard assumption of either susceptible or infected individuals. In some cases (e.g. the spread of *P. ramorum* and similar emerging pathogens in plant trade networks, the exchange of seeds among farmers, the trade of food commodities among nations), a continuum in the infection status of nodes can better capture the reality of epidemics in networks. In this paper, a Susceptible-Infected-Susceptible model along-a-continuum in the infection status (SIS_c) was presented, using as a case study directed networks and two parameters governing the epidemic process (probability of infection persistence (p_p) and of infection transmission (p_t)).

The previously empirically reported linear epidemic threshold in a plot of p_p as a function of p_t [29] was derived analytically and was shown to be governed by the leading eigenvalue of the network adjacency matrix Λ. Our results confirm the view that Λ is a key factor driving epidemics and percolation processes [48]. Also the previously observed negative correlation between the epidemic threshold and the correlation between links -in and links-out of nodes [30] was justified analytically under the assumption that edges degrees are uncorrelated. In addition, a simple algorithm to calculate the threshold conditions of any directed network was introduced. Finally, a preventive control strategy based on targeting market hierarchical categories such as producers, wholesalers and retailers was presented and applied to a realistic model of the UK horticultural trade network. The best strategy identified consists in targeting hubs, i.e. nodes with high number of links. The second best strategy identified consists in targeting wholesalers randomly, thus confirming the empirical findings obtained for local, random and small-world networks [25]. It follows that these findings on best control recommendations are likely to be independent of the network structure.

Interestingly, the epidemiological implications of the SIS_c model are likely to be different from the ones of the corresponding binary SIS model. In binary models, infectious nodes are blocked into an absorbing state: once infected, nodes cannot be re-infected, which is likely to slow down epidemics. On the contrary, in the SIS_c model, infectious nodes can be cumulatively infected, which is likely to make disease spread less affected by clusters of infected individuals. We can hence reasonably postulate that epidemic thresholds in the SIS_c model are lower than thresholds obtained from binary infectious models, other things being equal.

Whilst we consider the generalisations of the model to be reasonable, there are a number of caveats. The real world is not fully covered by this model when practical management is introduced. For instance, to make the SIS_c model more realistic (but less parsimonious), at least two epidemiological categories would need to be added: D (detected plants) and R (removed plants). The likelihood of disease detection is low when only a small fraction of the plants within a given nursery is infected. However, with an increasing disease prevalence, the disease is more likely to be detected (I to D), which may or not lead to control efforts (D to R), given the issue of unreported outbreaks. If the outbreak is reported to phytosanitary authorities, detected infected plants are removed and replaced by new susceptible plants (R to S), with the additional complication of quarantined neighbouring plants. Hence, the SIS_c model is valid at the beginning of an epidemic only and a $SIDRS_c$ model should be implemented instead when studying disease spread and control over the long term.

Nevertheless, p_p roughly incorporates detection and removal of infected plants, so that the SIS_c approach is still a useful first approximation of the system also later on in the epidemic. Moreover, detection, as part of a control strategy designed to mitigate disease spread is likely to miss some infectious plants. Consequently, it could be relevant to relate targeting of nodes to the degree of infection in such nodes. Furthermore, potential persistence of *P. ramorum* in the substrate at infected sites after eradication [49] may mean that the level of infection does not reach zero, even when all infectious plants are removed, and replaced by new susceptibles. This highlights an advantage of the SIS_c-related approaches, where the effective proportion of infectious individuals at a node may be less than one.

Another extension of interest is the case where p_p would be a function of the number of links. It can be assumed that p_p and

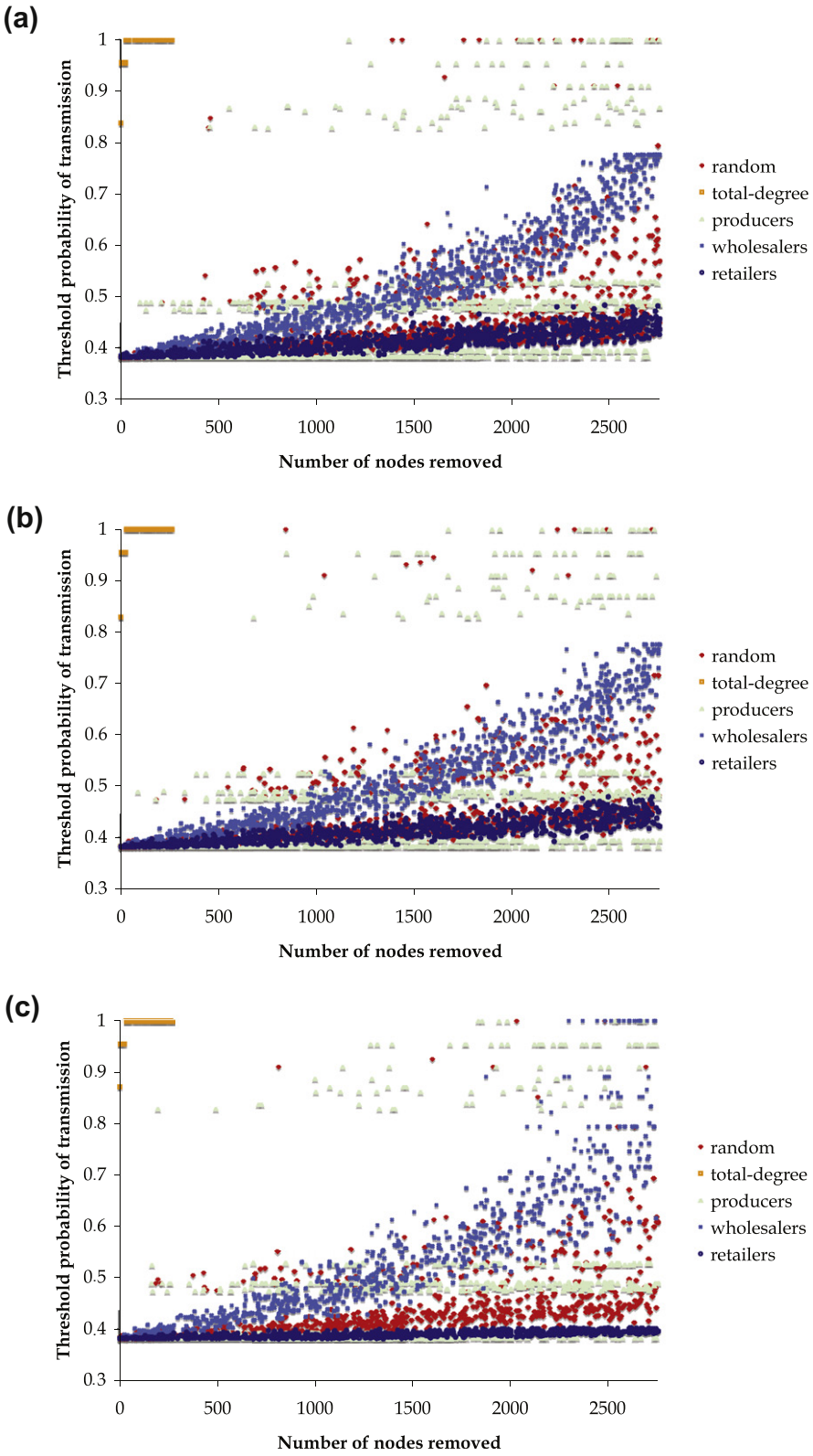


Fig. 2. Targeting strategies on network A1 with (a) $|\Delta_i| = 20\%$, (b) $|\Delta_i| = 60\%$ and (c) $|\Delta_i| = 100\%$.

the degree are negatively correlated if a more connected node tends to invest more resources for control. On the contrary, without control, we can expect a positive correlation between node connectivity and its associated p_p. Studying such trade-offs to find optimal levels of protection depending on the size/connectivity of the nodes is a key open question. To investigate this open issue, it is necessary to distinguish in p_p the intrinsic probability of recovery from control intensity.

Besides key applications in plant epidemiology [47] to e.g. avoid future plant health issues such as the rapid mortality due to *P.*

ramorum of Japanese larch plantations in Britain [20] and achieve a sustainable production of biofuel crops, which are at risk from new plant pathogens [50,51], the SIS_c model could be applied to other key dynamical systems with similar characteristics. Examples include invasive plants [52–54], diseases infecting cattle through trade networks [55,56], metapopulation dynamics (e.g. hospitals sharing patients [57]), seed exchange networks [58], international food trade [59] and the spread of rumours with partial memory recall [60]. Additionally, if the probability of persistence is set to 1, the SIS_c model becomes a SI_c model that can be useful to describe diseases where transmission is highly dependent on intra-host dynamics. An example of such a disease is HIV, as the probability of transmission is highly correlated with an individual's viral load [61]. In our model, the fraction of infection (I_c) can be interpreted as a viral load.

Further theoretical work could extend the SIS_c approach to other types of epidemiological models and identify the role of the leading eigenvalue of the adjacency matrix on dynamical systems spreading on various classes of networks (e.g. weighted and multipartite).

Acknowledgments

Many thanks to O. Holdenrieder, M. Parry, L. Pellis, M.W. Shaw, F. van den Bosch, E. Vergu and X. Xu for discussions and insights, and to two anonymous reviewers for helpful comments on a previous draft. This study was funded by the Department for Environment, Food and Rural Affairs (DEFRA) and the Rural Economy and Land Use Programme (RELU), (UK), and the Île-de-France Regional Council under the MIDEM project in the framework DIM Malinf and the French Ministries in charge of Agriculture and Sustainable Development.

References

[1] G.S. Cumming, Ö. Bodin, H. Ernstson, T. Elmqvist, Network analysis in conservation biogeography: challenges and opportunities, Diversity and Distributions 16 (2010) 414.
[2] C.J. Garroway, J. Bowman, D. Carr, P.J. Wilson, Applications of graph theory to landscape genetics, Evolutionary Applications 1 (2008) 620.
[3] E. Thebault, C. Fontaine, Stability of ecological communities and the architecture of mutualistic and trophic networks, Science 329 (2010) 853.
[4] J.K. Adou, Y. Billaud, D.A. Brou, J.P. Clerc, J.L. Consalvi, A. Fuentes, A. Kaiss, F. Nmira, B. Porterie, L. Zekri, N. Zekri, Simulating wildfire patterns using a small-world network model, Ecological Modelling 221 (2010) 1463.
[5] M. Keeling, The implications of network structure for epidemic dynamics, Theoretical Population Biology 67 (2005) 1.
[6] R.M. May, Network structure and the biology of populations, Trends in Ecology & Evolution 21 (2006) 394.
[7] P. Trapman, On analytical approaches to epidemics on networks, Theoretical Population Biology 71 (2007) 160.
[8] R.M. Anderson, R.M. May, Understanding the AIDS pandemic, Scientific American 266 (1992) 58.
[9] B.T. Grenfell, A.P. Dobson, Ecology of Infectious Diseases in Natural Populations, Cambridge University Press, Cambridge, New York, 1995.
[10] M.J. Keeling, K.T. Eames, Networks and epidemic models, Journal of the Royal Society Interface 2 (2005) 295.
[11] M.D.F. Shirley, S.P. Rushton, The impacts of network topology on disease spread, Ecological Complexity 2 (2005) 287.
[12] V. Colizza, M. Barthelemy, A. Barrat, A. Vespignani, Epidemic modeling in complex realities, Comptes Rendus Biologies 330 (2007) 364.
[13] F. Brauer, An introduction to networks in epidemic modeling, Mathematical Epidemiology 2008 (1945) 133.
[14] K.T. Eames, M.J. Keeling, Contact tracing and disease control, Proceedings of the Royal Society of London. B: Biological Sciences 270 (2003) 2565.
[15] K.T. Eames, Contact tracing strategies in heterogeneous populations, Epidemiology and Infection 135 (2007) 443.
[16] I.Z. Kiss, D.M. Green, R.R. Kao, The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing, Journal of the Royal Society Interface 5 (2008) 791.
[17] M.J. Jeger, M. Pautasso, O. Holdenrieder, M.W. Shaw, Modelling disease spread and control in networks: implications for plant sciences, New Phytologist 174 (2007) 279.
[18] C.A. Gilligan, An epidemiological framework for disease management, Advances in Botanical Research, Academic Press, 2002.
[19] D.M. Rizzo, M. Garbelotto, E.A. Hansen, *Phytophthora ramorum*: Integrative research and management of an emerging pathogen in California and Oregon forests, Annual Review of Phytopathology 43 (2005) 309.
[20] C. Brasier, J. Webber, Sudden larch death, Nature 466 (2010) 824.
[21] K. Dehnen-Schmutz, J. Touza, C. Perrings, M. Williamson, The horticultural trade and ornamental plant invasions in Britain, Conservation Biology 21 (2007) 224.
[22] K. Dehnen-Schmutz, J. Touza, Plant invasions and ornamental horticulture: pathway, propagule pressure and the legal framework, in: J.A. Teixeira da Silva (Ed.), Floriculture Ornamental and Plant Biotechnology: A dva nces and Topical Issues, Global Science Books, Isleworth, UK, 2008, pp. 15–21.
[23] S. Sugiura, Species interactions-area relationships: biological invasions and network structure in relation to island area, Proceedings of the Royal Society B: Biological Sciences 277 (2010) 1807.
[24] C. Vacher, J.-J. Daudin, D. Piou, M.-L. Desprez-Loustau, Ecological integration of alien species into a tree-parasitic fungus network, Biological Invasions 12 (2010) 3249.
[25] M. Pautasso, X. Xu, M.J. Jeger, T.D. Harwood, M. Moslonka-Lefebvre, L. Pellis, Disease spread in small-size directed trade networks: the role of hierarchical categories, Journal of Applied Ecology 47 (2010) 1300.
[26] M.E.J. Newman, S.H. Strogatz, D.J. Watts, Random graphs with arbitrary degree distributions and their applications, Physical Review E 6402 (2001) 026118.
[27] L.A. Meyers, M.E.J. Newman, B. Pourbohloul, Predicting epidemics on directed contact networks, Journal of Theoretical Biology 240 (2006) 400.
[28] S.M. Park, B.J. Kim, Dynamic behaviors in directed networks, Physical Review E 74 (2006) 026114.
[29] M. Pautasso, M.J. Jeger, Epidemic threshold and network structure: The interplay of probability of transmission and of persistence in small-size directed networks, Ecological Complexity 5 (2008) 1.
[30] M. Moslonka-Lefebvre, M. Pautasso, M.J. Jeger, Disease spread in small-size directed networks: epidemic threshold, correlation between links to and from nodes, and clustering, Journal of Theoretical Biology 260 (2009) 402.
[31] L. Allen, A. Burgin, Comparison of deterministic and stochastic SIS and SIR models in discrete time, Mathematical Biosciences 163 (2000) 1.
[32] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society. A: Mathematical, Physical and Engineering Sciences 115 (1927) 700.
[33] A.W. Park, S. Gubbins, C.A. Gilligan, Invasion and persistence of plant parasites in a spatially structured host population, Oikos 94 (2001) 162.
[34] J. Segarra, M.J. Jeger, F. van den Bosch, Epidemic dynamics and patterns of plant diseases, Phytopathology 91 (2001) 1001.
[35] F. Brauer, The Kermack–McKendrick epidemic model revisited, Mathematical Biosciences 198 (2005) 119.
[36] J.S. Tepan, D. Hlubinka, Kermack–McKendrick epidemic model revisited, Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, 2007.
[37] P.E. Parham, B.K. Singh, N.M. Ferguson, Analytic approximation of spatial epidemic models of foot and mouth disease, Theoretical Population Biology 73 (2008) 349.
[38] L.J.S. Allen, An introduction to stochastic epidemic models, Mathematical Epidemiology 2008 (1945) 81.
[39] D. Hiebeler, Moment equations and dynamics of a household SIS epidemiological model, Bulletin of Mathematical Biology 68 (2006) 1315.
[40] S.H. Orzack, E. Sober, A critical-assessment of Levins's *The strategy of model-building in population biology*, Quarterly Review of Biology 68 (1993) (1966) 533.
[41] D. Watts, Small Worlds: The Dynamics of Networks Between Order and Randomness, Princeton University Press, 1999.
[42] C.R. MacCluer, The many proofs and applications of Perron's theorem, Siam Review 42 (2000) 487.
[43] J.G. Restrepo, E. Ott, B.R. Hunt, Characterizing the dynamical importance of network nodes and links, Physical Review Letters 97 (2006) 094102.
[44] J.G. Restrepo, E. Ott, B.R. Hunt, Approximating the largest eigenvalue of network adjacency matrices, Physical Review E 76 (2007) 056119.
[45] N. Schwartz, R. Cohen, D. ben-Avraham, A.L. Barabasi, S. Havlin, Percolation in directed scale-free networks, Physical Review E 66 (2002) 015104.
[46] T.D. Harwood, X. Xu, M. Pautasso, M.J. Jeger, M.W. Shaw, Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK, Ecological Modelling 220 (2009) 3353.
[47] M. Moslonka-Lefebvre, A. Finley, I. Dorigatti, K. Dehnen-Schmutz, T. Harwood, M.J. Jeger, X. Xiangming, O. Holdenrieder, M. Pautasso, Networks in plant epidemiology: from genes to landscapes, countries and continents, Phytopathology 101 (2011) 392.
[48] D. Chakrabarti, Y. Wang, C.X. Wang, J. Leskovec, C. Faloutsos, Epidemic thresholds in real networks, ACM Transactions on Information and System Security 10 (2008) 1.
[49] E.J. Fichtner, D.M. Rizzo, S.A. Kirk, J.F. Webber, Root infections may challenge management of invasive *Phytophthora spp.* in UK woodlands, Plant Disease 95 (2011) 13.
[50] A. Stewart, M. Cromey, Identifying disease threats and management practices for bio-energy crops, Current Opinion in Environmental Sustainability 3 (2011) 675.
[51] L.J. Thomson, A.A. Hoffmann, Pest management challenges for biofuel crop production, Current Opinion in Environmental Sustainability 3 (2011) 95.

[52] H.J. Albers, C. Fischer, J.N. Sanchirico, Invasive species management in a spatially heterogeneous world: Effects of uniform policies, Resource and Energy Economics 32 (2010) 483.
[53] S. Brunel, F. Petter, The EPPO decision-support scheme for pest risk analysis and invasive alien plants, Plant Protection Quarterly 25 (2010) 42.
[54] C.G. Chimera, C.E. Buddenhagen, P.M. Clifford, Biofuels: the risks and dangers of introducing invasive species, Biofuels 1 (2010) 785.
[55] S. Rautureau, B. Dufour, B. Durand, Vulnerability of animal trade networks to the spread of infectious diseases: a methodological approach applied to evaluation and emergency control strategies in cattle, France, Transboundary and Emerging Diseases 58 (2011) (2005) 110.
[56] S. Widgren, J. Frössling, Spatio-temporal evaluation of cattle trade in Sweden: description of a grid network visualization technique, Geospatial Health 5 (2010) 119.
[57] D.L. Smith, S.A. Levin, R. Laxminarayan, Strategic interactions in multi-institutional epidemics of antibiotic resistance, Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 3153.
[58] J. Enjalbert, J.C. Dawson, S. Paillard, B. Rhone, Y. Rousselle, M. Thomas, I. Goldringer, Dynamic management of crop diversity: from an experimental approach to on-farm conservation, Comptes Rendus Biologies 334 (2011) 458.
[59] M. Barigozzi, G. Fagiolo, D. Garlaschelli, Multinetwork of international trade:a commodity-specific analysis, Physical Review E 81 (2010) 046104.
[60] L.M.A. Bettencourt, A. Cintrón-Arias, D.I. Kaiser, C. Castillo-Chávez, The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models, Physica A 364 (2006) 513.
[61] D.P. Wilson, M.G. Law, A.E. Grulich, D.A. Cooper, J.M. Kaldor, Relation between HIV viral load and infectiousness: a model-based analysis, Lancet 372 (2008) 314.

Annexe D

Supporting information for : *Market analyses of livestock trade networks to inform the prevention of joint economic and epidemiological risk*

D.1 Supplementary materials and methods

D.1.1 Standard indicators of epidemiological risk : definitions and implications for epidemics

We define standard indicators to assess epidemiological risk for networks : the *proportion of agents belonging to the largest strongly connected component (LSCC)* and *betweenness centrality.*

The proportion of agents belonging to the LSCC is a standard proxy to assess both the probability of an outbreak and the epidemic final size [e.g. Meyers *et al.*, 2006; Rautureau *et al.*, 2012]. Formally, the LSCC is the largest set of agents that can be reached by any other agent by following the direction of links over a time period $\mathcal{T}$. The larger the fraction of agents belonging to the LSCC, the larger the epidemiological risk. Since we calculate the proportion of agents belonging to the LSCC for each market category, we can estimate the contribution of each market category to the global epidemiological risk.

The betweenness centrality of an agent a, denoted BC_a, is the fraction of shortest path lengths that passes through a. The shortest path length from agent i to agent j is the smallest number of directed links needed to reach j from i. The larger BC_a, the larger the epidemiological risk of agent a [e.g. Brandes, 2001; Buttner *et al.*, 2013b]. Formally, BC_a, here normalised to account for networks differing

in total number of agents, is given by :

$$\mathrm{BC}_a(\mathcal{T}) = \frac{1}{(N-1)(N-2)} \sum_{s \neq a \neq t} \frac{\sigma_{sat}(\mathcal{T})}{\sigma_{st}(\mathcal{T})} \ , \tag{D.1}$$

where $\sigma_{st}(\mathcal{T})$ and $\sigma_{sat}(\mathcal{T})$ are the number of shortest paths from agent s to agent t and the number of shortest paths from s to t passing through a during a time period $\mathcal{T}$ respectively. Since calculating shortest paths on large networks is computationally intensive, we approximate BC_a at order 3, i.e. we only consider shortest paths of length 3 or less.

D.1.2 Multiple-criteria decision analysis based on the average infection chain as a proxy for prevention-effectiveness

An alternative proxy to assess the effectiveness of preventive strategies : the average infection chain

We alternatively measure effectiveness, for differing strategies $\mathcal{S}$, as the *percentage decrease of the average infection chain* as function of the fraction of agents to target F_n. For a given time period $\mathcal{T} = [t_1; t_2]$, the out-going infection chain of a seller i (in-going infection chain of a demander j respectively) is the number of demanders that can be reached and hence infected from i by following temporally-compatible links (the number of sellers leading to j and from which j can hence be infected by following temporally-compatible links respectively). The out-going infection chain is calculated forwards in time (from t_1 to t_2), while the in-going infectious chain is calculated backwards in time (from t_2 to t_1). Both the out- and in- infection chains are epidemiological proxies Dube *et al.* [2008]; Noremark *et al.* [2011]. Following Lal Dutta *et al.* [2014], we focus here on the average infectious chain, which is indifferently given, for a given $\mathcal{T}$, by the the average in-going infection chain or the average out-going infection chain as calculated over all nodes. In contrast with the largest strongly connected component, the average infection chain takes into account the sequentiality of the dates at which exchanges occur over period $\mathcal{T}$.

For each strategy explored, the percentage decrease of the average infection chain is evaluated at increasing values of F_n and we carry out 100 runs of random targeting.

Since we keep track of time, we further distinguish two types of interventions for a given $\mathcal{S}$: *early interventions* where all the F_n agents are surveilled at the beginning of the time period $\mathcal{T}$ where the strategy is evaluated, and *partly delayed interventions* where the first half of the ordered F_n agents is targeted at t_1 and the second half at $(t_1 + t_2)/2$.

Selection of geographical subsets of the cattle dataset to assess percentage decrease of the average infection chain

Even when carfully optimised, the evaluation of average infection chains remains computationally-intensive Noremark & Widgren [2014]. We hence restrict their calculations to the cattle market for small geographical scales, here French *Départements* (Dpt). We consider, on the one hand, exchanges occurring within the Ille-et-Vilaine Dpt and with the rest of the world (Dpt 35 + ROW) and, on the one hand, exchanges occurring within the Saône-et-Loire Dpt and the rest of the world (Dpt 71+ ROW). Dpt 35 and Dpt 71 are the largest French Dpt for dairy and beef production respectively (Tables D.1-D.2).

TABLE D.1 – **Overview of two contrasted *Départements* (Dpt) of the French cattle market in year 2009 : Ille-et-Vilaine (Dpt 35 ; essentially dairy production) and Saône-et-Loire (Dpt 71 ; essentially beef production)**

quantity	Dpt 35 + ROW (dairy)	Dpt 71 + ROW (beef)
average number of animals exchanged per transaction	6.3	7.6
number of animals exchanged per year	600,889	597,918
number of transactions per year	95,051	78,385
number of agents (% in the LSCC)	6,853 (37%)	4,957 (63%)
number of strict suppliers	3,386	1,450
number of wholesalers	3,035	3,318
number of strict demanders	432	189

The LSCC is the largest strongly connected component.

TABLE D.2 – **Specialisation of farms in Dpt 35 and Dpt 71 in year 2009**

	% dairy farms	% mixed farms	% beef farms	% other farms	% trading agents
Dpt 35 (dairy)	**46.0**	28.0	9.6	16.0	0.4
Dpt 71 (beef)	3.7	17.5	**64.2**	13.5	1.1

D.2 Supplementary results

D.2.1 Temporal evolution of market categories

We investigate to which extent market categories can be considered stable over time. For cattle, we explore temporal stability of market categories at a yearly scale for period 2005-2009 (Fig. D.1). For swine, we only have access to deliveries occurring during year 2010, so we can not explore stability of market categories on a yearly scale.

Over years 2005-2009, flow polarity is more variable than flow share (Fig. D.1A-B). Relative summaries such as flow share are expected to be highly stable over time and for any temporal scale used to calculate them [Perra *et al.*, 2012]. Flow polarity (flow share) is roughly stable (extremely stable)

from year 2008 to year 2009 (Fig. D.1C-D). For both flow polarity and flow share, stability tends to increase for agents with negligible flow polarity and larger values of flow share, i.e. wholesalers-leaders (Fig. D.1E-F).

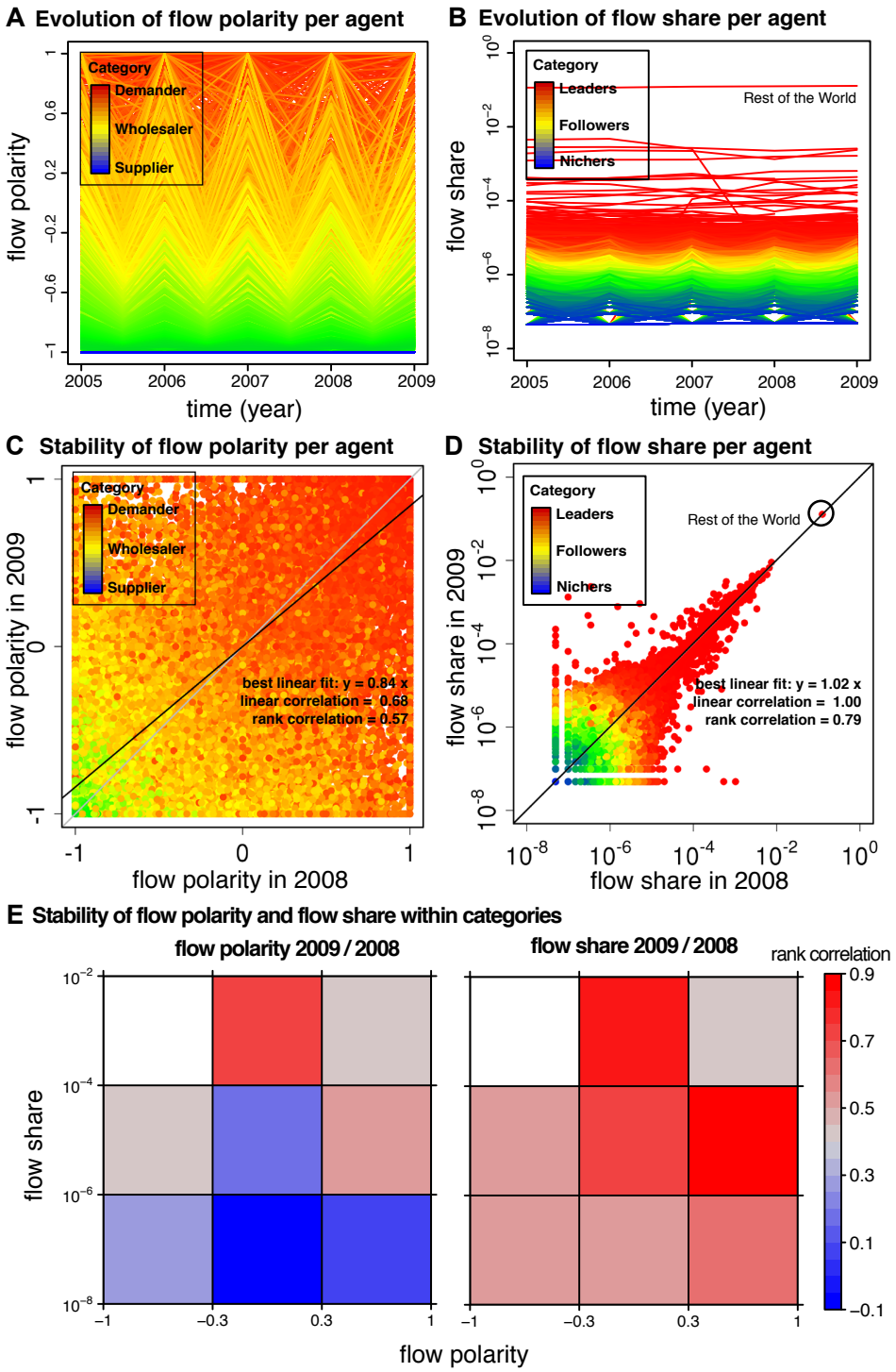


FIGURE D.1 – **Temporal stability of market categories for the cattle market. A**)-**B**) Annual evolution of flow polarity and flow share at the agent level from year 2005 to year 2009 for 5000 representative agents. Tenth of centiles in average annual values of flow polarity and flow share are used to cluster agents by color classes, from blue (minimal values) to red (maximal values). 5 agents are selected uniformly at random for each tenth of centile leading to a total of 5000 individual dynamics (about 2 % of total trajectories). **C**)-**D**) Stability of flow polarity and flow share at the agent level between years 2008 and 2009 for all agents. The color code is identical to (A-B). The black (grey) line represent the best linear fit $y = ax$ (the line $y = x$). Stability is assessed with Pearson's linear correlation coefficient and Spearman's rank correlation coefficient. **E**) Stability of flow polarity (left panel) and flow share (right panel) within market categories between years 2008 and 2009 for all agents. White cells correspond to empty categories. Market categories are defined here based on annual averages over five years and excluding exchanges with the rest of the world. Stability is assessed both with Pearson's linear and Spearman's rank correlation coefficients.

D.2.2 Standard indicators of epidemiological risk in relation with market categories

In addition to the proportion of agents belonging to the largest strongly connected component (Fig. 2 in the main text), we calculate the *betweenness centrality*, an additional standard indicator for epidemiological risk for both cattle and swine markets. We show how this measure scales with market categories, i.e. with flow polarity and flow share. For both cattle and swine, the betweenness centrality increases with flow share (Fig. D.2), implying a larger epidemiological risk associated with leaders compared with nichers. For a given flow share, the average betweenness centrality tends to be larger for agents with negligible flow polarity : wholesalers are probably stronger epidemiological drivers than suppliers, a finding in agreement with the theoretical results reported in Pautasso *et al.* [2010]; Moslonka-Lefebvre *et al.* [2012b].

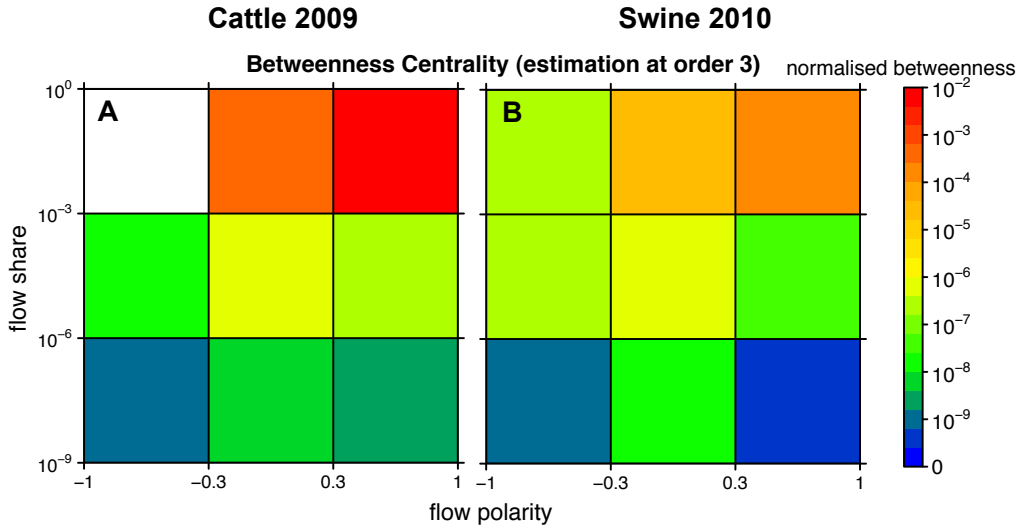


FIGURE D.2 – **Betweenness centrality as an indicator of epidemiological risk in the cattle and swine markets. A**)-**B**) Average normalised betweenness centrality per agent per market category. Betweenness centrality, here approximated at order 3, is considered a good indicator of epidemiological risk. The larger the betweenness centrality, the larger the risk. Market categories are defined as in Fig 2 of the main text. The white cell corresponds to an empty category.

D.2.3 Multiple-criteria decision analysis of contrasted preventive strategies based on the average infection chain

Measuring prevention-effectiveness based on the largest strongly connected component, we find, by carrying-out percolation experiments on the cattle market (Fig. 4 in the main text), that targeting suppliers-nichers first (SN strategy) induces lower relative flow-cost to the regulator and lower market distortions than targeting wholesalers-leaders first (WL strategy).

Here we test whether these findings stand when prevention-effectiveness is rather measured based on the average infection chain (Section D.1.2). We also explore impacts of delaying interventions. We focus our analyses on two contrasted geographical subsets of the French cattle market called Départements (Dpt) : Dpt 35 (essentially dairy production) and Dpt 71 (essentially beef production). We also consider, for each Dpt, exchanges with the Rest Of the World (ROW).

The results remain qualitatively unchanged compared to those reported in the main text for the cattle market (Fig. 4 in the main text) : the SN strategy outperforms the WL strategy as far as the relative flow-cost to the regulator and market distortions are concerned (Figs D.3,D.4,D.5,D.6). As expected, the WL strategy performs better than the SN strategy regarding relative agent-cost. Delaying interventions has a negligible effect on WL strategies and only affect SN strategies for very large values of prevention-effectiveness (Fig D.3 versus Fig D.4 and Fig D.5 versus Fig D.6).

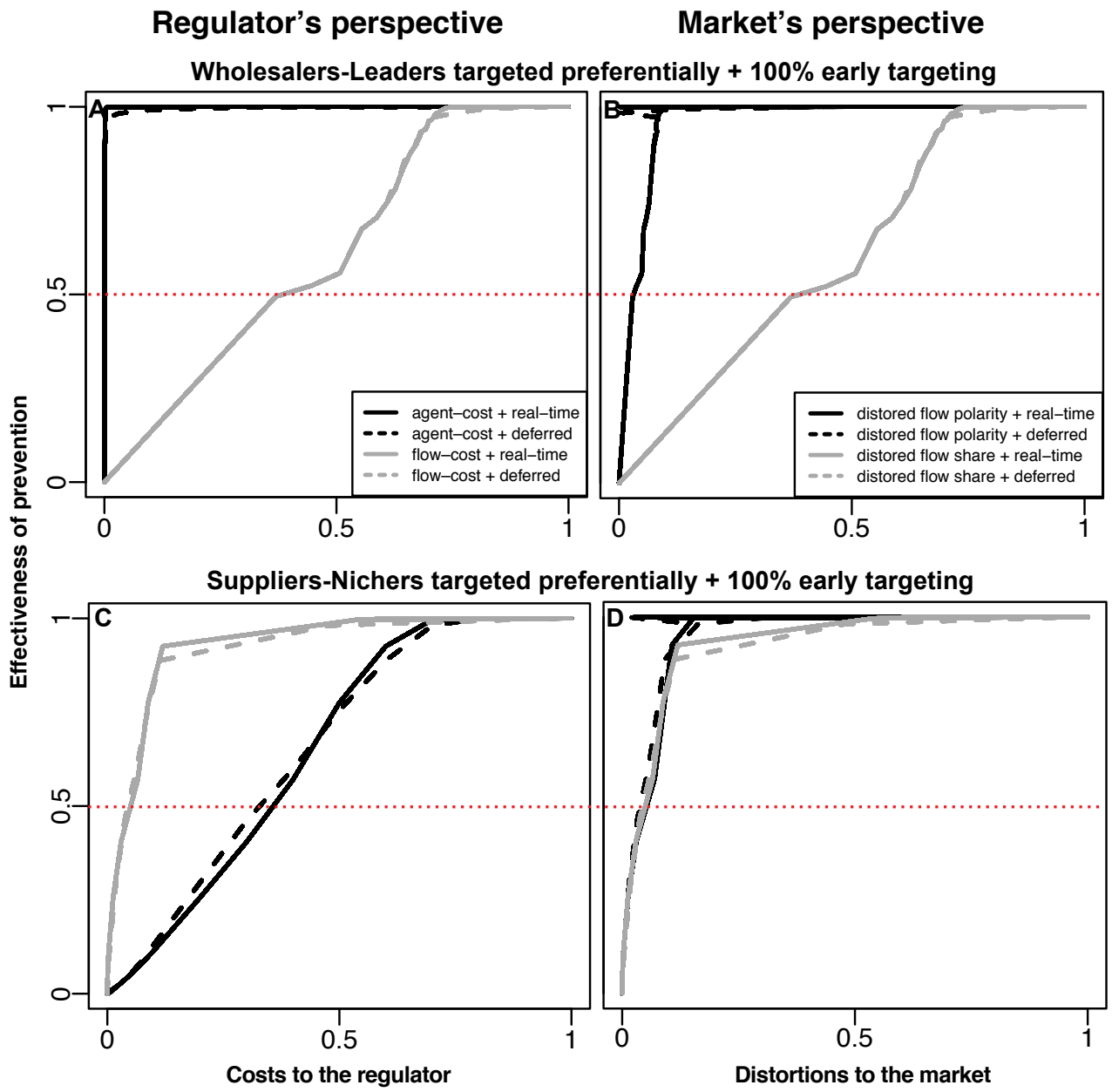


FIGURE D.3 – **Multiple-criteria decision analyses (MCDA) of contrasting targeted control strategies in Dpt 35 + ROW with early interventions.** MCDA of strategies targeting wholesalers-leaders first (**A**-**B**, with $z^{suppliers} = 1$; $z^{demanders} = 1$; $z^{nichers} = 0$; $z^{leaders} = 1$ in (3) of the main text) and suppliers-nichers first (**C**-**D**, with $z^{suppliers} = 1$; $z^{demanders} = 0$; $z^{nichers} = 1$; $z^{leaders} = 0$ in (3) of the main text). For each strategy, the x-axis quantifies prevention-efforts, i.e. the relative costs to the regulator (**A**-**C**) and relative distortions to the market (**B**-**D**) to reach a given prevention-effectiveness against epidemics as depicted on the y-axis (e.g. the red dotted lines to reach 50% of prevention-effectiveness). Prevention-effectiveness is measured by the relative decrease in the mean infection chain (Section D.1.2). We focus on Dpt 35 and associated exchanges with the rest of the world (ROW). Preventive strategies are implemented early, i.e at the beginning of year 2009, and evaluated over year 2009. Market categories are defined either over 2009 (real-time information available on agents, plain curves) or over 2008 (deferred information available on agents, dashed curves). Each case corresponds to 100 replicate simulations (notice the weak variability).

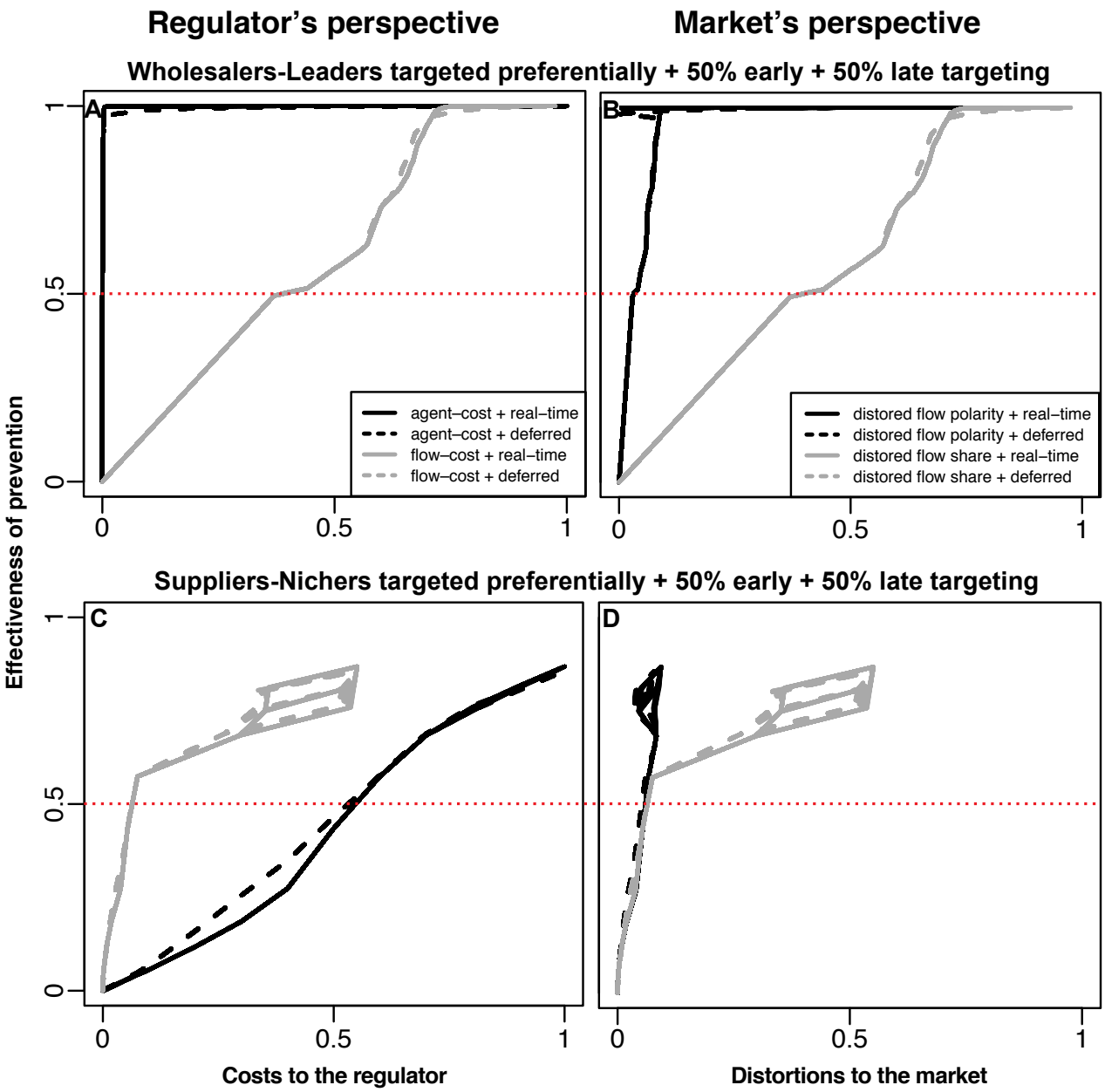


FIGURE D.4 – **MCDA of contrasting targeted control strategies in Dpt 35 + ROW with partly delayed interventions** We focus on Dpt 35 + ROW. Preventive strategies are implemented with a partial delay, i.e. the first half of agents to target are protected at the beginning of year 2009 and the second half on mid-2009, and evaluated over year 2009. The rest of the legend is identical to Fig. D.3.

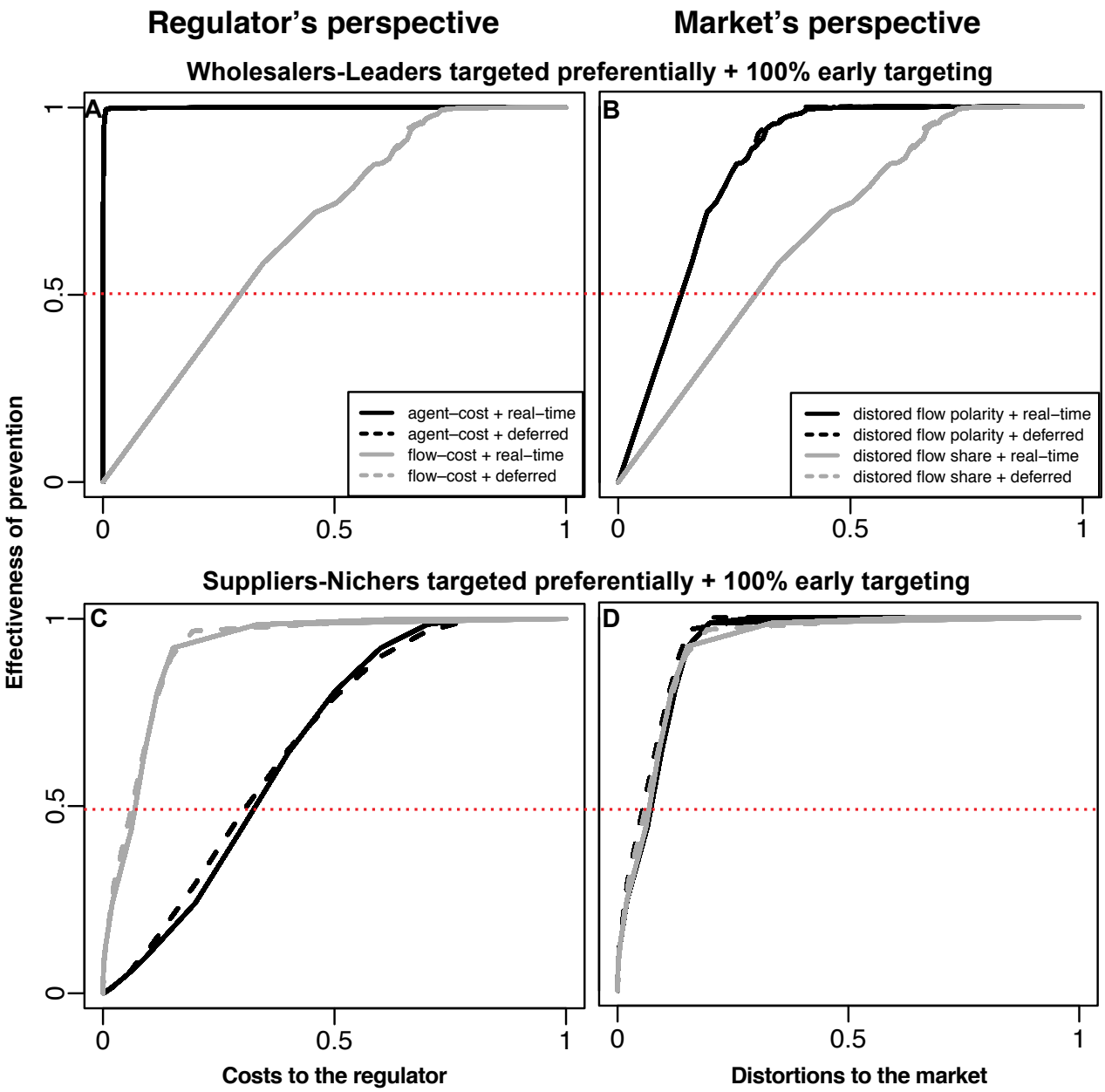


FIGURE D.5 – **MCDA of contrasting targeted control strategies in Dpt 71 + ROW with early interventions** We focus on Dpt 71 + ROW. Preventive strategies are implemented early, i.e. at the beginning of year 2009, and evaluated over year 2009. The rest of the legend is identical to Fig. D.3.

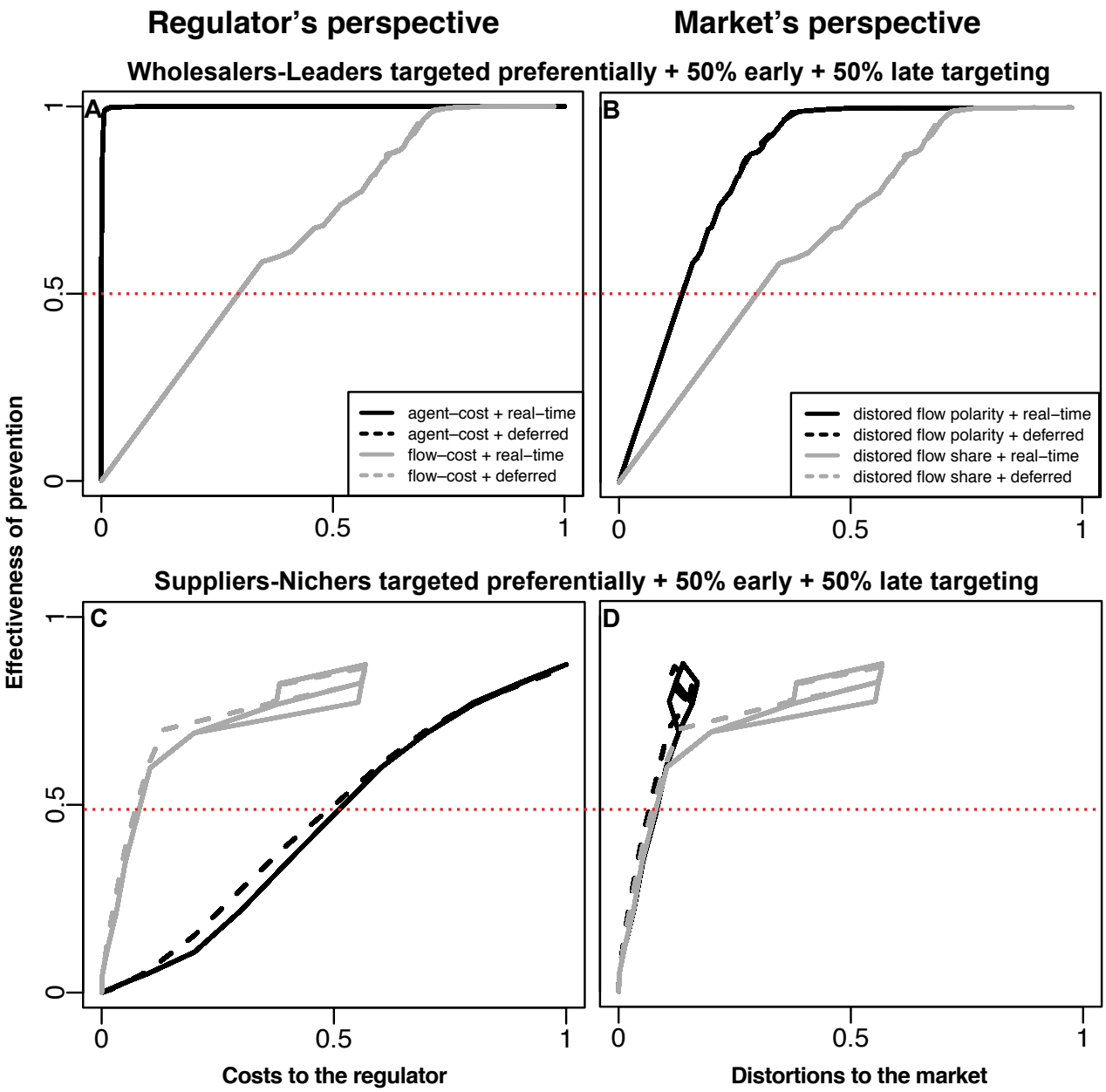


FIGURE D.6 – **MCDA of contrasting targeted control strategies in Dpt 71 + ROW with partly delayed interventions** We focus on Dpt 71 + ROW. Preventive strategies are implemented with a partial delay, i.e. the first half of agents to target are protected at the beginning of year 2009 and the second half on mid-2009, and evaluated over year 2009. The rest of the legend is identical to Fig. D.3.

Annexe E

Supporting information for : *Epidemics in markets with trade friction and imperfect transactions*

E.1 A comparative introduction to markets and market models

Understanding how markets emerge and operate remains a key open and controversial question in economics [e.g. Hahn, 1982; Katzner, 2010]. As trade-driven epidemics are impacted by market dynamics, we believe it is important to clarify the central concepts associated with markets. We also present market models widely used in the economic literature as a point of comparison to understand and justify the introduction of our frictional-trade market (FTM) model.

E.1.1 A tentative definition of markets

A market is an institution where voluntary exchanges of goods and services occur between economic agents. A market can hence be described as a network composed of agents in interaction [Rosenbaum, 2000; Goyal, 2009]. An agent is an entity which pursues its own interests through some kind of economic optimization. Agents have generally divergent interests resolved through exchanges and price definition [Guesnerie, 1996; Callon, 1998]. Examples of agents include individuals, businesses, countries or even sets of such entities [Goyal, 2009]. Note however that the concept of market is still debated among economists and remains largely ambiguous [Rosenbaum, 2000]. Besides the literature in economics, sociologists such as Callon [1998] argue that complex institutions such as markets cannot be reduced to networks. Operational fields such as marketing [Sissors, 1966] even define markets as peculiar group of people who do not necessarily form social networks (e.g. people who consume the same brand or people of similar age).

Two types of markets are distinguished in our study : *trading markets* where goods are exchanged

against money and *labour markets* where unemployed workers look for vacant jobs and companies seek to hire new workers.

E.1.2 A comparative review of existing market models

We review and compare existing market models that inspired our FTM model. As we focus on trade-driven epidemics, trading markets constitute the core of this review. We also briefly sketch search and matching labour models since we transpose their key concept of friction in our own model.

Trading market models

Mathematically, a trading market can be formalised explicitly at the agent-level through network models [Goyal, 2009; Atalay *et al.*, 2011] or implicitly through mass action compartmental models where agents are aggregated by categories [Mas-Colell *et al.*, 1995; Soliman *et al.*, 2010]. Here, markets are described as compartmental models corresponding to aggregated trade networks. We restrict ourselves to market models composed of two categories of agents with complementary interests : suppliers who wish to produce and sell goods if they receive money in compensation, and demanders who wish to buy and consume goods by providing money in exchange. We further assume that only two categories of units are exchanged against one another : goods against money. All goods are assumed to be identical. The price of one unit of money, known as the *numéraire* and denoted p_0, is assumed to remain constant and equal to 1. p_0 is the standard against which the relative value $p(t)/p_0$ of one unit of good is computed. A good hence represents here a typical product exchanged on the market against $p(t)/p_0$ units of money. We write throughout $p(t)$ for $p(t)/p_0$ that we simply refer to as the price. Apparently generic concepts such as supply and demand highly depend on the market model and are hence only defined on a case-by-case basis.

The partial equilibrium model

Economists essentially conceive market models working close to equilibrium rather than accounting for full transient dynamics and tipping points [May *et al.*, 2008; McCauley, 2009]. Quoting the reference textbook in microeconomics [Mas-Colell *et al.*, 1995] : *A characteristic feature that distinguishes economics from other scientific fields is that, for us, the equations of equilibrium constitute the center of our discipline. Other sciences, such as physics or even ecology, put comparatively more emphasis on the determination of dynamic laws of change. The reason, informally speaking, is that economists are good (or so we hope) at recognizing a state of equilibrium but are poor at predicting precisely how an economy in disequilibrium will evolve [...]. One of the difficulties in this area is the plethora of plausible disequilibrium models. Although there is a single way to be in equilibrium, there are many different ways to be in disequilibrium.*

Here, we introduce such an equilibrium-focused market model, called the partial equilibrium (PE) model, that still constitutes the basis of modern economic theory. In contrast with general equilibrium models, the term partial means that the PE model is restricted to a market where only one type of goods is produced. The PE model neglects population structure and transients to yield a simple relationship between the equilibrium values of supply, demand, and price [Mas-Colell *et al.*, 1995]. The market is said to be equilibrated at a price equalising supply and demand. To find such an equilibrium, the model specifies how supply and demand for a given good, denoted Q_S and Q_D respectively, change as a function of price p. $Q_S(p)$ means that suppliers will supply $Q_S(p)$ goods and demanders will demand $Q_D(p)$ goods at price p. Supply increases in price while demand decreases in price. The slopes associated with $Q_S(p)$ and $Q_D(p)$ are governed by quantities known as price elasticity of supply and demand and denoted respectively $\varepsilon_S(Q_S, p)$ and $\varepsilon_D(Q_D, p)$. Elasticities are defined as the relative change of supply or demand in response to the relative change in price :

$$\begin{aligned} \varepsilon_S(S,p) &\equiv \frac{dQ_S}{dp}\frac{p}{Q_S} \ , \\ \varepsilon_D(D,p) &\equiv -\frac{dQ_D}{dp}\frac{p}{Q_D} \ . \end{aligned} \tag{E.1}$$

Note the negative sign in the definition of ε_D so that both elasticities can only take positive values.

If we assume that elasticities are constants, it follows by integration that $Q_S(p)$ and $Q_D(p)$ follow simple power-laws with respect to price :

$$\begin{aligned} Q_S &= Q_{S_0}\left(\frac{p}{p_0}\right)^{\varepsilon_S} \ , \\ Q_D &= Q_{D_0}\left(\frac{p}{p_0}\right)^{-\varepsilon_D} \ , \end{aligned} \tag{E.2}$$

where Q_{S_0} and Q_{D_0} are the reference quantities supplied and demanded at the reference price $p = p_0 = 1$.

Supply and demand curves described by equations (E.2) correspond respectively to the sum of production functions when suppliers maximise their profits and the sum of consumption functions when demanders maximise their utilities [see Section E.3.1 and Chapter 10.C in Mas-Colell *et al.*, 1995].

The value of state variables at equilibrium, denoted by the sign eq in subscript, is found by solving the equation $Q_S(p) = Q_D(p)$, yielding :

$$\begin{aligned} Q_{S_{eq}} &= Q_{S_0} {p_{eq}}^{\varepsilon_S} \ , \\ Q_{D_{eq}} &= Q_{D_0} {p_{eq}}^{-\varepsilon_D} \ , \\ p_{eq} &= \left(\frac{Q_{D_0}}{Q_{S_0}}\right)^{\frac{1}{\varepsilon_S+\varepsilon_D}} \ . \end{aligned} \tag{E.3}$$

Equations (E.3) are in agreement with the so-called law of supply and demand (LSD). The latter

stipulates that the equilibrium price p_{eq} should increase if the reference demand Q_{D_0} is permanently increased (e.g. through a growing population). Conversely, p_{eq} should fall if the reference supply Q_{S_0} is increased (e.g. when new businesses enter the market). Notice however that relationships (E.3) tells us nothing about *how quickly* and *by which trading mechanisms* such an equilibrium emerge and shift when disrupted. While inapplicable to investigate the impact of market dynamics on epidemics, the PE model can still prove useful in practice. For instance, we can use a PE approach to quantify *a posteriori* the economic losses induced by epidemics. The PE method consists in comparing the initial and final state of the market with respect to a past outbreak [see e.g. Soliman *et al.*, 2010, for examples].

Tatônnement model of price dynamics

The partial equilibrium (PE) model neglects temporal dynamics. Building upon the PE model, Samuelson introduced an additional relationship between state variables, referred to in the literature as the Walras-Samuelson tatônnement (WST) model [see Chapter 17.H in Mas-Colell *et al.*, 1995]. The WST model postulates that price is updated based on the difference between demand and supply, known in the literature as the 'excess demand' and denoted $Q_E(t) \equiv Q_D(t) - Q_S(t)$. The full WST model is given by :

$$\begin{aligned} Q_S(t) &= Q_{S_0}\left(\frac{p(t)}{p_0}\right)^{\varepsilon_S} , \\ Q_D(t) &= Q_{D_0}\left(\frac{p(t)}{p_0}\right)^{-\varepsilon_D} , \\ Q_E(t) &= Q_D(t) - Q_S(t) , \\ \frac{dp}{dt} &= f(Q_E(t)) \quad \text{with : } \frac{\partial f}{\partial Q_E} \geq 0 . \end{aligned} \tag{E.4}$$

In practice, the relationship $\frac{dp}{dt} = f(Q_E(t))$ is usually assumed to scale either linearly or logarithmically with $Q_E(t)$ [Mas-Colell *et al.*, 1995; Anderson *et al.*, 2004]. Expressing price as function of supply and demand only, the price component of the WST model simplifies to :

$$\begin{aligned} &\textit{linear assumption :} \quad \frac{dp}{dt} = \lambda(Q_D - Q_S) , \\ &\textit{logarithmic assumption :} \quad \frac{dp}{dt} = \lambda\, p\, (Q_D - Q_S) , \end{aligned} \tag{E.5}$$

where $\lambda > 0$ is a rate parameter controlling the speed of adjustment. While the linear model is extremely simple, the logarithmic model has the advantage of always yielding positive prices when $p(t_0) > 0$.

The WST model (equations (E.4)-(E.5)) assumes that price will increase with respect to its current value when demand exceeds supply, and decrease when supply exceeds demands. This price adjustment process corresponds to the LSD. The LSD can hence refer both to price dynamics in a given market with a unique steady-state equilibrium (equations (E.4)-(E.5) of the WST model) or to the evolution to new equilibrium prices in response to external processes affecting the market (equations (E.3) of the

PE model).

While price is now a dynamical state variable, the total number of goods $Q(p)$ actually traded out of $Q_S(p)$ and $Q_D(p)$ is not specified by the WST model. Instead, the WST model assumes that no goods can actually be exchanged on the market prior equilibration, i.e. prior $Q_S = Q_D$. The weakness of this assumption is clearly stated and criticized in the economic literature and the conditions needed for the economic equilibrium $[Q_{S_{eq}}, Q_{D_{eq}}, p_{eq}]$ to be well-behaved seem very complicated [Hahn, 1982; Mas-Colell *et al.*, 1995; Anderson *et al.*, 2004; McCauley, 2009; Kitti, 2010]. We can conclude that even apparently simple concepts such as supply and demand have no clear real-world interpretation in simple trading market models such as the PE and the WST models.

Disequilibrium market models

In practice, trade also occurs when the market is not equilibrated [Hahn, 1982]. Although unsatisfactory from a conceptual perspective [Mas-Colell *et al.*, 1995; Katzner, 2010] and lacking empirical support [e.g. Hahn, 1982; Anderson *et al.*, 2004], disequilibrium (DE) market models were introduced to approximate the functioning of markets out of equilibrium [see the seminal contributions of Hahn & Negishi, 1962; Fair & Jaffee, 1972]. As for the WST model, a common assumption in DE models is that price dynamics are governed by the excess demand (see e.g. equations (E.5)). However, in contrast with the WST model, goods can be exchanged even when the market is not equilibrated.

Two types of contributions to DE dynamics should be distinguished due to their contrasted objectives and assumptions : theoretically-motivated and empirically-motivated studies. One the one hand, theoretical economists developed complex DE models to demonstrate the stability of general equilibrium (GE) market models. A GE model is a generalisation of the PE model to account for multiple markets in interaction. While insightful to investigate the convergence conditions in GE models where exchanges of goods are only based on initial endowments [Hahn & Negishi, 1962], theoretical approaches proved poorly applicable to GE models with production of new goods, the ones that matter for epidemics [see Hahn, 1982, for detailed explanations by a key author in the field]. One the other hand, applied economists developed simpler DE market models to estimate key economic parameters from empirical time series based on less restrictive assumptions. We hence only present in details the core model underlying empirically-motivated DE models : the DE model by Fair and Jaffee (FJ) [Fair & Jaffee, 1972; Quandt, 1988; Lee *et al.*, 2011]. The most parsimonious FJ model is given by [Chapter 2 of Quandt, 1988] :

$$\begin{aligned}
Q_S(\tau_k) &= w_{0,S}\ W_S(\tau_k) + w_{1,S}\ p(\tau_k) + \chi_S(\tau_k)\ , \\
Q_D(\tau_k) &= w_{0,D} W_D(\tau_k) - w_{1,D} p(\tau_k) + \chi_D(\tau_k)\ , \\
Q(\tau_k) &= \min\{Q_S(\tau_k), Q_D(\tau_k)\}\ , \\
p(\tau_k) &= p(\tau_{k-1}) + \lambda\,(Q_D(\tau_k) - Q_S(\tau_k)) + \chi_p(\tau_k)\ ,
\end{aligned} \tag{E.6}$$

where τ_k is the k^{th} period of time and $Q(\tau_k)$, $Q_S(\tau_k)$ and $Q_D(\tau_k)$ are the total quantities traded, supplied, demanded at price $p(\tau_k)$ during period τ_k respectively. $W_S(\tau_k)$ and $W_D(\tau_k)$ are vectors of variables other than price that influence $Q_S(\tau_k)$ and $Q_D(\tau_k)$ respectively. $w_{0,S}$, $w_{1,S}$, $w_{0,D}$, $w_{1,D}$, and λ are parameters to estimate and $\chi_S(\tau_k)$, $\chi_D(\tau_k)$ and $\chi_p(\tau_k)$ are error terms.

Interestingly, the FJ model emphasises that while supply and demand are non-observable, quantities traded and prices can be measured. Though some state variables are non-observable, parameter estimation is still possible because of a key dependency introduced between the state variables : at each time period, the total quantity traded (observed variable) is given by the *minimum* of supply and demand (non-observable variables). Thanks to the min function, supply and demand can now be interpreted as *willingnesses* to supply and demand goods. While far more realistic than the WST model to understand trade dynamics, the FJ model suffers from a major drawback : the residuals in willingnesses to trade, i.e. the leftovers implied by $Q_S \neq Q_D$, are not re-injected in the evolution of supply and demand. The functional shapes of supply and demand should have been modified to account for disequilibrium in trade. In addition, the contact process underlying disequilibrium trade is not specified, which makes the model inapplicable to epidemiological settings.

Labour market models : key insights from search and matching theory

The importance of friction : the labour market model of Diamond, Mortensen and Pissarides

While inapplicable as such to model disease dynamics on trading markets, labour market models provide inspiring concepts to grasp the contact process underlying market dynamics. Following Economic Sciences Prize Committee [2010], we focus on the reference labour market model of Diamond, Mortensen and Pissarides (DMP). The DMP model assumes a labour market operating on continuous time with a fixed number of labour force participants L. Let u denote the fraction of unemployed workers. It follows uL workers are unemployed and $(1-u)L$ workers are employed. Let v denote the fraction of L that corresponds to vacant positions, so that vL is the number of vacant positions while $(1-v)L$ is the number of non-vacant positions. In the general case, $u \neq v$. Here, we restrict our analysis of the DMP model to unemployment dynamics, i.e. we only specify a model for $\frac{du}{dt}(u,v)$. This restriction is sufficient to highlight the key contributions of the DMP model that we transpose in the FTM model.

In the DMP model, jobs are destroyed at per capita rate γ. Unemployed workers find a job at per capita rate ν. In contrast with γ that is assumed constant, ν depends on uL and vL. To fully specify ν, the DMP model introduces a matching function $\Psi = m(uL, vL, \kappa)$ where Ψ represents the rate of successful job matches at the population level that result from the joint-search efforts of the uL unemployed workers to find jobs and of companies to fill their vL vacancies. κ is a parameter controlling

the intensity of search and matching friction in the market. As expected intuitively, m is chosen so that Ψ increases with uL and vL and decreases with κ. By construction, ν is given by $\nu = \frac{\Psi}{uL}$. The dynamics of u are hence described by the differential equation :

$$\frac{du}{dt} = \frac{m(uL, vL, \kappa)}{uL} uL - \gamma(1-u)L \tag{E.7}$$

We now assume that unemployment reaches a steady-state. Setting $\frac{du}{dt} = 0$ yields a key relationship between u_{eq} and v_{eq}, known in the labour economics as the Beveridge curve :

$$u_{eq} = \frac{\gamma}{\gamma + \frac{m(u_{eq}L, v_{eq}L, \kappa)}{u_{eq}L}} \tag{E.8}$$

Based on the properties of the matching function m, the Beveridge curve implies that steady-state equilibrium unemployment u_{eq} and vacancies v_{eq} are negatively related. If we set the ratio v_{eq}/u_{eq} constant, increasing the level of friction κ will increase both u_{eq} and v_{eq}. In other words, by decreasing the rate of successful matches $\Psi = m(uL, vL, \kappa)$ between agents, friction makes the labour market worse.

Understanding the causes of friction : the mechanistic determinants of the matching function

While providing us with intuitions on the effects of imperfect matches on labour markets, the DMP model neither specifies the origins of κ nor the shape of the matching function m. m and κ are essentially black boxes [Pissarides, 2001]. Recent studies managed to uncover the mechanisms underlying m and κ [Lagos, 2000; Burdett *et al.*, 2001; Stevens, 2007], i.e. the local contact structure underlying friction in labour markets. Such studies assume time is discrete so that $M = m(uL, vL, \kappa)$ now corresponds to a number of successful matches and is not a rate any more. Here, we only present the taxicab model of Lagos [2000] as its matching function $M = m(uL, vL, \kappa)$ closely resembles our transaction rate Θ.

The taxicab model explicitly describes the labour market with an agent-based model based on taxicabs transporting passengers to their desired location on a spatial grid. In this model, uL represents the number of people who want to exit their current location to move to another location. vL represents the number of taxicabs. A passenger can only move on the grid if a free taxicab is present on its current location. Each taxicab can only transport one passenger at a time. When inside a taxicab, a passenger decides it destination. Free taxicabs decide where they want to pick up passengers. M denotes the total number of successful cab-passenger meetings occurring over the grid. Lagos [2000] shows that at the population-level, M can be simply expressed as :

$$M = \min\{uL, \frac{1}{\kappa} vL\} \leq \min\{uL, vL\}. \tag{E.9}$$

where $\kappa \geq 1$ increases with the heterogeneity in preferences expressed by taxicabs and passengers for specific spatial locations.

The case $\kappa = 1$ corresponds to the minimal amount of friction in the taxicab model where $M = \min\{uL, vL\}$. In contrast, the case $\kappa \to \infty$ corresponds to the maximal amount of friction where $M = 0$. The case where $\kappa \leq 1$ is not allowed in the taxicab model because taxicabs and people cannot be divided in smaller units. In contrast with the transaction rate Θ of the FTM model, notice that friction only affects vL and not uL. This stems from the asymmetric properties of taxicabs (that can move freely and decide where they want to pick new passengers) and passengers (who can decide where they want to go but cannot move by themselves). In the FTM model, both suppliers and demanders can move freely and search for each other, so we assume friction applies both to suppliers and demanders, in a symmetric way.

E.1.3 Correspondence between the FTM model and existing market models

While developing new concepts such as imperfect transactions with friction, the FTM model builds upon the four existing market models reviewed above (PE, WST, FJ and DMP). We highlight here the key differences and commonalities between the five models. In the following, supply stock, demanded stock and the total number of goods traded are denoted respectively T, S and D when they are generated based on continuous processes, while they are denoted Q, Q_S and Q_D in the discrete case.

Correspondence with the partial equilibrium (PE) model

The FTM model assumes that supply and demanded stocks are respectively generated at rates $\Sigma_\oplus(p, N_S)$ and $\Delta_\oplus(p, N_S)$ given by :

$$\begin{aligned} \Sigma_\oplus(p, N_S) &= N_S \ \sigma_0 \ p^{\varepsilon_S} \ , \\ \Delta_\oplus(p, N_D) &= N_D \ \delta_0 \ p^{-\varepsilon_D} \ . \end{aligned} \tag{E.10}$$

where elasticities are defined the same way as the PE model by replacing stocks with rates (see equations (E.1) - (E.3)). As it defines generation rates rather than supply and demanded stocks, the FTM model is a generalisation of the PE model to dynamical settings. Further details on the microeconomic foundations underlying the generation rates of supply and demanded stocks (E.10) are available in Section E.3.1.

Correspondence with the Walras-Samuelson tatônnement model (WST) of price dynamics

The FTM model assumes that variations in price are directly related, via a dimensionless coefficient μ, to changes in net willingness to transact :

$$\frac{dp}{dt} = \mu \frac{d(D - S)}{dt} p = \mu(\Delta - \Sigma)p \ . \tag{E.11}$$

As the WST model requires that trade only occurs at equilibrium, assuming that price updating is based on the excess demand stock ($Q_D - Q_S$; see equations (E.5)) is identical to assuming that price updating is based on the excess demand creation rate ($\Delta_\oplus - \Sigma_\oplus$). In other words, the WST model implicitly assumes that stocks of supply (Q_S) and demand (Q_D) are identical to our creation rates of supply ($\Sigma_\oplus$) and demand ($\Delta_\oplus$) as far as price dynamics are concerned. The two models are hence fully equivalent when losses and external flows are neglected in the FTM model, i.e. when $\Sigma = \Sigma_\oplus$ and $\Delta = \Delta_\oplus$.

Correspondence with the disequilibrium market model by Fair and Jaffee (FJ)

Our definition of the average transaction stock exchanged from a supplier to a demander (q in the main text) is directly inspired by the FJ model. Here we show that the two market approaches are not equivalent : the FTM model involves less restrictive assumptions on how trade is carried out compared with the FJ model. Our market model is more general than the FJ model as it accounts for various levels of friction.

Implications of the hypotheses of Fair and Jaffee

The key contribution of the discrete-time FJ model (see equations (E.6)) essentially boils down to :

$$Q(\tau_k) = \min\{Q_S(\tau_k), Q_D(\tau_k)\} \ . \tag{E.12}$$

We extend the FJ model to continuous time. To give a fair comparison of the models, we assume that supply and demanded stocks are generated at net rates $\Sigma(t)$ and $\Delta(t)$, the same as in the FTM model. We denote by $\Phi_{\mathrm{FJ}}(t)$ the trade flow generated by the FJ model. Our objective is to characterise the value of $\Phi_{\mathrm{FJ}}(t)$ or its integral $T_{\mathrm{FJ}}(t)$ as function of parameters and functions common to both models.

Generalized to continuous time, equation (E.12) implies :

$$\underbrace{\int_{t-dt}^{t} \Phi_{\mathrm{FJ}}(u)du}_{\text{Total number of goods traded from } t-dt \text{ to } t} = \min\Bigg[\underbrace{S(t_0) + \int_{t_0}^{t-dt} \Sigma(u)du - \int_{t_0}^{t-dt} \Phi_{FJ}(u)du}_{\text{Residual supply stock at } t-dt} + \underbrace{\int_{t-dt}^{t} \Sigma(u)du}_{\text{Creation of supply stock from } t-dt \text{ to } t} ;$$
$$\underbrace{D(t_0) + \int_{t_0}^{t-dt} \Delta(u)du - \int_{t_0}^{t-dt} \Phi_{FJ}(u)du}_{\text{Residual demanded stock at } t-dt} + \underbrace{\int_{t-dt}^{t} \Delta(u)du}_{\text{Creation of demanded stock from } t-dt \text{ to } t}\Bigg] . \tag{E.13}$$

After simplifications, equation (E.13) leads to :

$$T_{\mathrm{FJ}}(t) = T(t_0) + \min\left[S(t_0) + \int_{t_0}^{t} \Sigma(u)du \ ; \ D(t_0) + \int_{t_0}^{t} \Delta(u)du\right] , \tag{E.14}$$

where $T_{\mathrm{FJ}}(t)$ is the total number of goods traded from time t_0 to time t under the FJ model.

Implications of the hypotheses of the FTM model

The generic version of the FTM model reads :

$$\begin{aligned} \frac{dS}{dt} &= \Sigma(t) - \Phi(t), \\ \frac{dD}{dt} &= \Delta(t) - \Phi(t) \ , \\ \frac{dT}{dt} &= \Phi(t) \ . \end{aligned} \tag{E.15}$$

This directly implies :

$$\begin{aligned} \int_{t_0}^{t} \Phi(u)du &= S(t_0) + \int_{t_0}^{t} \Sigma(u)du - S(t) \ , \\ \int_{t_0}^{t} \Phi(u)du &= D(t_0) + \int_{t_0}^{t} \Delta(u)du - D(t) \ . \end{aligned} \tag{E.16}$$

In particular, equalities (E.16) lead to :

$$T(t) = T(t_0) + \min\left\{S(t_0) + \int_{t_0}^{t} \Sigma(u)du - S(t) \ ; \ D(t_0) + \int_{t_0}^{t} \Delta(u)du - D(t)\right\} , \tag{E.17}$$

where $T(t)$ is the total number of goods traded from time t_0 to time t under the FTM model.

Comparison of the two models

We immediately notice :

$$T(t) \leq T_{\mathrm{FJ}}(t) \quad \forall t \ . \tag{E.18}$$

The two quantities are equal at market steady-state equilibrium when the residual supply and demand stocks are equal and vanish. The FJ model is hence a special case of our reference market model where : 1) we assume that the LSD is respected (so that $S^* = D^* = \max\{\kappa N_S, \kappa N_D\}$; see the analysis of the integral price model with $S_0 = D_0$ in section E.3.3) ; 2) we neglect friction (so that $\kappa \to 0$). Taken together, special cases 1) and 2) imply $S^* = D^* = 0$, which leads to $T \to T_{\mathrm{FJ}}$.

Intuitively, the two approaches are equivalent only when the market fully clears, i.e. when each good

produced is instantly shipped and consumed. Our approach allows for the more general case where the LSD is not necessarily respected and where the time needed for a transaction is not negligible any more. In other words, accounting for imperfect transactions with friction implies that supply and demand stocks can never vanish, even at steady-state market equilibrium.

Correspondence with the labour market model of Diamond, Mortensen and Pissarides (DMP)

Both the DMP and the FTM model postulate the existence of friction. Our motivation for considering trade friction in the FTM model is the following : the exchange of goods is conditional on goods having been produced. For livestock, production takes substantial time (usually months for swine and years for cattle). Farmers also need to search for business partners and negotiate deals, which can generate further delay in trade. Finally, transportation of animals is costly, so farmers tend to exchange batches of animals. Such constraints on trade are sources of friction that, as far as we know, have not been examined with models of trade-driven epidemics, particularly for livestock markets. We show in the main text that they can have important consequences that should be considered in the regulation and management of disease outbreaks in livestock markets. Our concept of friction aggregates the rigidities affecting trade and disease outbreak in livestock markets. In practice, there is biological delay of months (for swine) or years (for cattle) until farm animals can be profitably traded. This delay is accounted for through our rate of production of supply stock and can lead to accumulation of stock if friction is not negligible. In addition, the average batch size q (number of animals exchanged per transaction) relates to our friction parameter κ (in steady state). We show that the constraint $\kappa \geq 1$, reflecting that the minimum unit of trade is one living animal, has key economic and epidemiological implications. Moreover, it is in fundamental contrast with models that assume friction is negligible ($\kappa \approx 0$). Specifically, the interplay between slow generation of livestock (months, years), and the potentially much faster pace (days, weeks) of trade transactions and associated spread of infection (such as foot-and-mouth disease), can have important consequences for disease control as well as for the market in the long term.

The DMP model was designed to approximate labour markets while the FTM model was designed to approximate trading markets. It follows the two models differ in nature, notably as regard to the type of exchanges and friction involved. We also show that their conclusions are different in scope and highlight what we believe is our key economic contribution in contrast with existing market models : the notion of imperfect trade transaction with friction.

Both models involve two market sides and interaction events between them that depend on friction. The DMP model introduces a matching function $\Psi(uL, vL, \kappa_{\mathrm{DMP}})$ that closely resembles our transaction rate $\Theta(N_S, N_D, \sigma_0, \delta_0, p, \kappa)$, where κ_{DMP} and κ are the coefficients of friction associated with the DMP model and the FTM model respectively. We recall that Ψ represents the rate of successful job

matches at the population level that result from the joint-search efforts of the uL unemployed workers to find jobs and of companies to fill their vL vacancies. Θ represents the transaction rate between the N_S suppliers and the N_D demanders.

The types of interaction between pairs of agents involved in labour and trading markets are not equivalent. Labour markets involve on-off matches between workers and employers : either an encounter is successful and a vacancy is filled by a worker or the pairing is rejected. Trading markets involve recurrent transactions between suppliers and demanders, and each transaction is associated with a variable number of goods exchanged q. More precisely :

- A *job match* is stable for a certain period of time (a new employee is likely to keep her/his job for a few days at the very least), is usually unrepeated for a given pair of agents (a person who gets fired by a company is unlikely to get hired again in the same company) and corresponds to a binary process (a person is either employed or unemployed).
- In contrast *a trade transaction* is a transient process (once a trade is realised, the physical interaction stops), is generally repeated in time for a given pair of agents (a supplier and demander who get on well are likely to interact multiple times per time period) and corresponds to a continuous process (a variable number of goods is exchanged from a given supplier to a given demander).

The types of friction hence spanned by κ_{DMP} and κ differ : while κ_{DMP} aggregates job search and matching friction, κ represents trade friction (= production friction + partner search friction + stock delivery friction). The ranges encompassed by κ_{DMP} and κ are also different : though κ_{DMP} must satisfy $\kappa_{\text{DMP}} \geq 1$ because you cannot exchange less than one 'unit' of labour, κ satisfies $\kappa \geq 0$ because you can have more transaction events per time unit than production/consumption events per time unit. Note that the average transaction stock $q \equiv \min\{\frac{S}{N_S}, \frac{D}{N_D}\}$ of the FTM model is associated with another type of dynamical "friction" : misbalance in stock-matching, since the per capita supply does not necessarily match the per capita demand. This type of friction is not accounted for by the DMP model.

The models finally differ in their conclusions. The DMP model shows that friction in job matches is essential to explain the existence of persistent unemployment in labour markets. Our model not only shows that friction in transactions explains the persistence of residual supply and demanded stocks in trading markets, but friction can also modify the joint-dynamics of markets and epidemics (see results in the main text). This point stands whether epidemics are included or not, and suggests that the FTM model can also prove useful from a purely economic point of view. Generally speaking, we believe the notion of dynamical transactions with friction is an improvement over existing trading market models.

E.2 Model parametrisation and estimation of friction from French cattle and swine livestock exchange data

E.2.1 Presentation of the datasets

To parameterise the FTM model and estimate the range of friction encompassing real markets, we analyse trading datasets of two livestock markets : cattle (BDNI dataset) and swine (BDPorc dataset). The BDNI and BDPorc datasets are respectively managed by the French ministry in charge of Agriculture (`http://agriculture.gouv.fr/identification-et-tracabilite`) and the French professional union BDPorc (`http://www.bdporc.com/`). Each dataset details movements of living animals occurring in France among all economic agents involved in the supply chain, from strictly breeding farms to slaughterhouses with various categories of wholesalers in between (e.g. breeding-fattening farms, strictly fattening farms, markets, dealers). Imports and exports are also included. Traceability is imposed by the regulator at the scale of individual animals for cattle and batches of animals for swine. Declaration of cattle and swine movements is compulsory in France since January 1999 and July 2009 respectively.

E.2.2 Extraction of individual transactions from the datasets

To estimate the parameters of the FTM model, we need to extract for each dataset a table of individual transactions detailing : the pair of supplier-demander involved during each transaction, the date and the associated volume of goods exchanged. We only extract such detailed transactions for a subset of the datasets. We focus on cattle data for civil year 2009 and pig data for civil year 2010. Following the literature on networks of livestock exchanges in France [Rautureau *et al.*, 2011, 2012], we neglect imports, exports and movements of animals to slaughterhouses and rendering plants. Imports and exports are already accounted for in the FTM model as free parameters (E_S for supply and E_D for demand, with $E > 0$ for imports and $E < 0$ for exports). While concentrating a large part of the total number of goods exchanged (about 50% for cattle [Rautureau *et al.*, 2011] and 80 % for pigs [Rautureau *et al.*, 2012]), slaughterhouses and rendering plants are positioned at the very bottom of the supply chain. They are hence probably associated with a negligible risk of infection [Rautureau *et al.*, 2012]. Note that transactions involving topological dead-ends such as slaughterhouses could be easily included in the FTM model as we distinguish wholesalers from strict demanders. However, as the FTM model describes mass-action interactions among identical agents within their respective hierarchical categories (strict suppliers, wholesalers and strict demanders), we would largely overestimate the infectious risk associated with trade by over-estimating the fractions of transactions and average transaction stocks that contribute to infection. In other words, even if such dead-end movements are important for the market, we neglect them in favor of more epidemiological realism.

Extraction of transactions from the BDNI dataset (cattle)

Since the scale of traceability is known at the animal level for cattle, transactions can be directly extracted from the BDNI dataset.

Extraction of transactions from the BDPorc dataset (swine)

Swine traceability is only known at the scale of batches of pigs transported by the same truck during a round of transportation. A round is a movement of swine involving the same truck that starts a journey empty, loads pigs from one or more holdings and finishes the journey empty after successful delivery to one or more holdings. Loading and delivery events may occur several times and in any order during a given round. Inferring detailed transactions from the BDPorc dataset hence requires some reconstruction hypotheses that are detailed here.

The inference problem is the following : we know how much each agent is contributing, with whom and when it is interacting, but we do not know which pigs precisely reach a given destination. In other words, we need to infer the quantities of pigs exchanged $q^r_{i,j}$ from agent i^r to agent j^r during round r. The problem is identical to the reconstruction of sex acts on weighted sexual contact networks, where we know the partnerships, how much sex acts each individual has, but we do not know how many sex acts are associated with each existing partnership. We hence use the same methodology of reconstruction as Moslonka-Lefebvre *et al.* [2012a]. In the following, the letter i denotes individual suppliers and the letter j individual demanders. $\widehat{x}$ means that quantity x is a raw data that will be subsequently modified by a reconstruction treatment.

Data reconstruction is applied to each round r associated with *at least* n^r suppliers (since we remove some suppliers such as importers) and *at least* m^r demanders (since we remove some demanders such as slaughterhouses). For each round r, we know that : i^r is a supplier supplying $\widehat{s^r_i}$ pigs to at least m^r demanders, j is a demander receiving $\widehat{d^r_j}$ pigs from at least n^r sellers and that i^r may be connected to j^r because the same truck visited these two agents during the same round. However, we do not know for sure if i^r actually shipped goods to j^r as more than two agents can be involved per round, i.e. we do not know the value of $q^r_{i,j}$. The reconstruction of $q^r_{i,j}$ operates in four steps :

- **Step 1 : Calculation of the maximal number of goods exchanged.** To account for cut-offs e.g. induced by neglected slaughterhouses and imports, we calculate the total number of goods q^r that is exchanged during round r among the A^r *non-neglected* agents. A^r is made of n^r suppliers and m^r demanders. The n^r suppliers supply $\widehat{s^r} = \sum_{i=1}^{n^r} \widehat{s^r_i}$ goods. The m^r demanders receive $\widehat{d^r} = \sum_{j=1}^{m^r} \widehat{d^r_j}$ goods. We assume $q^r = \min\{\widehat{s^r}; \widehat{d^r}\}$.
- **Step 2 : Normalisation of the number of goods exchanged at the round level.** The quantities $\widehat{s^r_i}$ and $\widehat{d^r_j}$ are normalised so that the total quantity of pigs supplied matches the total

quantity of pigs demanded within a round :

$$s_i^r \leftarrow \widehat{s_i^r} \frac{q^r}{\widehat{s^r}} \ ,$$
$$d_j^r \leftarrow \widehat{d_j^r} \frac{q^r}{\widehat{d^r}} \ ,$$

where the notation $x \leftarrow y$ means that we replace the value x by the value y. After normalisation, we have by construction : $\sum_i s_i^r = \sum_j d_j^r = q^r$.

– **Step 3 : Partitioning of potential goods exchanged.** Since traceability is only known at the round-level, we do not know the precise destination of each pig. We hence need to specify a preference function describing for each agent how much goods she/he wishes to trade with each of its potential partners. Let the asymmetric preferences functions $[i^r \to j^r]^{\text{sell}}$ and $[j^r \to i^r]^{\text{buy}}$ represent the total amount of goods that i^r wishes to sell to j^r and j^r buy from i^r during round r respectively. Here we assume that agents favour preferentially their biggest potential partners with a proportional model :

$$[i^r \to j^r]^{\text{sell}} = s_i^r \frac{d_j^r}{\sum_j d_j^r} \ ,$$
$$[j^r \to i^r]^{\text{buy}} = d_j^r \frac{s_i^r}{\sum_i s_i^r} \ .$$

– **Step 4 : Reconstruction of actual goods exchanged.** For each pair of trading agents (i^r,j^r), the actual number of goods exchanged $q_{i,j}^r$ is assumed to be the minimum of their respective preferences :

$$q_{i,j}^r = \min\{[i^r \to j^r]^{\text{sell}}; [j^r \to i^r]^{\text{buy}}\} = \frac{s_i^r d_j^r}{q^r} \ .$$

All $q_{i,j}^r$ are reconstructed by iterating Steps 1-4 for all pairs of agents involved in round r. Straightforward calculations show that $\sum_{i,j} q_{i,j}^r = q^r$.

E.2.3 Estimation of model parameters

The assumptions on the FTM model

To enable quick estimations, we assume that the two livestock markets analysed are at steady-state equilibrium and are not submitted to any kind of major shock. We also neglect losses ($L_S = L_D = 0$) and external flows ($E_S = E_D = 0$). We hence recover the reference market model (see main text). We finally assume that the reference price is equal $p^* = 1$. We deduce the following relationships between model parameters :

$$\Phi^* = \Theta^* \kappa \ ,$$
$$\Sigma^* = \Sigma_\oplus^* = N_S \sigma_0 = \Phi^* \ ,$$
$$\Delta^* = \Delta_\oplus^* = N_D \delta_0 = \Phi^* \ ,$$

with $N_S = N_{S\setminus D} + N_{S\cap D}$ and $N_D = N_{D\setminus S} + N_{S\cap D}$.

We only need to estimate Φ^*, Θ^*, $N_{S\setminus D}$, $N_{S\cap D}$ and $N_{D\setminus S}$ to deduce all the other parameters.

Determination of key parameters from data

Let t_1 and t_2 be the extreme dates of the data we wish to consider for parameter estimation. The corresponding interval of time is given by : $\Delta t = t_2 - t_1$. For each market, agents are numbered from 1 to N. Let $a \in [|1; N|]$ denote the index of an agent. When referring to a pair of agents, we reserve the index $i \in [|1; N|]$ for sellers and the index $j \in [|1; N|]$ for demanders. For cattle data, a round r corresponds to a unique transaction involving a supplier and a demander. For swine data, a round r is a set of transactions involving potentially multiple suppliers and demanders. By definition, $q^r_{i,j} = 0$ means that i and j are not connected during round r, while $q^r_{i,j} > 0$ means that supplier i and demander j are connected by a unique transaction.

Let $s_a([t_1; t_2])$ and $d_a([t_1; t_2])$ denote the total number of goods supplied and received by a over $[t_1; t_2]$ respectively. By definition we have :

$$s_a([t_1; t_2]) = \sum_{r\in[t_1;t_2]} \sum_j q^r_{a,j}$$
$$d_a([t_1; t_2]) = \sum_{r\in[t_1;t_2]} \sum_i q^r_{i,a}$$

where $r \in [t_1; t_2]$ means that round r occurs within period $[t_1; t_2]$.

It follows $N_{S\setminus D}$, $N_{S\cap D}$ and $N_{D\setminus S}$ are given by :

$$N_{S\setminus D}([t_1; t_2]) = \sum_a \mathbb{1}_{\{s_a([t_1;t_2])\neq 0\}\&\{d_a([t_1;t_2])=0\}}$$
$$N_{S\cap D}([t_1; t_2]) = \sum_a \mathbb{1}_{\{s_a([t_1;t_2])\neq 0\}\&\{d_a([t_1;t_2])\neq 0\}}$$
$$N_{D\setminus S}([t_1; t_2]) = \sum_a \mathbb{1}_{\{s_a([t_1;t_2])=0\}\&\{d_a([t_1;t_2])\neq 0\}}$$

where $\mathbb{1}_O$ is the indicator function of event O, i.e. $\mathbb{1}_O = 1$ when O is true and $\mathbb{1}_O = 0$ when O is false.

Reference trade flow Φ^* and transaction rate Θ^* are estimated as :

$$\Phi^*([t_1; t_2]) = \frac{\sum_{r\in[t_1;t_2]} \sum_{i,j} q^r_{i,j}}{\Delta t}$$
$$\Theta^*([t_1; t_2]) = \frac{\sum_{r\in[t_1;t_2]} \sum_{i,j} \mathbb{1}_{q^r_{i,j}\neq 0}}{\Delta t}$$

Average values of model parameters

Apart from the coefficient of friction (see below), we are solely interested by the average value of model parameters over one year of transactions. Table 1 in the main text contains the parameter values derived from the French cattle and swine markets with $\Delta t = 1$ year.

Dynamics of the coefficient of friction

For each market, we calculate the monthly value of the coefficient of friction κ and set $\Delta t = 1/12$ year. We show in Fig. E.1 that κ is constant for each market.

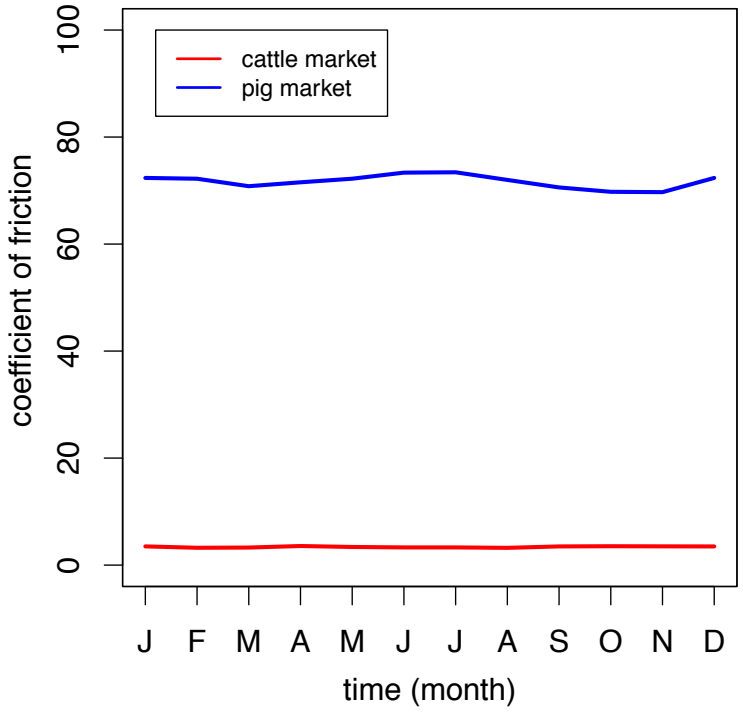


FIGURE E.1 – **Dynamics of the coefficients of friction for cattle and swine markets.** For each market, the coefficient of friction is calculated every month.

E.3 Supplementary results on the frictional-trade market (FTM) model

In addition to detailed proof and additional results, this section presents microeconomic insights on the FTM model.

E.3.1 Microeconomic foundations

We highlight the microeconomic mechanisms underlying the FTM model and show that there are in agreement with common sense and mainstream economic theories [Mas-Colell *et al.*, 1995; Acemoglu, 2008].

At the population level, the FTM model assumes that supply and demand rates are generated at global rates $\Sigma_{\oplus}(p, N_S)$ and $\Delta_{\oplus}(p, N_S)$ respectively (E.10). The FTM model is an aggregated economic model, which implies that each supplier and demander is identical to an average representative supplier and demander respectively. Let $\sigma(p)$ and $\delta(p)$ represent the *per-capita* rates of generation of supply

and demanded stocks. By construction we have :

$$\begin{aligned} \sigma(p) &= \frac{\Sigma_{\oplus}(p, N_S)}{N_S} = \sigma_0 \left(\frac{p}{p_0}\right)^{\varepsilon_S} , \\ \delta(p) &= \frac{\Delta_{\oplus}(p, N_D)}{N_D} = \delta_0 \left(\frac{p}{p_0}\right)^{-\varepsilon_D} , \end{aligned} \tag{E.19}$$

where p_0 is the numéraire, i.e. the price of one unit of money (in practice, we assume $p_0 = 1$ to simplify calculations).

Below, we derive the elementary profit and utility functions that underly (E.19) and hence (E.10) after summation. We also show that the transaction rate Θ of the FTM model (equation (6) in the main text) stems from the aggregation of elementary transactions rates for each supplier-demander pair in the market. The average transaction stock q (equation (5) in the main text) is already defined at the microscopic level. Evolution of per-capita supply and demanded stocks can be straightforwardly deduced from the previous considerations. We finally discuss some limitations of the FTM model and potential extensions.

Profit maximisation underlying the generation rate in supply stock

Let $\pi(p, \sigma_x)$ represent the expected profit per time unit of the representative supplier when producing σ_x sellable goods per time unit. In the general case, there is no reason that equality $\sigma_x = \sigma(p)$ is satisfied (E.19). By definition, expected profit is the difference between expected benefits per time unit, i.e. $p\ \sigma_x$, and production costs per time unit, denoted $c(\sigma_x)$:

$$\begin{aligned} [\text{expected profit}] &= [\text{expected benefits}] - [\text{production costs}] , \\ \pi &= p\ \sigma_x - c(\sigma_x) . \end{aligned} \tag{E.20}$$

Property 1 : For a given price p, the representative supplier maximises his/her expected profit π by producing $\sigma_x = \sigma(p)$ sellable goods per time unit. The underlying optimisation program of the supplier is given by :

$$\begin{aligned} &\text{The supplier sets } \sigma_x \text{ to maximise } \pi(p, \sigma_x), \\ &\text{with corresponding production costs : } c(\sigma_x) = c_{\text{sunk}} + \frac{p_0\ \sigma_0}{1 + 1/\varepsilon_S} \left(\frac{\sigma_x}{\sigma_0}\right)^{1+1/\varepsilon_S} , \end{aligned} \tag{E.21}$$

where c_{sunk} is a positive constant representing sunk costs per time unit (implicitly including costs of search and delivery).

Note that since $\dfrac{d^2 c(\sigma_x)}{d{\sigma_x}^2} = \dfrac{p_0}{\sigma_0\ \varepsilon_S} \left(\dfrac{\sigma_x}{\sigma_0}\right)^{1/\varepsilon_S - 1} \geq 0\ \forall \varepsilon_S \geq 0$, production costs correspond to a strictly convex technology, i.e. marginal costs of production $\dfrac{dc(\sigma_x)}{d\sigma_x}$ increase with production σ_x. $\sigma(p)$ hence

exhibits decreasing returns to scale in agreement with the law of diminishing returns [Chapter 5 in Mas-Colell *et al.*, 1995],

Proof of Property 1 is as follows :

For a given price p, the producer maximises his/her profit π for σ_x satisfying [Chapter 5 in Mas-Colell *et al.*, 1995] :

$$\frac{\partial \pi}{\partial \sigma_x} = 0 \ . \tag{E.22}$$

Condition (E.22) implies that σ_x is set by the supplier so that the marginal cost of production equals the price :

$$\frac{\partial c(\sigma_x)}{\partial \sigma_x} = \frac{dc(\sigma_x)}{d\sigma_x} = p \ . \tag{E.23}$$

The per-capita generation rate of supply stock (E.19) can be rewritten :

$$p_0 \left(\frac{\sigma(p)}{\sigma_0} \right)^{1/\varepsilon_S} = p \ . \tag{E.24}$$

By identification, the left-hand sides of (E.23) and (E.24) are necessarily equal. The marginal costs of production are hence given by :

$$\frac{dc(\sigma_x = \sigma(p))}{d\sigma_x} = p_0 \left(\frac{\sigma_x}{\sigma_0} \right)^{1/\varepsilon_S} \ , \tag{E.25}$$

which directly yields $(E.21)$ after integration.

Utility maximisation underlying the generation rate in demand stock

Let $u(\delta_x, ...)$ denote the expected utility per time unit of the representative demander when he/she expects to buy δ_x focal goods per time unit. $u(\delta_x, ...)$ quantifies the expected satisfaction per time unit of the representative demander when he/she expects to consume δ_x. In the general case, $\delta_x \neq \delta(p)$ (E.19). Let $\widehat{\delta_y}$ represent the expected consumption rate for goods on other markets than the focal market. We hence have : $u(\delta_x, ...) = u(\delta_x, \widehat{\delta_y})$, because the utility of the representative demander depends both on the consumption of δ_x and $\widehat{\delta_y}$.

Property 2 : For a given price p, the representative demander maximises his/her expected utility u by consuming $\delta_x = \delta(p)$ goods per time unit. The underlying optimisation program of the demander

is given by :

$$\begin{aligned}
&\text{The demander chooses } (\delta_x, \widehat{\delta_y}) \text{ to maximise } u(\delta_x, \widehat{\delta_y}\ ) \ , \\
&\text{subject to the budget constraint : } p\delta_x + p_0\widehat{\delta_y} = b \ , \\
&\text{with corresponding utility : } u(\delta_x, \widehat{\delta_y}\ ) = \underbrace{\frac{\delta_0}{1-1/\varepsilon_D}\left(\frac{\delta_x}{\delta_0}\right)^{1-1/\varepsilon_D}}_{u_1(\delta_x)} + \underbrace{\widehat{\delta_y}}_{u_2(\widehat{\delta_y}\ )} \ ,
\end{aligned} \tag{E.26}$$

where b is the budget generation rate of the demander (implicitly including costs of search and delivery) and u_1 and u_2 represent the utilities per time unit associated with the expected consumptions δ_x and $\widehat{\delta_y}$ respectively. Note that the optimal level of $\widehat{\delta_y}$ is given by solving the budget constraint for $\delta_x = \delta(p)$.

From the expression of u_1 in (E.26), we notice, $\forall \varepsilon_D \geq 0$, $\frac{du_1(\delta_x)}{d\delta_x} = \left(\frac{\delta_x}{\delta_0}\right)^{-1/\varepsilon_D} \geq 0$ and $\frac{d^2u_1(\delta_x)}{d{\delta_x}^2} = -\frac{1}{\delta_0\ \varepsilon_D}\left(\frac{\delta_x}{\delta_0}\right)^{-1/\varepsilon_D-1} \leq 0$. It follows u_1 is increasing and concave, in agreement with the economic literature on utility functions [Chapter 3 in Mas-Colell *et al.*, 1995]. Though u_1 increases with δ_x, concavity implies diminishing marginal utility, i.e. the representative demander tends to get satiated for large δ_x. Interestingly, u_1 can also be interpreted as the constant relative (economic) risk aversion utility, where $1/\varepsilon_D$ is the coefficient of relative risk aversion [Henderson & Hobson, 2002] and ε_D the inter-temporal elasticity of substitution which quantities the willingness of the representative demander to substitute consumption over time [Section 8.2.3. in Acemoglu, 2008]. As expected for numéraire goods [Konishi & Fishburn, 1996], u_2 is a linear function (hence also increasing and concave).

The budget constraint in (E.26) implies that the shape of $\delta(p)$ is governed by an optimal trade-off between the utility derived from the focal goods (u_1) and the utility derived from all the other goods not explicitly accounted for in the FTM model (u_2).

Proof of Property 2 is as follows :

Following standard microeconomic models [Mas-Colell *et al.*, 1995], we assume the demander chose δ_x and $\widehat{\delta_y}$ to maximise his/her utility $u(\delta_x, \widehat{\delta_y}\ )$ subject to a budget constraint. Let b represent the budget generation rate of the demander. p (p_y) is the price at which the focal good (the other goods) is (are) sold. Since markets other than the focal are not explicitly accounted for in the FTM model, we can reasonably assume $p_y = p_0$, i.e. the price of the other goods is set to the price of the numéraire.

Assuming the budget is entirely invested at each time step, the budget constraint reads : $p\ \delta_x + p_0\ \widehat{\delta_y} = b$. Formally, the optimisation problem of the demander can be written :

$$\begin{aligned}
&\max_{(\delta_x, \widehat{\delta_y})} \{u(\delta_x, \widehat{\delta_y}\ )\} \ , \\
&\text{subject to : } p\delta_x + p_0\widehat{\delta_y} = b \ .
\end{aligned} \tag{E.27}$$

We solve problem (E.27) with the Lagrangian multiplier method [e.g. Rowthorn *et al.*, 2009]. Let $\mathcal{L}$ represent the Lagrangian function associated with (E.27). By definition, we have :

$$\mathcal{L}(\delta_x, \widehat{\delta_y}\ , \zeta) = u(\delta_x, \widehat{\delta_y}\ ) + \zeta(b - p\delta_x - p_0\widehat{\delta_y})\ , \tag{E.28}$$

where ζ is the Lagrangian multiplier.

The first order conditions associated with (E.28) are :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \delta_x} &= \frac{\partial u}{\partial \delta_x} - \zeta p = 0\ , \\ \frac{\partial \mathcal{L}}{\partial \widehat{\delta_y}} &= \frac{\partial u}{\partial \widehat{\delta_y}} - \zeta p_0 = 0\ , \\ \frac{\partial \mathcal{L}}{\partial \zeta} &= b - p\delta_x - p_0\widehat{\delta_y} = 0\ . \end{aligned} \tag{E.29}$$

The first two conditions in (E.29) imply :

$$\zeta = \frac{\frac{\partial u}{\partial \delta_x}}{p} = \frac{\frac{\partial u}{\partial \widehat{\delta_y}}}{p_0}\ . \tag{E.30}$$

Equation (E.30) leads to :

$$\frac{\frac{\partial u}{\partial \delta_x}}{\frac{\partial u}{\partial \widehat{\delta_y}}} = \frac{p}{p_0}\ . \tag{E.31}$$

The per-capita generation rate in demanded stock (E.19) can be rewritten :

$$\left(\frac{\delta(p)}{\delta_0}\right)^{-1/\varepsilon_D} = \frac{p}{p_0}\ . \tag{E.32}$$

By identification, the left-hand sides of (E.31) and (E.32) are equal and we have :

$$\frac{\partial u}{\partial \delta_x}(\delta_x = \delta(p)) = \frac{\partial u}{\partial \widehat{\delta_y}}\left(\frac{\delta_x}{\delta_0}\right)^{-1/\varepsilon_D}\ . \tag{E.33}$$

Assuming the utility function u is quasilinear (separable and linear in the numéraire good) as in partial equilibrium models [Singh & Vives, 1984; Konishi & Fishburn, 1996], we have the decomposition :

$$\begin{aligned} \text{[total expected utility]} &= \text{[expected utility from the focal goods]} + \text{[expected utility from the other goods]}\ , \\ u(\delta_x, \widehat{\delta_y}\ ) &= u_1(\delta_x) + \underbrace{\widehat{\delta_y}}_{u_2(\widehat{\delta_y}\ )}\ . \end{aligned} \tag{E.34}$$

From (E.34), we have $\frac{\partial u}{\partial \widehat{\delta_y}} = 1$. Equation (E.33) hence reduces to :

$$\frac{\partial u}{\partial \delta_x}(\delta_x = \delta(p)) = \frac{du_1}{d\delta_x}(\delta_x = \delta(p)) = \left(\frac{\delta_x}{\delta_0}\right)^{-1/\varepsilon_D} \quad . \tag{E.35}$$

which directly yields $(E.26)$ after integration.

Elementary transactions underlying the transaction rate

Let $\theta_{i,j}$ represent the elementary transaction rate from the representative seller i to the representative demander j. We highlight here the implicit microeconomic assumptions on $\theta_{i,j}$ that lead to Θ, the transaction rate of the FTM model (equation (6) in the main text).

The total generation rate of supply stock of i (generation rate of demanded stock of j) is given by $\sigma(p)$ ($\delta(p)$). We denote by $f_{i,j}$ ($g_{i,j}$) the share of $\sigma(p)$ allocated to j by i (the share of $\delta(p)$ allocated to i by j). We assume agents share their respective generation rates of supply and demanded stocks equally among their business partners but without allowing self-loops (i.e. a wholesaler cannot exchange goods with himself/herself), which translates into :

$$f_{i,j} = \frac{1}{N_D - \langle \mathbb{1}_{i \in \mathcal{N}_D} \rangle} = \frac{1}{N_D - \frac{N_{S \cap D}}{N_S}} \quad \text{(supplier } i \text{ to demander } j\text{)} \ , \tag{E.36}$$

$$g_{i,j} = \frac{1}{N_S - \langle \mathbb{1}_{j \in \mathcal{N}_S} \rangle} = \frac{1}{N_S - \frac{N_{S \cap D}}{N_D}} \quad \text{(demander } j \text{ to supplier } i\text{)} \ , \tag{E.37}$$

where $\langle \mathbb{1}_{i \in \mathcal{N}_D} \rangle$ and $\langle \mathbb{1}_{j \in \mathcal{N}_S} \rangle$ represent the probabilities for the supplier i and the demander j, respectively, to be a wholesaler.

Since an indivisible good cannot be further divided, the current maximal possible rate of exchange of indivisible goods from i to j is $\min\{f_{i,j}\ \sigma(p)\ ; g_{i,j}\ \delta(p)\}$. Let $\kappa_{i,j}$ represent elementary friction, for instance the minimal shipment size from i to j. We assume $\kappa_{i,j} = \kappa$ as in the FTM model. It follows :

$$\theta_{i,j} = \frac{\min\{f_{i,j}\ \sigma(p); g_{i,j}\ \delta(p)\ \}}{\kappa} = \frac{\min\{N_S \sigma(p); N_D \delta(p)\ \}}{\kappa(N_S N_D - N_{S \cap D})} \ . \tag{E.38}$$

The macroscopic transaction rate Θ is obtained by summing the elementary transaction rates $\theta_{i,j}$ over all supplier-demander pairs without self-loops :

$$\Theta = \sum_{i,j,j \neq i} \theta_{i,j} = \theta_{i,j}(N_S N_D - N_{S \cap D}) = \frac{\min\{N_S \sigma(p); N_D \delta(p)\ \}}{\kappa} \ , \tag{E.39}$$

which is the definition of the transaction rate of the FTM model ((6) in the main text).

The FTM model at the microscopic scale : evolution of supply, demand and price

Let $s_i = s$ and $d_j = d$ represent the elementary supply and demanded stocks of the representative supplier i and the representative demander j respectively. From the above considerations, the microeconomic FTM model reads :

$$\begin{aligned} \frac{ds_i}{dt} &= \sigma(p) - \sum_{j \neq i} \theta_{i,j}\ q_{i,j}\ , \\ \frac{dd_j}{dt} &= \delta(p) - \sum_{i \neq j} \theta_{i,j}\ q_{i,j}\ , \\ p(t) &= p(0) \exp\{\mu\,[D(t) - S(t) - (D(0) - S(0))]\}\ , \end{aligned} \tag{E.40}$$

where $q_{i,j} = \min\{s_i; d_j\} = \min\{s; d\}$ by definition. Note that price dynamics can only be defined at the macroscopic level since agents are assumed to be price-takers.

By summation over the N_S suppliers and the N_D demanders, we immediately notice that system (E.40) is identical to system (E.41) with corresponding stocks $S = s\ N_S$ and $D = d\ N_D$. In other words, we recover the macroscopic FTM model from the microscopic FTM model.

Potential extensions of the FTM model

The FTM model implicitly assumes that friction κ does not directly influence $\sigma(p)$ and $\delta(p)$, which implies that changing κ leads to identical trade flow at steady-state equilibrium (Section E.3.3). The latter assumption is reasonable as first approximation, because it is difficult to foresee in practice whether changing friction will actually increase or decrease overall costs of production in supply and demand. Let us take the simple case where trade friction κ reduces to the minimal shipment size, i.e. $\kappa = \kappa_{\text{delivery}}$ (equation (9) in the main text). On the one hand, a larger κ means that agents can ship more units of goods per transaction. A larger minimal shipment size is hence expected to reduce transaction costs [see Williamson & Masten, 1999; Klos & Nooteboom, 2001; Tesfatsion, 2001; Klaes, 2008, for further details on transaction costs and their influence on trade dynamics]. On the other hand, increasing κ implies that agents need to accumulate more units of goods for a transaction to occur. It follows a greater minimal shipment size probably translates into larger maintenance costs and special investments to allow for large stocks at equilibrium, i.e. increasing maximal capacities of holdings. The microeconomic optimisation programs of the representative supplier (E.21) and representative demander (E.26) could be easily generalised to quantify the direct influence of κ on σ and δ. For instance, explicit costs depending on κ could be added in the production costs for the supplier (E.21) and in the budget constraint for the demander (E.26). Since we lack empirical data on the cost structure underpinning friction, we assume that costs associated with κ are included in sunk costs (c_{sunk} in (E.21)) and the budget generation rate (b in (E.26)). We leave the extension to more complex cost functions for future work.

Other potential extensions of interest to the FTM model include price stickiness, another type of friction induced by rigid prices [see Silvestre, 2008, for further details], and the generalisation of instant optimisation programs ((E.21) and (E.26)) to account for inter-temporal optimisation [Mas-Colell *et al.*, 1995; Acemoglu, 2008].

E.3.2 Existence and unicity of solutions

For any $a \geq 0$ and $b \geq 0$, we consider the function f defined as :

$$\begin{aligned} \mathbb{R}^{+2} &\to \mathbb{R}^{+} \\ f : (x, y) &\mapsto f(x, y) = \min(ax; by) \ . \end{aligned}$$

Property 3 : f is Lipschitz continuous with Lipschitz constant $M = \max(a; b)$.

Proof of Property 3 is as follows :

We first rewrite f as :

$$f(x, y) = \frac{ax + by - |ax - by|}{2} \ .$$

Then we have :

$$\begin{aligned} f(x_2, y_2) - f(x_1, y_1) &= \frac{a(x_2 - x_1) + b(y_2 - y_1) - (|ax_2 - by_2| - |ax_1 - by_1|)}{2} \\ &\leq \frac{a|x_2 - x_1| + b|y_2 - y_1| + |\ |ax_2 - by_2| - |ax_1 - by_1|\ |}{2} \ . \end{aligned}$$

By the reverse triangle inequality, we have :

$$f(x_2, y_2) - f(x_1, y_1) \leq \frac{a|x_2 - x_1| + b|y_2 - y_1| + |a(x_2 - x_1) - b(y_2 - y_1)|}{2} \ .$$

The triangle inequality finally yields :

$$f(x_2, y_2) - f(x_1, y_1) \leq a|x_2 - x_1| + b|y_2 - y_1| \leq \max(a; b)(|x_2 - x_1| + |y_2 - y_1|) \ .$$

Apart from the min function, the other functions used in the FTM and ME models are canonical and known to be Lipschitz on sets of epidemiological and economic relevance. We deduce directly, by usual operations (composition, multiplication, addition), that all functions implemented in the FTM and ME models (the right-hand sides of the ordinary differential equations in the main text) are Lipschitz.

From the existence and uniqueness theorem [Chapters 7 and 17 in Hirsch *et al.*, 2004], it follows that solutions to the associated ordinary differential equations are unique and exist for all initial conditions.

E.3.3 Steady-state equilibria and stability analyses

We analyse the FTM model analytically with a bottom-up approach. First, we define and analyse a *reference market* without stock loss nor external flows. Second we show that the *law of supply and demand* (LSD) is spanned by the reference market. Third we explore the impacts of *stock loss* on trade dynamics. Fourth we analyse the influence of *external flows* on trade dynamics.

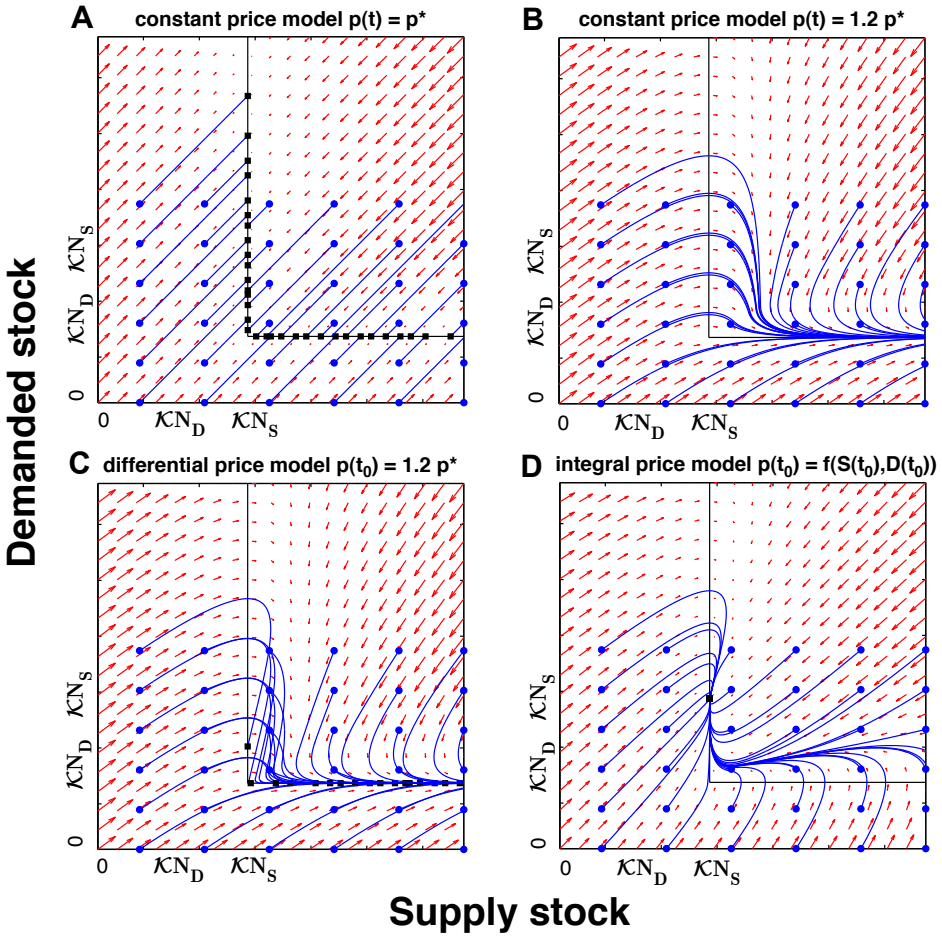


FIGURE E.2 – **Impacts of various price models on reference market dynamics.** Joint-evolution of the demanded stock $D(t)$ (y-axis) as function of the supply stock $S(t)$ (x-axis) for various price dynamics : constant price model ($\mu = 0$) with $p = p^*$ and $p = 1.2\ p^*$ (Panels **A-B**), differential price model with $p = 1.2\ p^*$ (Panel **C**) and integral price model corresponding to the LSD (Panel **D**). For the four cases, market dynamics correspond to the reference market with $\kappa = 1$. We set $\mu = 0$ for the constant price model and $\mu = 10^{-6}$ for the differential and integral price models. Other parameters are from the cattle dataset. Red arrows represent the vector field governing the dynamics of S and D. For any initial condition $(S(t_0), D(t_0))$, each arrow hence represents the vector $\mathbf{v} : (\frac{dS}{dt}, \frac{dD}{dt})_{S=S(t_0),D=D(t_0),p=p(t_0)}$. The blue dots represent a subset of the latter initial conditions for which we explored the full trajectories of the model numerically ; the blue lines and the black squares represent the associated dynamics and steady-states respectively. The vertical ($S = S^*_{min} = \kappa N_S$) and horizontal ($D = D^*_{min} = \kappa N_D$) black lines represent the theoretical steady-state equilibria (S^*, D^*) of the differential price model. The point $S^* = D^* = \max\{\kappa N_S, \kappa N_D\}$ is the only steady-state equilibrium of the integral price model that does not violate the LSD, as predicted theoretically and observed numerically.

Trade without stock loss nor external flows : the reference market

We neglect stock loss ($L_S = L_D = 0$) and external flows ($E_S = E_D = 0$). This case is referred to as the *reference market* and the value of its variables at steady-state equilibrium are denoted throughout with a *star* in superscript. Market dynamics reduce to :

$$\begin{aligned}
\frac{dS}{dt} &= \overbrace{N_S\ \sigma_0\ p^{\varepsilon_S}}^{\Sigma_\oplus} - \frac{\overbrace{\min\{\Sigma_\oplus;\Delta_\oplus\}}^{\Theta}}{\kappa}\ \overbrace{\min\left\{\frac{S}{N_S};\frac{D}{N_D}\right\}}^{q}\ , \\
\frac{dD}{dt} &= \overbrace{N_D\ \delta_0\ p^{-\varepsilon_D}}^{\Delta_\oplus} - \overbrace{\Theta q}^{\Phi}\ , \\
\frac{dp}{dt} &= \mu\frac{d(D-S)}{dt}\ p = \mu(\Delta_\oplus - \Sigma_\oplus)p\ .
\end{aligned} \tag{E.41}$$

The case $\mu = 0$ corresponds to a model where price is assumed to remain constant and equal to its initial value, i.e. $p(t) = p(t_0)$. Such a model is known in the economic literature as a 'fixprice model' and corresponds to an example of 'price stickiness' [Silvestre, 2008]. The subcase $p(t_0) = p^*$ is treated below (see also Fig. E.2**A**). Due to the functional shapes of $\Sigma_\oplus$ and $\Delta_\oplus$, the subcase $p(t_0) \neq p^*$ implies that either $\frac{dS}{dt}$ or $\frac{dD}{dt}$ is never equal to 0. The market is hence diverging in either supply or demand and $p(t_0)$ is a bifurcation parameter. For instance, the constant price model with $p = 1.2p^*$ diverges in supply (Fig. E.2**B**). The more interesting case $\mu > 0$ is analysed below.

Steady-state equilibria.

We assume from now that $\mu > 0$, i.e. we allow for price dynamics.

By construction, $p = p^* = \left(\frac{N_D\delta_0}{N_S\sigma_0}\right)^{\frac{1}{\varepsilon_S+\varepsilon_D}}$ equalises $\Sigma_\oplus$ and $\Delta_\oplus$. Given trade flow $\Phi = \Theta q$ acts symmetrically on supply and demand, it follows $p = p^*$ is the only value of p for which the market is equilibrated. As $p = p^*$ is unique, reference trade flow Φ^* is equal to $\Sigma_\oplus(p^*) = \Delta_\oplus(p^*)$ and is unique. $\Theta(p^*) = \Theta^*$ is also unique.

Φ^* is the only value of Φ that satisfies $\frac{dS}{dt} = \frac{dD}{dt} = 0$. To determine the steady-state equilibrium values of supply and demanded stocks, we hence solve $q(S^*, D^*) = \min\{\frac{S^*}{N_S};\frac{D^*}{N_D}\} = \frac{\Phi^*}{\Theta^*} = \kappa$. Due to the min function, the system is characterised by two sets of infinite steady-state equilibria : either $\{S^* = \kappa N_S \text{ and } D^* \geq \kappa N_D\}$ or $\{S^* \geq \kappa N_S \text{ and } D^* = \kappa N_D\}$. $S^*_{min} = \kappa N_S$ and $D^*_{min} = \kappa N_D$ are the minimal supply and demanded stocks for which the market is at steady-state equilibrium. Trade flow at steady-state equilibrium is hence given by :

$$\Phi^* = [N_S\sigma_0]^{\frac{\varepsilon_D}{\varepsilon_S+\varepsilon_D}}\ [N_D\delta_0]^{\frac{\varepsilon_S}{\varepsilon_S+\varepsilon_D}}\ . \tag{E.42}$$

While trade flow at the reference market steady-state equilibrium is unique and does not depend on κ, S^* and D^*, i.e. persistent frustrations, are positively associated with κ.

Stability.

From system (E.41), we notice that $\frac{dp}{dt}$ is a function of variables p, $\Sigma_\oplus(p, N_S)$ and $\Delta_\oplus(p, N_D)$. Since N_S and N_D are constant in the FTM model, $\frac{dp}{dt}$ only depends on p.

We first analyse the global stability of p. We define the function $V(p) = (p - p^*)^2$. V has the following properties :

$$\begin{aligned} &V(p^*) = 0 \text{ and } V(p) > 0 \text{ for } p \neq p^* \ , \\ &\frac{dV}{dt} = 2\frac{dp}{dt}(p - p^*) = 2\mu p \left(\Delta_\oplus(p, N_D) - \Sigma_\oplus(p, N_S)\right)(p - p^*) < 0 \text{ for } p \neq p^* \ , \end{aligned} \tag{E.43}$$

since by construction $\Delta_\oplus(p, N_D) < \Sigma_\oplus(p, N_S) \Leftrightarrow p > p^*$.

We conclude V is a strict Liapunov function and the point $p = p^*$ is asymptotically stable [Section 9.2 of Hirsch *et al.*, 2004].

Given that the point $p = p^*$ is eventually reached by the system, we only need to explore the joint-dynamics of (S, D) from the point of time t_x where $p(t_x) = p^*$. We do not know the precise values of t_x and $[S(t_x), D(t_x)]$. We can still conclude on the stability of system (E.41) if we can analyse the dynamics of the system (S, D) with $p = p^*$ for *any* initial conditions $[S(t_0), D(t_0)]$ because $[S(t_0), D(t_0)] \supset [S(t_x), D(t_x)]$.

From now on, we analyse such a system (S, D) with initial conditions $[S(t_0), D(t_0)]$ and where $p = p^*$. Since $p = p^*$, we know that : $\Sigma_\oplus = \Delta_\oplus = \Phi^*$. It follows $\frac{dS}{dt} = \frac{dD}{dt}$, $\forall t$. The equality of the derivatives means that there is a linear vector field associated with any initial values of (S, D). In other words, the dynamics of (S, D) are hence fully described by a line $D(t) = S(t) + a$ where a is a constant that depends on initial conditions. For $t = t_0$, $D(t_0) = S(t_0) + a$, which leads to $a = D(t_0) - S(t_0)$. Full market dynamics are hence given by the line :

$$D(t) = S(t) + D(t_0) - S(t_0) \ . \tag{E.44}$$

The direction of the dynamics along the line described by equation (E.44) are found by solving the inequality $\frac{dS}{dt} = \frac{dD}{dt} < 0$:

$$\begin{aligned} &\frac{dS}{dt} < 0 \Leftrightarrow \Phi^* - \Theta^* q < 0 \ , \\ &\frac{dS}{dt} < 0 \Leftrightarrow q(t) = \min\left\{\frac{S(t)}{N_S}; \frac{D(t)}{N_D}\right\} > \kappa \ , \\ &\frac{dS}{dt} < 0 \Leftrightarrow [S(t) > \kappa N_S \quad \text{AND} \quad D(t) > \kappa N_D] \ , \end{aligned}$$

which is equivalent by logical negation to : $\frac{dS}{dt} > 0 \Leftrightarrow [S(t) < \kappa N_S \quad \text{OR} \quad D(t) < \kappa N_D]$. Those equivalences prove that the average-exchange stock $q(t)$ always converges to κ. It follows the system is globally stable in trade flow, i.e. $\Phi(t) \to \Phi^*$ for any initial conditions.

Given relation (E.44) and our stability analysis, the position of $(D(t_0) - S(t_0))$ with respect to $(D^*_{min} - S^*_{min})$ separates the state space (S, D) in two subparts that determine the final state of the system :

$$\begin{aligned}
&\text{If } (D(t_0) - S(t_0)) \geq (D^*_{min} - S^*_{min}) : \quad (S^* = S^*_{min} \ , \ D^* = S^*_{min} + (D(t_0) - S(t_0))) \ . \\
&\text{If } (D(t_0) - S(t_0)) \leq (D^*_{min} - S^*_{min}) : \quad (S^* = S^*_{min} - (D(t_0) - S(t_0)) \ , \ D^* = D^*_{min}) \ .
\end{aligned}$$

The case $(D(t_0) - S(t_0)) \geq (D^*_{min} - S^*_{min})$ corresponds to an initial positive excess demand which implies a limiting supply stock at steady-state equilibrium. The case $(D(t_0) - S(t_0)) \leq (D^*_{min} - S^*_{min})$ corresponds to an initial negative excess demand which implies a limiting demanded stock at steady-state equilibrium.

Since the steady-state equilibria in supply and demanded stocks depend on initial conditions, there is an infinite number of unstable steady-state equilibria (S^*, D^*) with a switched fixed point : either $S^* = S^*_{min}$ or $D^* = D^*_{min}$.

In conclusion, the system always converges to reference flow $(\Phi(t) \to \Phi^*)$ and reference price $(p(t) \to p^*)$ for any initial conditions or external perturbations. In other words, our model market is globally stable in (measurable) price and flow while extremely unstable in (hidden) frustrations. Our key findings are summarized graphically in Fig. E.2**C**.

The law of supply and demand (LSD) is spanned by the reference market

One could argue that our pricing model is unsatisfactory since S and D are not necessarily identical at equilibrium. The LSD indeed stipulates that price should increase in response to any excess demand $D - S > 0$. There is little empirical support for the LSD [McCauley, 2009], which is consistent with the fact that S and D are not observable. For comparative purposes with the existing literature, we nevertheless investigate the LSD analytically and show that it is a special case spanned by the reference market.

By integration of system (E.41), we have the equivalence :

$$\begin{aligned}
&(\mathcal{P}_d) : \quad \frac{dp}{dt} = \mu \frac{d(D-S)}{dt} \ p \text{ with initial conditions : } [S(t_0) = S_0, D(t_0) = D_0, p(t_0) = p_0] \ , \\
&\qquad\qquad\qquad\qquad \Leftrightarrow \\
&(\mathcal{P}_i) : \quad p(t) = p_0 \exp\left(\mu \left[D(t) - S(t) - (D_0 - S_0)\right]\right) \ ,
\end{aligned} \tag{E.45}$$

where $(\mathcal{P}_d)$ and $(\mathcal{P}_i)$ represent the differential and integral formulations of our pricing model $\mathcal{P}$, and S_0, D_0 and p_0 are constants.

System (E.41) is hence equivalent to a new system where :

$$\begin{aligned}
\frac{dS}{dt} &= \overbrace{N_S\ \sigma_0\ p^{\varepsilon_S}}^{\Sigma_\oplus} - \overbrace{\frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}}^{\Theta}\overbrace{\min\left\{\frac{S}{N_S};\frac{D}{N_D}\right\}}^{q}\ ,\\
\frac{dD}{dt} &= \overbrace{N_D\ \delta_0\ p^{-\varepsilon_D}}^{\Delta_\oplus} - \overbrace{\Theta q}^{\Phi}\ ,\\
p(t) &= p_0 \exp\left(\mu\left[D(t) - S(t) - (D_0 - S_0)\right]\right)\ .
\end{aligned} \tag{E.46}$$

Importantly, (E.46) is only identical to (E.41) is we take the same initial conditions, i.e. if we set $[S(t_0) = S_0, D(t_0) = D_0, p(t_0) = p_0]$. Now we analyse system (E.46) in the case where the constants $[S_0, D_0, p_0]$ are not necessarily equal to $[S(t_0), D(t_0), p(t_0)]$. From now, we only consider the case where $p_0 = p^*$. We want to find the values $[S_0, D_0]$ that are compatible with the LSD.

Since $(\mathcal{P}_i) \Rightarrow (\mathcal{P}_d)$ (the converse proposition is not true any more), we have the same properties as already described in the analysis of system (E.41). But we have new properties thanks to the $(\mathcal{P}_d)$ model. We know : $p(t) = p^* \exp\left(\mu\left[D(t) - S(t) - (D_0 - S_0)\right]\right)$. Since $p = p^*$ is asymptotically stable, we solve equation $p(t) = p^*$ at steady-state equilibrium to yield a new relationship :

$$D^* - S^* = D_0 - S_0\ , \tag{E.47}$$

that is specific to the $(\mathcal{P}_d)$ model.

Steady-state equilibria.

At steady-state equilibrium, we have now two relationships between S^* and D^* :

$$\begin{aligned}
&D^* - S^* = D_0 - S_0\ ,\\
&q(S^*, D^*) = \min\left\{\frac{S^*}{N_S};\frac{D^*}{N_D}\right\} = \kappa\ ,
\end{aligned} \tag{E.48}$$

which lead to a unique steady-state equilibrium whose value depends on the position of $(D_0 - S_0)$ with respect to $(D^*_{min} - S^*_{min})$:

$$\begin{aligned}
&\text{If } (D_0 - S_0) \geq (D^*_{min} - S^*_{min}) : \quad (S^* = S^*_{min}\ ,\ D^* = S^*_{min} + (D_0 - S_0))\ .\\
&\text{If } (D_0 - S_0) \leq (D^*_{min} - S^*_{min}) : \quad (S^* = D^*_{min} - (D_0 - S_0)\ ,\ D^* = D^*_{min})\ .
\end{aligned} \tag{E.49}$$

From (E.49), we notice that the LSD corresponds to the case $(D_0 - S_0) = 0$, i.e. $S_0 = D_0$. If $N_D > N_S$, the LSD is hence given by $S^* = D^* = \kappa N_S$. While if $N_D < N_S$, the LSD is given by $S^* = D^* = \kappa N_D$. For both cases, the LSD leads to $S^* = D^* = \max\{\kappa N_D, \kappa N_S\}$.

Note that if agents were rational, they would try to minimise their stocks while keeping trade flow Φ^* constant : this corresponds instead to the case $(D_0 - S_0) = (D^*_{min} - S^*_{min})$, i.e. $S_0 = D_0 - (D^*_{min} - S^*_{min})$, for which $(S^* = S^*_{min}, D^* = D^*_{min})$ according to (E.49).

For any constants (S_0, D_0) taken as a reference, the mismatch between supply and demanded stocks can be assessed with a generalisation of the excess demand, denoted $E(S_0, D_0)$, that accounts for imperfect transactions with friction :

$$E(S_0, D_0) \equiv D(t) - S(t) - (D_0 - S_0) \tag{E.50}$$

Stability.

We know $(\mathcal{P}_i) \Rightarrow (\mathcal{P}_d)$. It follows our former analysis stands and the LSD steady-state equilibrium $[p^*, S^* = D^* = \max\{\kappa N_D, \kappa N_S\}]$ is asymptotically stable. The integral version of our pricing model is hence compatible with the expected properties of the LSD. As predicted, the integral price model with $S_0 = D_0$ converges to a unique steady-state equilibrium $S^* = D^* = \max\{\kappa N_S, \kappa N_D\} = \kappa N_S$ in the case of cattle data (Fig. E.2**D**).

Trade with stock loss : the general case

We now allow for stock loss in the FTM model, which leads to the system :

$$\begin{aligned}
\frac{dS}{dt} &= \overbrace{N_S\ \sigma_0\ p^{\varepsilon_S}}^{\Sigma_\oplus} - r_S S - \overbrace{\frac{\min\{\Sigma_\oplus; \Delta_\oplus\}}{\kappa}}^{\Theta} \overbrace{\min\left\{\frac{S}{N_S}; \frac{D}{N_D}\right\}}^{q}, \\
\frac{dD}{dt} &= \overbrace{N_D\ \delta_0\ p^{-\varepsilon_D}}^{\Delta_\oplus} - r_D D - \overbrace{\Theta q}^{\Phi}, \\
\frac{dp}{dt} &= \mu \frac{d(D-S)}{dt}\ p = \mu\left(\Delta(D,p) - \Sigma(S,p)\right)\ p,
\end{aligned} \tag{E.51}$$

with $\Sigma(S,p) = \Sigma_\oplus - r_S S$ and $\Delta(D,p) = \Delta_\oplus - r_D D$. Notice that in contrast with the reference market (system (E.41)), $\dfrac{dp}{dt}$ now depends on S, D and p.

Steady-state equilibria.

When stock loss is included, we only know that the number of steady-state equilibria is not unique and we cannot derive the steady-state equilibrium values of state variables explicitly. Let *eq* in subscript denote steady-state equilibrium values in the general case (in contrast with the special case of the reference market where steady-state equilibrium is denoted by a star in superscript). At steady-state equilibrium, $\Sigma(S_{eq}, p_{eq}) = \Delta(D_{eq}, p_{eq})$ is indeed the only independent equation of system (E.51) which has three unknowns (S_{eq}, D_{eq}, p_{eq}). However we can still extract useful information on the system. If

we e.g. only consider positive stock loss ($r_S \geq 0$ and $r_D \geq 0$), we know for sure from (E.51) that :

$$\begin{aligned} \Phi_{eq} &= \Sigma_{\oplus}^{eq} - r_S S_{eq} \leq \Sigma_{\oplus}^{eq} \ , \\ \Phi_{eq} &= \Delta_{\oplus}^{eq} - r_D D_{eq} \leq \Delta_{\oplus}^{eq} \ . \end{aligned} \tag{E.52}$$

From (E.52) we conclude that equilibrium flow is necessarily suboptimal as :

$$\Phi_{eq} \leq \min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} \leq \Phi^* = \min\{\Sigma_{\oplus}^*, \Delta_{\oplus}^*\} \tag{E.53}$$

A deeper analysis enables us to quantify the effects of κ, r_S and r_D on steady-state equilibrium trade flow. We know in the general case that :

$$\begin{aligned} &\Phi_{eq} = \Sigma_{eq} = \Delta_{eq} = \Sigma_{\oplus}^{eq} - r_S S_{eq} = \Delta_{\oplus}^{eq} - r_D D_{eq} \ , \\ &q_{eq} \equiv \frac{\Phi_{eq}}{\Theta_{eq}} \Rightarrow \min\{\frac{S_{eq}}{N_S}; \frac{D_{eq}}{N_D}\} = \frac{\kappa \Phi_{eq}}{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\}} \ . \end{aligned} \tag{E.54}$$

Then two cases need to be considered depending on the limiting factor of q_{eq}.

If q_{eq} is limited by the per capita supply stock, i.e. if $q_{eq} = \frac{S_{eq}}{N_S}$, we find from (E.54) that $S_{eq} = \frac{\kappa N_S \Sigma_{\oplus}^{eq}}{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} + r_S \kappa N_S}$. We hence deduce the value of q_{eq}. Given that $\Phi_{eq} = \Theta_{eq} q_{eq}$, we find :

$$\Phi_{eq} = \Sigma_{\oplus}^{eq} \frac{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\}}{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} + r_S \kappa N_S} \tag{E.55}$$

If q_{eq} is limited by the per capita demanded stock, i.e. if $q_{eq} = \frac{D_{eq}}{N_D}$. We directly find by symmetry :

$$\Phi_{eq} = \Delta_{\oplus}^{eq} \frac{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\}}{\min\{\Sigma_{\oplus}^{eq}, \Delta_{\oplus}^{eq}\} + r_D \kappa N_D} \tag{E.56}$$

Importantly, equations (E.55)-(E.56) imply that r_S, r_D and κ have a negative impact on steady-state equilibrium flow. In particular, equations (E.55)-(E.56) imply :

$$\lim_{r_S \kappa \to \infty} \Phi_{eq} = \lim_{r_D \kappa \to \infty} \Phi_{eq} = 0 \tag{E.57}$$

Since the dynamics with stock loss are not fully analytically tractable in the general case, we resort to extensive numerical simulations to confirm the key influence of r_S, r_D and κ on trade dynamics (see Global Sensitivity Analysis (GSA) of the FTM model). In addition, the special case analysed thereafter enables us to get exhaustive analytical insights on the impact of stock loss on market dynamics.

Stability.

We cannot conclude analytically on the stability of system (E.51) in the general case.

Trade with stock loss : detailed analysis of a special case.

To study market transient behaviour and confirm our general findings, we consider initial conditions and parameter values that enable us to solve system (E.51) analytically. We set $[S(t_0) = D(t_0) \geq 0; p(t_0) = p^*]$ at initial time t_0 and track trade flow Φ until steady-state equilibration. This set of initial conditions is compatible with the LSD since $S(t_0) = D(t_0)$ and $p(t_0) = p^*$. Once accumulated over time at rates $\Sigma_\oplus$ and $\Delta_\oplus$, supply and demanded stocks are converted through trade (Φ) and losses (at rates $r_S S$ and $r_D D$). For simplicity, we consider symmetrical losses ($r_S = r_D = r$). Since we start from $S(t_0) = D(t_0)$ and $p(t_0) = p^*$ as initial conditions, and as N_S and N_D do not change over time, we have $p(t) = p^*$. Hence, when r is set, the coefficient of friction κ is the only factor influencing market dynamics. Since $p(t) = p^*$, it follows that $\Sigma_\oplus(t) = \Delta_\oplus(t) = \Phi^*$ and $\Theta(t) = \frac{\Phi^*}{\kappa}$. Using symmetry arguments in equations describing S and D in system (E.51) and given that we assumed $S(t_0) = D(t_0)$ and $r_S = r_D = r$, we have $S(t) = D(t)\ , \forall t$. We hence have $q(t) = aS(t)$ where $a = \min\{\frac{1}{N_S}, \frac{1}{N_D}\}$ is a dimensionless constant. From the expression of Θ and q, trade flow Φ can be written : $\Phi(t) = \dfrac{a\Phi^* S(t)}{\kappa}$. Hence, system (E.51) reduces to :

$$\dot{S} = \Phi^* - \left(r\ + \frac{a\Phi^*}{\kappa}\right) S\ . \tag{E.58}$$

Solving (E.58) enables us to explicitly describe the market at each point of time. For $t_0 = 0$ and $r \neq -\dfrac{a\Phi^*}{\kappa}$ we have :

$$S(t) = S(t_0) e^{-t(r + \frac{a\Phi^*}{\kappa})} + \frac{\Phi^*}{r + \frac{a\Phi^*}{\kappa}} \left[1 - e^{-t(r + \frac{a\Phi^*}{\kappa})}\right]\ . \tag{E.59}$$

We notice that (E.59) converges if $r > -\dfrac{a\Phi^*}{\kappa}$ and we find :

$$S_{eq} = \frac{\Phi^*}{r + \frac{a\Phi^*}{\kappa}}\ . \tag{E.60}$$

From equation (E.59) with $S(t_0) = 0$, we deduce the analytical expression of $\Phi(t)$ where :

$$\Phi(t) = \Phi^* \frac{\Phi^*}{\Phi^* + \frac{r\kappa}{a}} \left[1 - e^{-t(r\ + \frac{a\Phi^*}{\kappa})}\right]\ . \tag{E.61}$$

If $r > -\dfrac{a\Phi^*}{\kappa}$, $\Phi(t)$ converges to :

$$\Phi_{eq} = \Phi^* \frac{\Phi^*}{\Phi^* + \frac{r\kappa}{a}}\ . \tag{E.62}$$

As expected from the analysis of the general case, we reach reference flow ($\Phi_{eq} = \Phi^*$) when stock loss is neglected ($r = 0$) and sub-optimal flow ($\Phi_{eq} < \Phi^*$) when stock loss is positive ($r > 0$). If $r \in]-\dfrac{a\Phi^*}{\kappa}, 0[$ flow converges and is over-optimal ($\Phi_{eq} > \Phi^*$). Finally if $r \leq -\dfrac{a\Phi^*}{\kappa}$, flow diverges to infinity ($\Phi \to \infty$).

For the realistic case where stock loss is strictly positive ($r > 0$), κ and r have a negative impact on steady-state equilibrium flow and we recover the key results of equations (E.55)-(E.56).

Trade with external flows

For sake of concision, we neglect stock loss ($r_S = r_D = 0$) and only allow for positive and symmetric external flow ($E_S = E_D = E > 0$). This corresponds to the case explored in the main text to show the impact of imports on epidemics (Figure 3C). We hence have :

$$\begin{aligned}
\frac{dS}{dt} &= \overbrace{N_S\ \sigma_0\ p^{\varepsilon_S}}^{\Sigma_\oplus} + E - \overbrace{\frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}}^{\Theta} \overbrace{\min\left\{\frac{S}{N_S};\frac{D}{N_D}\right\}}^{q}\ , \\
\frac{dD}{dt} &= \overbrace{N_D\ \delta_0\ p^{-\varepsilon_D}}^{\Delta_\oplus} + E - \overbrace{\Theta q}^{\Phi}\ , \\
\frac{dp}{dt} &= \mu(\Delta_\oplus - \Sigma_\oplus)p\ .
\end{aligned} \tag{E.63}$$

As a side remark, the case where any values of E_S and E_D are allowed is not difficult to solve since the dynamics of $\frac{dp}{dt}$ do not depend on S and D. As we do not use it in the main text, we do not develop it in the SI.

steady-state equilibria.

Since we assume symmetrical external flow, the equilibrium in price is the same as the reference case. We hence have $\Theta_{eq} = \Theta^*$. $\Phi_{eq} = \Phi^* + E$ is the only value of Φ that satisfies $\frac{dS}{dt} = \frac{dD}{dt} = 0$. To determine the steady-state equilibrium values of supply and demanded stocks, we hence solve $q(S_{eq}, D_{eq}) \equiv \min\{\frac{S_{eq}}{N_S};\frac{D_{eq}}{N_D}\} = \frac{\Phi_{eq}}{\Theta_{eq}} = \kappa\frac{\Phi^* + E}{\Phi^*}$. Due to the min function, the system is characterised by two sets of infinite steady-state equilibria : either $\{S_{eq} = \kappa N_S\frac{\Phi^* + E}{\Phi^*}$ and $D_{eq} \geq \kappa N_D\frac{\Phi^* + E}{\Phi^*}\}$ or $\{S_{eq} \geq \kappa N_S\frac{\Phi^* + E}{\Phi^*}$ and $D_{eq} = \kappa N_D\frac{\Phi^* + E}{\Phi^*}\}$. It follows $q_{eq} = \kappa\frac{\Phi^* + E}{\Phi^*} > \kappa$

This simple example shows that external flows only act on the supply and demanded stock and hence on the average stock exchanged per transaction q. The transaction rate Θ is unchanged, since it only depends on limited search and delivery budgets that depend on the numbers of suppliers, demanders and price. This finding has implications for epidemics since both Θ and q determine the actual force of infection : if the probability of infection per good is already very large, increasing external flow will increase transmission but only moderately (see Figure 3C in the main text).

Stability.

The proof is the same as the reference market.

E.3.4 Global sensitivity analysis (GSA) of the FTM model

The FTM model includes many parameters, including initial conditions of state variables. To assess the robustness of our key analytical findings to uncertainty and variability in parameter values, we carry out a GSA on key economic outputs.

We rank the relative importance of all parameters of potential importance with an improved version of the Morris method, a GSA technique used to screen the importance of parameters in high-dimensional models [Campolongo *et al.*, 2007]. In a nutshell, the improved Morris method can discriminate the sign and overall influence of factors at a low computational cost and minor risk of error (see below for further details).

Inputs and outputs explored

Table E.1 describes the input parameters of the FTM model and associated ranges explored in the GSA. Each parameter d with set in a range $[a;b]$ is assumed to follow a probability distribution which is either uniform ($d \sim \mathcal{U}(a,b)$) or 'log-uniform' ($\log(d) \sim \mathcal{U}(\log_{10}(a), \log_{10}(b))$. We assume a constant number of suppliers ($N_S = 193,354$) and demanders ($N_D = 118,503$), but allow for a variable proportion of wholesalers in the system ($p_{N_{S\cap D}} \equiv \frac{N_{S\cap D}}{\min\{N_S, N_D\}}$). Baseline rates at which supply and demanded stocks are generated ($\sigma_0 = 39$ per year and $\delta_0 = 64$ per year respectively) are also kept constant. Values of N_S, N_D, σ_0 and δ_0 are derived from the French cattle market (Table 1 in the main text).

TABLE E.1 – **Input parameters and associated ranges explored in the GSA of the FTM model.**

Parameter	Formal notation	Meaning	Range explored	Distribution
k	κ	coefficient of friction	$[10^{-2}\,;\,10^{2}]$	log-uniform
rs	r_S	supply loss rate	[-0.05 ; 33] per year	uniform
rd	r_D	demand loss rate	[-0.05 ; 33] per year	uniform
is	E_S	supply import rate	$[0\,;\,\Phi^*]$ per year	uniform
id	E_D	demand import rate	$[0\,;\,\Phi^*]$ per year	uniform
pNsd	$p_{N_{S\cap D}}$	proportion of wholesalers	[0.5 ; 1]	uniform
St0	$S(t_0)$	initial supply stock	$[0.8\kappa N_S\,;\,1.2\kappa N_S]$	uniform
Dt0	$D(t_0)$	initial demanded stock	$[0.8\kappa N_D\,;\,1.2\kappa N_D]$	uniform
pt0	$p(t_0)$	initial price	[0.8 ; 1.2]	uniform
es	ε_S	price elasticity of supply	[0 ; 3]	uniform
ed	ε_D	price elasticity of demand	[0 ; 3]	uniform
mu	μ	pricing scale parameter	$[0\,;\,10^{-6}/\kappa]$	uniform

The following economic outputs are analysed with the GSA :

- the normalised steady-state equilibrium flow, $\frac{\Phi_{eq}}{\Phi^*} \in \mathbb{R}^+$,
- the time to reach 95 % of steady-state equilibrium flow, $t_{95}(\Phi_{eq}) \in [0;200]$ year,
- the normalised flow at final time $t_f = 200$ year, $\frac{\Phi(t_f)}{\Phi^*} \in \mathbb{R}^+$,
- the excess demand per agent at steady-state equilibrium (E.50), $\frac{E_{eq}(D_0-(D^*_{min}-S^*_{min}),D_0)}{N} = \frac{D_{eq}-S_{eq}-\kappa(N_D-N_S)}{N} \in \mathbb{R}$,
- the steady-state equilibrium price $p_{eq} \in \mathbb{R}^+$,
- and the time to reach 95 % of steady-state equilibrium price, $t_{95}(p_{eq}) \in [0;200]$ year.

Analysis with the improved Morris method

Details on the method

The Morris method is a very efficient GSA technique to explore numerically models with a large parametric space at a low computational cost and minor risk of error. The method is referred to as a global SA technique since it analyses distributions of local elementary effects based on various numerical evaluations of inputs sampled from the parametric space. The improved Morris method ensures that the coverage of the parametric space is improved compared with the standard Morris method by selecting a subset of evaluations that satisfy a maximin-type distance criterion in the parametric space.

More precisely, the Morris method relies on calculating random trajectories with incremental variations along the d parameters $1, ..., i, ..., d$ [see Campolongo *et al.*, 2007, for details]. For a given trajectory, each parameter is selected randomly across l levels, so that a set of $d+1$ values are explored sequentially by changing each one of the parameter at a time (OAT). A total of $r_{\max}$ trajectories are sampled randomly to constitute a basic OAT design. To ensure an optimal coverage of the parameter space, r out $r_{\max}$ trajectories are then chosen to maximise the minimal Haussdorf distance between the $r_{\max}$ selected trajectories (maximin-type optimised design). Incremental ratios called elementary effects EE are then computed. An EE is formally defined as the difference in a given output induced by the difference in a jump of one level for a given parameter. Finally, a distribution of r elementary effects is obtained for each parameter i, denoted $EE(i)$. Three simple metrics are derived from $EE(i)$:

- m^*, the mean of the absolute value of $EE(i)$, which represents the overall influence of the parameter i on the output,
- m, the mean of $EE(i)$, which gives the overall sign of the influence of i (and overall influence of i provided the effect in monotonous),
- s, the standard deviation of $EE(i)$, which jointly quantities the non-linear behaviour of i and interactions of i with other parameters j.

Notice that as various trajectories are explored in the parameter space, the Morris method is a global SA tool. The improvement on the original Morris method consists in the maximin space-filling design and the introduction of m^*. The improved Morris method is implemented in the *morris* function of the *R* package *sensitivity*. In our case we take $r_{\max} = 1000$, $r = 200$ (the value recommended in the literature is 50), $l = 5$. Hence, the impact of $d = 12$ parameters are explored at the cost of $(d+1) * r = 2600$ simulations. In contrast, a standard SA based on an analysis of variance would require $l^d = 244140625$ simulations.

Results

The GSA confirms that the coefficient of friction κ is a key parameter governing trade dynamics.

Sensitivity of $\frac{\Phi_{eq}}{\Phi^*}$

	m*	m	s
k	1.34	-1.27	0.88
is	0.37	0.36	0.33
id	0.35	0.34	0.33
rs	0.28	-0.27	0.82
rd	0.25	-0.25	1.13
mu	0.25	0.07	0.83
pt0	0.15	0.02	0.28
es	0.15	-0.05	0.32
ed	0.13	0.01	0.23
St0	0.01	0.00	0.02
Dt0	0.01	-0.00	0.01
pNsd	0.00	0.00	0.00

Sensitivity of $t_{95}(\Phi_{eq})$

	m*	m	s
k	3.00	3.00	14.30
rs	1.60	-1.01	12.13
rd	1.29	-1.17	10.76
mu	0.48	-0.41	3.02
es	0.28	-0.25	2.45
id	0.08	0.04	0.35
is	0.06	-0.04	0.16
ed	0.05	-0.04	0.39
pt0	0.05	-0.01	0.13
St0	0.01	-0.00	0.03
Dt0	0.01	0.00	0.03
pNsd	0.00	0.00	0.00

Sensitivity of $\frac{\Phi(t_f)}{\Phi^*}$

	m*	m	s
ed	5.26	-5.12	62.64
mu	5.08	0.21	41.98
rs	3.08	-3.06	30.63
k	1.66	-0.95	4.13
id	1.45	1.45	13.58
rd	0.44	-0.43	2.55
pt0	0.36	-0.12	2.21
is	0.36	0.36	0.32
es	0.15	-0.05	0.32
St0	0.06	-0.05	0.66
Dt0	0.05	0.04	0.50
pNsd	0.00	0.00	0.00

Sensitivity of $\frac{D_{eq}-S_{eq}-\kappa(N_D-N_S)}{N}$

	m*	m	s
rs	107.45	107.42	923.38
mu	103.58	42.85	748.16
k	70.25	16.57	326.82
rd	62.45	-61.45	626.41
es	16.68	-16.26	162.17
pNsd	6.46	0.97	32.02
pt0	3.59	-3.58	19.70
id	2.70	2.60	9.54
is	2.24	-2.24	7.68
ed	1.23	0.73	8.75
St0	1.16	-1.15	6.14
Dt0	0.90	0.84	4.23

Sensitivity of p_{eq}

	m*	m	s
es	0.77	-0.71	5.89
rd	0.73	-0.73	7.04
k	0.67	-0.24	2.82
mu	0.66	0.14	3.63
id	0.39	0.39	2.39
is	0.36	-0.36	2.38
pt0	0.22	0.22	0.17
rs	0.21	0.21	0.61
ed	0.09	0.03	0.24
St0	0.02	0.02	0.05
Dt0	0.01	-0.01	0.01
pNsd	0.00	0.00	0.00

Sensitivity of $t_{95}(p_{eq})$

	m*	m	s
k	0.55	0.50	0.92
mu	0.30	-0.29	0.60
pt0	0.08	0.07	0.22
rs	0.05	-0.00	0.08
id	0.03	0.03	0.06
is	0.01	-0.01	0.02
rd	0.01	0.00	0.03
ed	0.01	-0.01	0.01
Dt0	0.01	-0.01	0.02
es	0.01	-0.01	0.01
St0	0.01	0.01	0.01
pNsd	0.00	0.00	0.00

E.3.5 Comparison of supply and demanded stocks with inventories

To ensure our definitions of supply and demanded stocks are economically sound, we investigate the relationship between simulated supply and demanded stocks on the one hand, and livestock inventories as reported in French cattle data on the other hand. Assuming steady-state equilibrium, we explore two aggregation scales : the macroscopic scale implied by the FTM model, and the microscopic scale, based on an agent-based extension of the FTM model.

Macroscopic comparison

Assuming the law of supply and demand is satisfied ($S^* = D^* = \max\{\kappa N_D, \kappa N_S\}$; see section E.3.3 for details), and taking the values of κ, N_D, N_S calculated for French cattle and swine markets (Table 3 in the main text), we find that supply and demanded stocks at steady-state equilibrium are about 60 (20 times) lower than total cattle within holdings in 2009 (total swine farms capacity in 2010), as estimated from data. This feature of our model is in agreement with empirical observations, as agents cannot sell or buy more goods than they own or can stock.

Microscopic comparison

The results at the macroscopic scale appear to hold at the microscopic scale (Figure E.3).

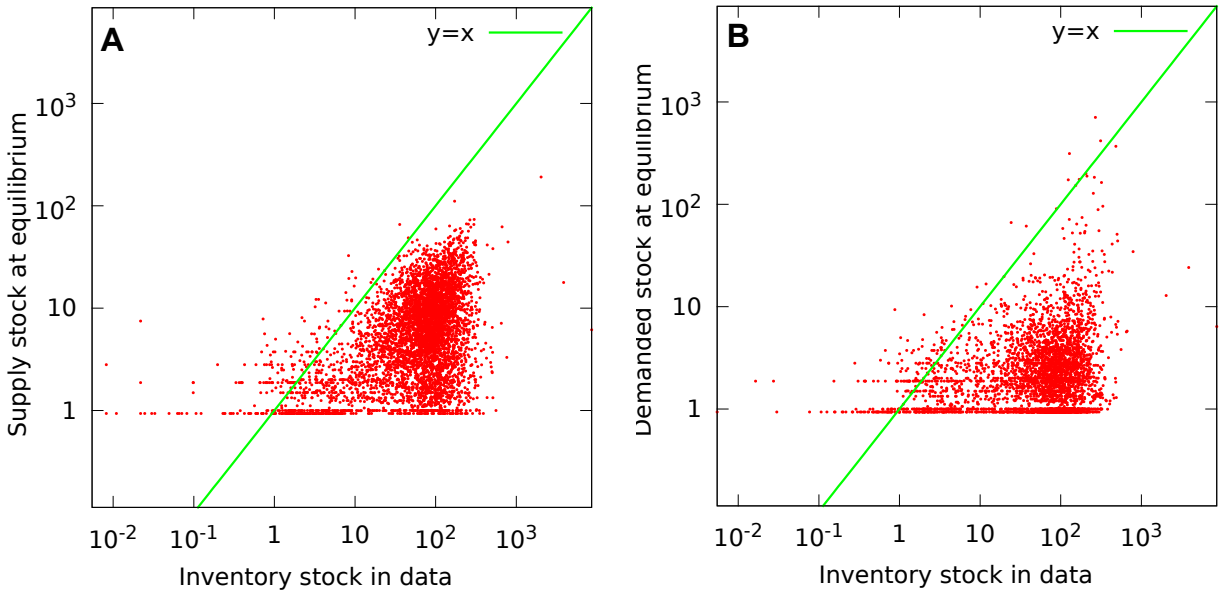


FIGURE E.3 – **Supply and demanded stock as function of inventory stock.** Supply stock (**A**) and demanded stock (**B**) at steady-state equilibrium in a microeconomic extension of the FTM model plotted against recorded inventory stock of cattle holdings in France [see Part III of Moslonka-Lefebvre, 2014]. Each point represents a farm. Supply and demanded stocks are typically lower that the inventory stock, as they appear below the $x = y$ line (in green). To facilitate numerical exploration, we focused on internal animal movements within the Finistère Département of Brittany, one of the most important areas of dairy cattle production in France, over period 2005-2009.

E.4 Supplementary results on the market-epidemiological (ME) model

E.4.1 Analytical insights on the ME model

Presentation of the complete ME model

Keeping the notations of the main text, the complete ME dynamics are given by the following set of ordinary differential equations :

$$
\begin{aligned}
\dot{S} &= \overbrace{\underbrace{N_S^{XY}\ \sigma_0\ p^{\varepsilon_S}}_{\Sigma_\oplus} - \gamma \frac{N_S^Y}{N_S^{XY}} S + E_S}^{\Sigma} - \overbrace{\underbrace{\frac{\min\{\Sigma_\oplus;\Delta_\oplus\}}{\kappa}}_{\Theta} \underbrace{\min\{\frac{S}{N_S^{XY}};\frac{D}{N_D^{XY}}\}}_{q}}^{\Phi} \ , \\
\dot{D} &= \overbrace{\underbrace{N_D^{XY}\ \delta_0\ p^{-\varepsilon_D}}_{\Delta_\oplus} - \gamma \frac{N_D^Y}{N_D^{XY}} D + E_D}^{\Delta} - \overbrace{\Theta q}^{\Phi} \ , \\
\dot{p} &= \mu(\Delta - \Sigma) p \ , \\
\dot{N}_{S\setminus D}^X &= \nu N_{S\setminus D}^Z - \Lambda_{\overline{tr}} P_{RA} N_{S\setminus D}^X \ , \\
\dot{N}_{S\setminus D}^Y &= \Lambda_{\overline{tr}} P_{RA} N_{S\setminus D}^X - \gamma N_{S\setminus D}^Y \ , \\
\dot{N}_{S\setminus D}^Z &= \gamma N_{S\setminus D}^Y - \nu N_{S\setminus D}^Z \ , \\
\dot{N}_{S\cap D}^X &= \nu N_{S\cap D}^Z - (\Lambda_{tr} + \Lambda_{\overline{tr}}) P_{RA} N_{S\cap D}^X \ , \\
\dot{N}_{S\cap D}^Y &= (\Lambda_{tr} + \Lambda_{\overline{tr}}) P_{RA} N_{S\cap D}^X - \gamma N_{S\cap D}^Y \ , \\
\dot{N}_{S\cap D}^Z &= \gamma N_{S\cap D}^Y - \nu N_{S\cap D}^Z \ , \\
\dot{N}_{D\setminus S}^X &= \nu N_{D\setminus S}^Z - (\Lambda_{tr} + \Lambda_{\overline{tr}}) P_{RA} N_{D\setminus S}^X \ , \\
\dot{N}_{D\setminus S}^Y &= (\Lambda_{tr} + \Lambda_{\overline{tr}}) P_{RA} N_{D\setminus S}^X - \gamma N_{D\setminus S}^Y \ , \\
\dot{N}_{D\setminus S}^Z &= \gamma N_{D\setminus S}^Y - \nu N_{D\setminus S}^Z \ .
\end{aligned}
\tag{E.64}
$$

Here we restrict ourselves to the case of symmetric imports, i.e. $E_S = E_D = E \geq 0$.

The forces of infection via trade routes (Λ_{tr}) and non-trade routes ($\Lambda_{\overline{tr}}$) are given by :

$$
\begin{aligned}
\Lambda_{tr} &= \overbrace{\underbrace{[1-(1-\phi)^q]}_{P_{tr}(q)} \frac{\Theta}{N_D^{XY}}}^{\beta_{tr}} \frac{N_S^Y}{N_S^{XY}} \ , \\
\Lambda_{\overline{tr}} &= \beta_{\overline{tr}} \frac{N^Y}{N^{XY}} \ .
\end{aligned}
\tag{E.65}
$$

The risk aversion factor $P_{RA} \in [0,1]$ is given by :

$$P_{RA} = \left(1 - \frac{N^Z}{N}\right)^{\alpha} \ . \tag{E.66}$$

Predicting the global behaviour of the ME model with the basic reproduction number R_0

R_0, the basic reproduction number, is the average number of susceptible that will be infected along the course of an epidemic by a single infectious agent propagated in an initially disease-free population [Anderson & May, 1991]. The global behavior of the coupled model depends on the position of R_0 with respect to 1. If $R_0 \leq 1$, the epidemic will eventually die out. If $R_0 > 1$ the epidemic will invade the population.

We derive R_0 using the next-generation matrix approach from Diekmann *et al.* [2010].

We take $S = \kappa N_S \frac{\Phi^* + E}{\Phi^*}$ and $D = \kappa N_D \frac{\Phi^* + E}{\Phi^*}$ as initial values for supply and demanded stocks, i.e. the steady-state equilibrium values of S and D in the FTM model when imports are not negligible (system (E.63)). It follows the market is equilibrated at steady-state before the epidemic onset, which implies, $\Theta = \Theta^* = \dfrac{\Phi^*}{\kappa}$, $q = \kappa \frac{\Phi^* + E}{\Phi^*}$ and $\Phi = \Phi^* + E$.

To calculate R_0, we focus on the 'disease-free equilibrium' (DFE) where a small number of agents are initially infected. Since the market is initially equilibrated at steady-state, we assume : $\Sigma \approx \Phi^* + E$ and $\Delta \approx \Phi^* + E$. Solving $\gamma \frac{N_S^Y}{N_S} S << \Phi^* + E$ and $\gamma \frac{N_D^Y}{N_D} D << \Phi^* + E$ with $S = \kappa N_S \frac{\Phi^* + E}{\Phi^*}$ and $D = \kappa N_D \frac{\Phi^* + E}{\Phi^*}$, we see that this approximation stands when $\max\{N_S^Y ; N_D^Y\} << \dfrac{\Phi^*}{\gamma\kappa}$. We now check if this assumption is in agreement with our livestock data. In practice, we have $\gamma_{\max} = 33$ per year and $\kappa_{\max} = 100$, so $\gamma_{\max}\kappa_{\max} \approx 10^3$ per year. $\Phi^* \approx 10^7$ per year (see Table **??** for reference values on κ and Φ^*). It follows $\dfrac{\Phi^*}{\gamma_{\max}\kappa_{\max}} \approx 10^4 \leq \dfrac{\Phi^*}{\gamma\kappa}$. Our assumption on the values of Σ and Δ at the DFE hence seems reasonable.

At the DFE, we notice from system (E.64) and equations (E.65)-(E.66) that :

$$\begin{aligned}
&\Lambda_{tr} \approx \overbrace{P_{tr}(\kappa \tfrac{\Phi^* + E}{\Phi^*}) \frac{\Phi^*}{\kappa N_D}}^{\beta_{tr}} \frac{N_S^Y}{N_S} \ , \\
&\Lambda_{\overline{tr}} \approx \beta_{\overline{tr}} \frac{N^Y}{N} \ , \\
&P_{RA} \approx 1 \ .
\end{aligned} \tag{E.67}$$

We can hence derive the infectious subsystem at the DFE :

$$\begin{aligned}
\dot{N}^Y_{S\setminus D} &= \beta_{\overline{tr}} \frac{N_{S\setminus D}}{N} N^Y - \gamma N^Y_{S\setminus D} \ , \\
\dot{N}^Y_{S\cap D} &= \beta_{tr} \frac{N_{S\cap D}}{N_S} N_S^Y + \beta_{\overline{tr}} \frac{N_{S\cap D}}{N} N^Y - \gamma N^Y_{S\cap D} \ , \\
\dot{N}^Y_{D\setminus S} &= \beta_{tr} \frac{N_{D\setminus S}}{N_S} N_S^Y + \beta_{\overline{tr}} \frac{N_{D\setminus S}}{N} N^Y - \gamma N^Y_{D\setminus S}.
\end{aligned} \tag{E.68}$$

We notice the right-hand side of system (E.68) is only a function of N_S^Y and N^Y. We can simplify system (E.68) by direct summation over suppliers ($N_S^Y = N_{S\setminus D}^Y + N_{S\cap D}^Y$) and agents ($N^Y = N_S^Y + N_{D\setminus S}^Y$) :

$$\begin{aligned}\dot{N}_S^Y &= \beta_{tr}\frac{N_{S\cap D}}{N_S}N_S^Y + \beta_{\overline{tr}}\frac{N_S}{N}N^Y - \gamma N_S^Y \ ,\\ \dot{N}^Y &= \beta_{tr}\frac{N_D}{N_S}N_S^Y + \beta_{\overline{tr}}N^Y - \gamma N^Y \ .\end{aligned} \tag{E.69}$$

We decompose the Jacobian matrix J from subsystem (E.69) into the sum $T + F$, where $T = \begin{pmatrix} \beta_{tr}\frac{N_{S\cap D}}{N_S} & \beta_{\overline{tr}}\frac{N_S}{N} \\ \beta_{tr}\frac{N_D}{N_S} & \beta_{\overline{tr}} \end{pmatrix}$ and $F = \begin{pmatrix} -\gamma & 0 \\ 0 & -\gamma \end{pmatrix}$ are the matrices accounting respectively for transmission and transitions from an epidemiological point-of-view.

We finally derive the next-generation matrix with large domain given by :

$$K_L = -TF^{-1} = \begin{pmatrix} \frac{\beta_{tr}}{\gamma}\frac{N_{S\cap D}}{N_S} & \frac{\beta_{\overline{tr}}}{\gamma}\frac{N_S}{N} \\ \frac{\beta_{tr}}{\gamma}\frac{N_D}{N_S} & \frac{\beta_{\overline{tr}}}{\gamma} \end{pmatrix} \tag{E.70}$$

R_0 is the leading eigenvalue of K_L. The two eigenvalues of K_L, denoted a, are obtained by solving the equation $\det(K_L - aI_2) = 0$, i.e. by solving the equation :

$$(\frac{\beta_{tr}}{\gamma}\frac{N_{S\cap D}}{N_S} - a)(\frac{\beta_{\overline{tr}}}{\gamma} - a) - (\frac{\beta_{tr}}{\gamma}\frac{N_D}{N_S})(\frac{\beta_{\overline{tr}}}{\gamma}\frac{N_S}{N}) = 0 \ . \tag{E.71}$$

Equation (E.71) is quadratic in a with a positive discriminant. It follows equation (E.71) has two real roots $a_1 \leq a_2$ with $R_0 = a_2$ given by :

$$\begin{aligned}R_0 &= \frac{R_0^{tr} + R_0^{\overline{tr}} + \sqrt{(R_0^{tr} - R_0^{\overline{tr}})^2 + 4R_0^{tr}R_0^{\overline{tr}}\frac{N_S N_D}{N N_{S\cap D}}}}{2}, \text{ with } R_0^{tr} \text{ and } R_0^{\overline{tr}} \text{ given by :}\\ R_0^{tr} &= \frac{\beta_{tr}}{\gamma}\frac{N_{S\cap D}}{N_S} \ ,\\ R_0^{\overline{tr}} &= \frac{\beta_{\overline{tr}}}{\gamma} \ .\end{aligned} \tag{E.72}$$

If we assume that all agents are wholesalers ($N_S = N_D = N_{S\cap D} = N$), R_0 simplifies to $R_0 = R_0^{tr} + R_0^{\overline{tr}}$.

Insights on the impacts of epidemics on market dynamics

Epidemics impact the market by decreasing the number of active suppliers and demanders and depleting the supply and demanded stocks. Assuming trade flow is equilibrated at the reference market level, it follows epidemics act by decreasing the net generation rates of supply and demanded stocks : $\Sigma \leq \Phi^*$ and $\Delta \leq \Phi^*$. The latter implies $\Phi \leq \Phi^*$, i.e. trade flow is decreased by epidemics.

The evolution of price is less intuitive. Generally speaking, an equilibrium price will satisfy $\Sigma(S, N_S^{XY}, p_{eq}) = \Delta(D, N_D^{XY}, p_{eq})$. We assume trade is the only path of transmission. Since trade-related transmission is directed from suppliers to demanders, demanders will be more impacted than suppliers. From the constraint $\Sigma = \Delta$, it directly follows that equilibrium prices will tend to decrease when submitted to epidemiological shocks (Fig. E.4). This is not necessarily true any more for measures such as quarantine or if other paths of transmission are involved.

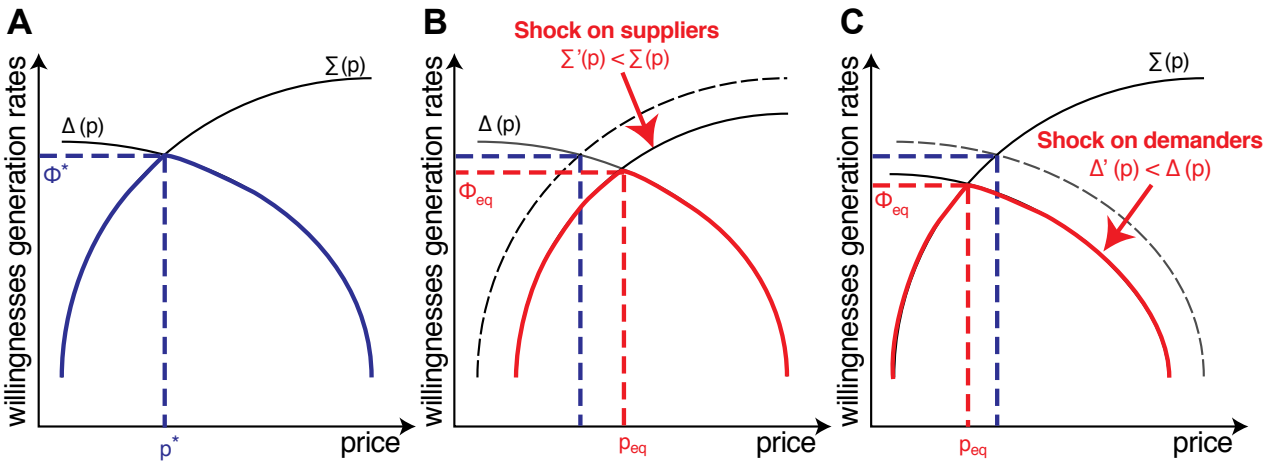


FIGURE E.4 – **Impact of shocks such as epidemics on trade flow and prices**

E.4.2 Numerical exploration of the ME model

Impacts of frictional-trade dynamics with risk aversion on disease dynamics

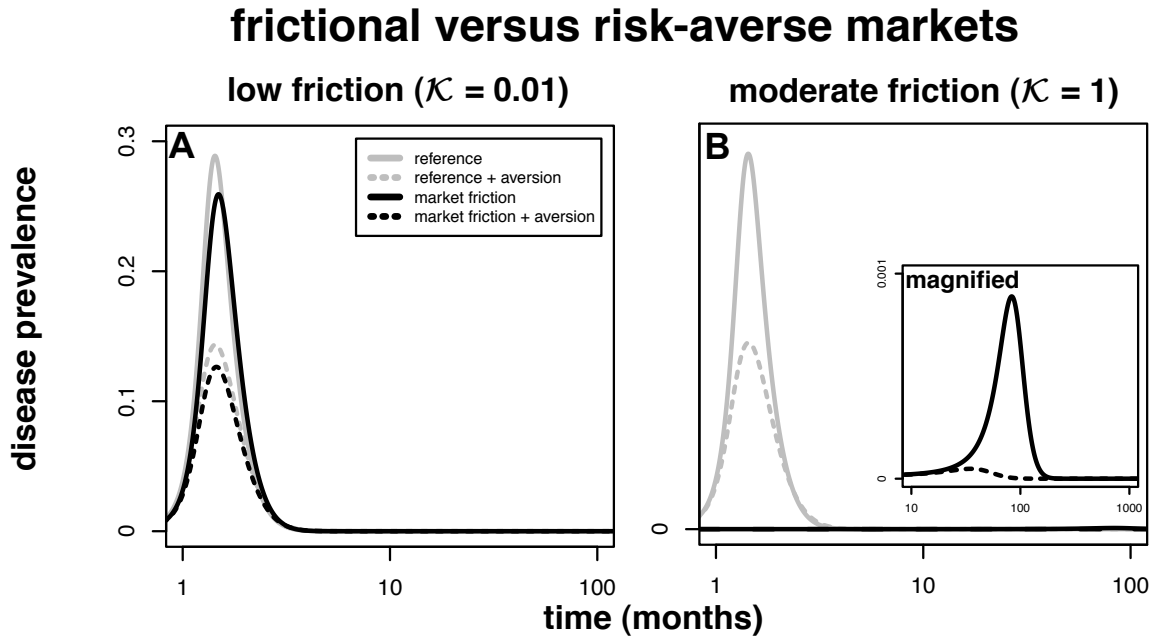


FIGURE E.5 – **Impacts of frictional-trade dynamics with risk aversion on disease dynamics (additional results).** Evolution of disease prevalence ($Y(t)/N$) as function of time with low (**A**) and moderate friction (**B**). Meaning of colours/lines and values of parameters are the same as in Figure 3A-B.

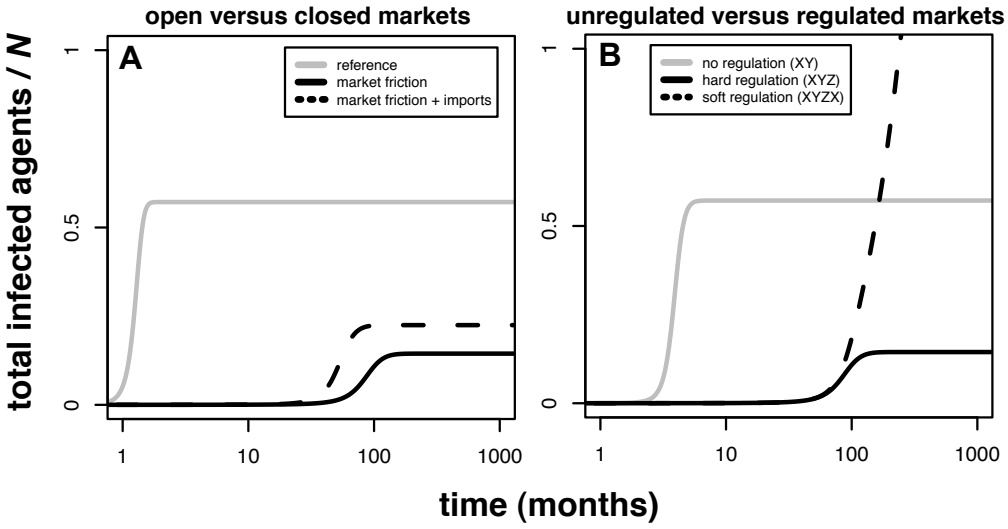


FIGURE E.6 – **Impacts of imports and different types of disease regulation on disease dynamics.** Evolution of total infected agents normalised by the number of agents as function of time with and without imports and with various types of regulation. *Impacts of open versus closed markets (**A**).* Regulation is hard ($1/\gamma$ = 11 days, $1/\nu \to \infty$ days), the level of friction is set to $\kappa = 1$ and the dashed line represents the impact of doubling trade flow by setting $E = \Phi^*$. Other lines are as in Figure 3B of the main text. *Impacts of unregulated versus regulated markets (**D**).* Regulation is inexistent ($1/\gamma \to \infty$ days ; in grey), soft ($1/\gamma$ = 11 days, $1/\nu$ = 365 days ; in dashed black) or hard ($1/\gamma$ = 11 days, $1/\nu \to \infty$ days ; in plain black). The level of friction is set to $\kappa = 1$ and the transmission rate to its frictional-market value. Other values of parameters are the same as in Figure 3A-B.

Impacts of epidemics on trade dynamics

The main text focuses on the impacts of frictional-trade dynamics on epidemics. Here we show an alternative point-of-view : the impact of epidemics on overall trade dynamics captured by the evolution of trade flow and price.

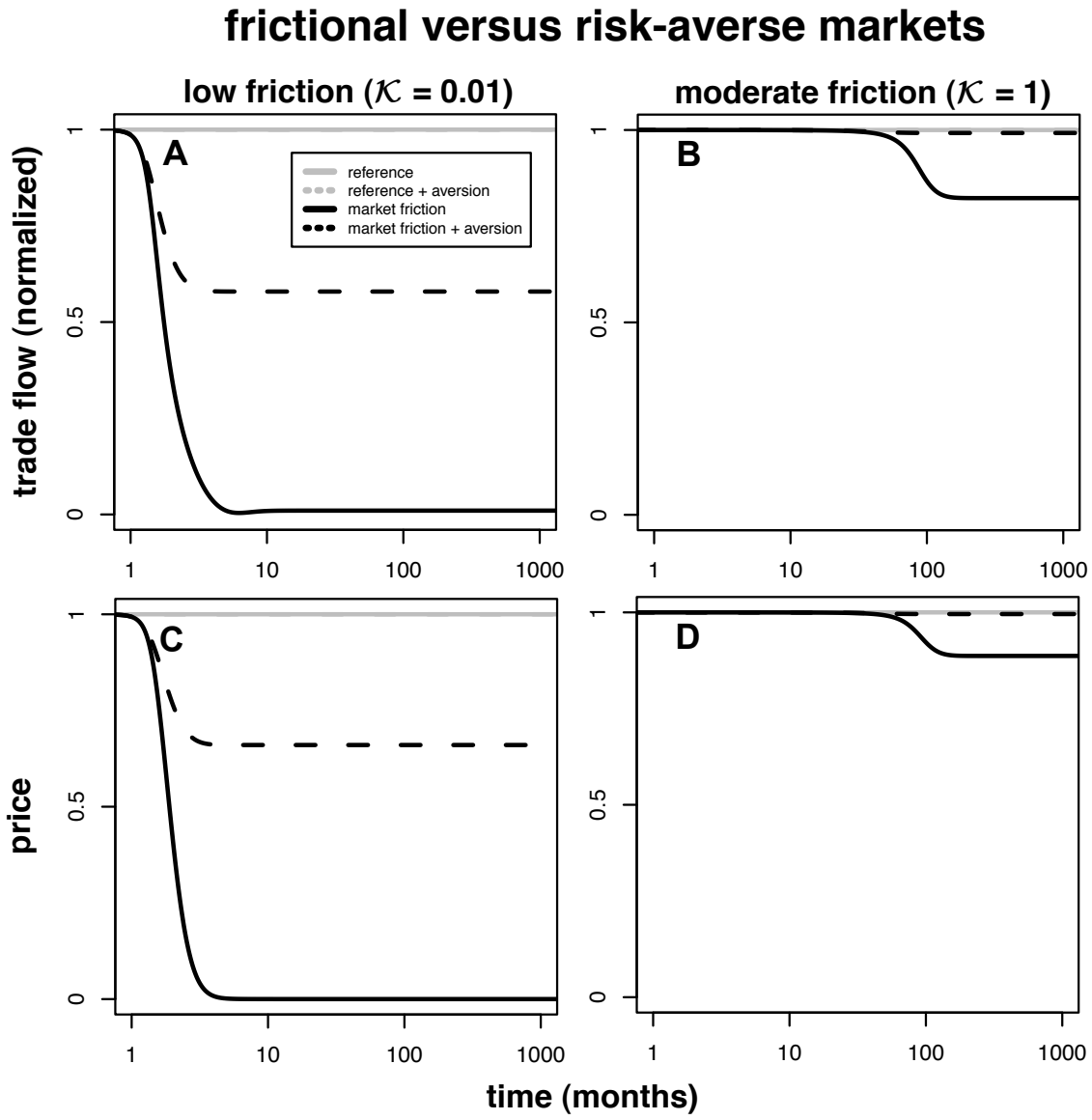


FIGURE E.7 – **Impacts of epidemics on trade dynamics** Evolution of normalised trade flow ($\frac{\Phi(t)}{\Phi^*}$, **A-B**) and price ($p(t)$, **C-D**) as function of time with low (**A-C**) and moderate friction (**B-D**). Meaning of colours/lines and values of parameters are the same as in Figure 3A-B of the main text.

GSA of the ME model

As for the FTM model, we assess the robustness of our conclusions on key economic and epidemiological outputs with the Morris method.

Inputs and outputs explored

Table E.2 describes the input parameters and associated ranges explored in the sensitivity analyses. Except when stated otherwise, notations and parameters are the same as described in the GSA of the FTM model.

TABLE E.2 – **Input parameters and associated ranges explored in the GSA of the ME model.**

Parameter	Formal notation	Meaning	Range explored	Distribution
k	κ	coefficient of friction	$[10^{-2}\,;\,10^{2}]$	log-uniform
i	$E_S = E_D = E$	supply and demand import rates (assumed equal)	$[0\,;\,\Phi^*]$ per year	uniform
pNsd	$p_{N_{S\cap D}}$	proportion of wholesalers	$[0.5\,;\,1]$	uniform
St0	$S(t_0)$	initial supply stock	$[0.8\kappa N_S\frac{\Phi^*+E}{\Phi^*}\,;\,1.2\kappa N_S\frac{\Phi^*+E}{\Phi^*}]$	uniform
Dt0	$D(t_0)$	initial demanded stock	$[0.8\kappa N_D\frac{\Phi^*+E}{\Phi^*}\,;\,1.2\kappa N_D\frac{\Phi^*+E}{\Phi^*}]$	uniform
pt0	$p(t_0)$	initial price	$[0.8\,;\,1.2]$	uniform
es	ε_S	price elasticity of supply	$[0\,;\,3]$	uniform
ed	ε_D	price elasticity of demand	$[0\,;\,3]$	uniform
mu	μ	pricing scale parameter	$[0\,;\,10^{-6}/\kappa]$	uniform
alpha	α	level of aversion to risk	$[0\,;\,8]$	uniform
gamma	γ	rate of detection and removal	$[0.5\,;\,33]$ per year	uniform
nu	ν	rate of market reentry	$[0\,;\,33]$ per year	uniform
phi	ϕ	probability of infection per good exchanged	$[0\,;\,\phi^{\max}]$	uniform
betabar	$\beta_{\overline{tr}}$	non-trade-related rate of transmission	$[0\,;\,\beta_{\overline{tr}}^{\max}]$ per year	uniform

To avoid unrealistic epidemics, we first set upper bounds on the trade-related (R_0^{tr}) on non-trade-related ($R_0^{\overline{tr}}$) components of R_0 : $R_0^{tr,\max} = 10$ and $R_0^{\overline{tr},\max} = 10$ respectively. Then we calculate the corresponding upper bounds for the probability of infection per good exchanged ($\phi^{\max}$) and the non-trade-related rate of transmission ($\beta_{\overline{tr}}^{\max}$) where $\phi^{\max} = 1-(1-P_{tr}^{\max})^{\frac{\Phi^*}{\kappa(\Phi^*+E)}}$ with $P_{tr}^{\max} = \min\{1;\frac{\gamma R_0^{tr,\max}\kappa N_S N_D}{\Phi^* N_{S\cap D}}\}$ and $\beta_{\overline{tr}}^{\max} = \gamma\min\{R_0^{\overline{tr},\max}, 2R_0^{tr}\}$.

The following economic and epidemiological outputs are analysed with the GSA :

- the endemic proportion of susceptible (at steady-state), $x_{eq} \in [0;1]$,
- the time to reach a real coupled at steady-state equilibrium (in supply, demand, price and epidemiological categories) or an apparent coupled at steady-state equilibrium (in trade flow, price and epidemiological categories), $t_{eq} \in [0;200]$ year,
- the ratio of steady-state equilibrium flow to disease-free flow at steady-state equilibrium, $\frac{\Phi_{eq}}{\Phi^*} \in \mathbb{R}^+$,
- the at steady-state equilibrium price, $p_{eq} \in \mathbb{R}^+$,
- the basic reproduction number, $R_0 \in \mathbb{R}^+$,
- the endemic proportion of infected (at steady-state), $y_{eq} \in [0;1]$,
- the proportion of infected at the peak, $y_{\max} \in [0;1]$,
- and the total number of cases at at steady-state equilibrium normalised by total agents, $[\text{total cases}]_{eq}/N \in \mathbb{R}^+$.

Results

The GSA confirms that the coefficient of friction κ is a key parameter governing trade and epidemiological dynamics. For all outputs explored, the effect of κ systematically outweighs the effect of adaptive risk aversion (RA) controlled by parameter α.

Sensitivity of x_{eq}

	m*	m	s
k	0.45	0.45	0.56
phi	0.42	-0.42	0.53
gamma	0.34	0.34	0.45
betabar	0.23	-0.23	0.31
alpha	0.15	0.15	0.25
nu	0.07	0.03	0.19
pNsd	0.06	-0.06	0.10
i	0.04	-0.04	0.10
pt0	0.01	-0.00	0.03
mu	0.01	-0.01	0.03
ed	0.00	-0.00	0.01
es	0.00	0.00	0.00
St0	0.00	0.00	0.00
Dt0	0.00	-0.00	0.00

Sensitivity of t_{eq}

	m*	m	s
k	65.44	57.95	97.86
mu	34.44	-0.54	79.03
gamma	24.55	-15.57	55.10
es	21.11	-2.34	55.29
ed	20.68	2.25	57.11
pt0	17.89	1.59	47.08
nu	14.40	-5.87	41.73
i	11.85	2.57	37.64
St0	11.56	0.99	38.74
pNsd	11.25	-2.12	35.31
betabar	10.70	1.75	30.25
phi	7.59	0.51	19.83
Dt0	7.47	-1.55	28.68
alpha	3.93	-2.13	20.82

Sensitivity of $\frac{\Phi_{eq}}{\Phi^*}$

	m*	m	s
i	0.94	0.94	0.13
k	0.35	0.20	0.56
phi	0.29	-0.29	0.42
gamma	0.23	0.02	0.42
nu	0.16	0.14	0.37
betabar	0.13	-0.13	0.24
alpha	0.12	0.12	0.22
mu	0.10	0.10	0.19
pt0	0.09	0.01	0.22
pNsd	0.05	-0.05	0.12
ed	0.05	-0.02	0.11
es	0.04	-0.03	0.11
St0	0.00	0.00	0.01
Dt0	0.00	0.00	0.01

Sensitivity of p_{eq}

	m*	m	s
pt0	0.13	0.13	0.17
k	0.11	0.09	0.33
mu	0.11	-0.04	0.23
phi	0.10	-0.08	0.25
nu	0.06	0.06	0.22
gamma	0.06	-0.01	0.18
ed	0.05	0.04	0.19
alpha	0.04	0.02	0.16
es	0.04	0.02	0.13
i	0.02	0.00	0.14
betabar	0.02	0.01	0.06
pNsd	0.01	0.01	0.05
St0	0.01	0.01	0.02
Dt0	0.00	-0.00	0.01

Sensitivity of R_0

	m*	m	s
k	7.13	-6.84	9.88
phi	6.92	6.92	8.37
gamma	5.35	-5.35	7.48
betabar	3.31	3.31	4.14
pNsd	0.90	0.51	1.38
i	0.40	0.40	0.76
St0	0.00	0.00	0.00
Dt0	0.00	0.00	0.00
pt0	0.00	0.00	0.00
es	0.00	0.00	0.00
ed	0.00	0.00	0.00
mu	0.00	0.00	0.00
alpha	0.00	0.00	0.00
nu	0.00	0.00	0.00

Sensitivity of y_{eq}

	m*	m	s
gamma	0.35	-0.35	0.51
k	0.22	-0.22	0.38
phi	0.21	0.21	0.39
nu	0.14	0.14	0.30
betabar	0.12	0.12	0.21
alpha	0.06	-0.06	0.10
pNsd	0.03	0.03	0.06
i	0.02	0.02	0.04
pt0	0.00	-0.00	0.02
mu	0.00	0.00	0.02
ed	0.00	-0.00	0.00
es	0.00	-0.00	0.00
Dt0	0.00	0.00	0.00
St0	0.00	-0.00	0.00

Sensitivity of y_{max}

	m*	m	s
gamma	0.35	-0.35	0.46
k	0.34	-0.34	0.44
phi	0.31	0.31	0.44
betabar	0.18	0.18	0.23
alpha	0.07	-0.07	0.11
nu	0.06	0.06	0.11
pNsd	0.05	0.05	0.08
i	0.03	0.03	0.06
pt0	0.01	0.00	0.02
mu	0.00	0.00	0.02
ed	0.00	-0.00	0.00
es	0.00	-0.00	0.00
Dt0	0.00	0.00	0.00
St0	0.00	-0.00	0.00

Sensitivity of $[\text{total cases}]_{eq}/N$

	m*	m	s
k	4.50	-2.15	8.26
phi	3.42	2.97	6.68
mu	3.26	-3.19	8.27
gamma	2.31	-1.13	5.60
betabar	2.13	1.01	5.70
nu	1.98	1.36	5.28
pt0	1.26	0.31	4.93
alpha	1.21	-0.63	3.21
i	0.90	0.60	3.70
es	0.85	-0.45	3.36
pNsd	0.83	0.33	2.62
St0	0.60	0.33	3.64
ed	0.55	-0.42	2.58
Dt0	0.37	0.32	1.91

Annexe F

Internship report : *Construction et implémentation d'un modèle d'épidémiologie économique fondé sur des réseaux dynamiques et adaptatifs*

This Master report was written by Sébastien Geeraert as part of an internship supervised by Elisabeta Vergu, Mathieu Moslonka-Lefebvre and Patrick Hoscheit.

École Polytechnique

Promotion 2011

Geeraert Sébastien

Rapport de stage de recherche

Construction et implémentation d'un modèle d'épidémiologie économique fondé sur des réseaux dynamiques et adaptatifs

NON CONFIDENTIEL

Option :	Mathématiques appliquées
Champ de l'option :	Modélisation probabiliste et statistique
Directeur de l'option :	Vincent Bansaye
Directeurs de stage :	Elisabeta Vergu, Mathieu Moslonka-Lefebvre
Co-encadrant :	Patrick Hoscheit
Dates du stage :	31/03/14–15/08/14
Nom et adresse de l'organisme :	INRA Unité MIA (UR341) Domaine de Vilvert 78352 Jouy-en-Josas, FRANCE

Abstract :

Commercial trade contributes to infectious diseases transmission. Conversely, epidemics can destabilize trade. However, interactions between trade and epidemics have not been extensively explored. In this study, we propose, analyse and numerically explore a mathematical model to describe commercial trade dynamics outside equilibrium. We first present an existing aggregated model of trade-based disease transmission, defined by ordinary differential equations and assuming the existence of representative suppliers and demanders. In order to take into account the inter-individual heterogeneity of economical agents involved in exchanges, we generalize the model at the individual scale, where the interactions lie on a directed, weighted and dynamical graph. We develop a deterministic model based on ordinary differential equations and a stochastic model based on markovian jump processes, and we analytically study the correspondence between the two models. We parametrize and numerically explore the model using a database of commercial cattle movements between French holdings. In particular, the data allow us to estimate key parameters of our model, the microscopic frictions, which delay exchanges without necessarily altering exchanged quantities over a given time period. Simulation results, in terms of study of model equilibrium and qualitative consistency with the data, are interpretated to suggest measures of pathogens propagation control. We show, using inhomogeneous percolation experiments, that frictions can be manipulated in order to increase the available critical time for control policies implementation. In practice, frictions can be increased by taxing transport and imposing a minimal size on animals shipments. This work will be continued on the one hand with a global sensibility analysis of the model to the parameters, particularly to frictions and to the functions defining the interactions network at the individual scale (in terms of neighbourhood and of links weights), and on the other hand with the integration of epidemical dynamics to validate the percolation results.

Résumé :

Les échanges commerciaux concourent à la transmission de maladies infectieuses. Réciproquement, les épidémies peuvent déstabiliser le commerce. Cependant, les interactions entre le commerce et les épidémies n'ont pas été explorées en profondeur. Dans cette étude, nous proposons, analysons et explorons numériquement un modèle mathématique pour décrire la dynamique en déséquilibre des échanges commerciaux. Nous présentons d'abord un modèle agrégé existant de transmission de maladie par le commerce, défini par des équations différentielles ordinaires, et supposant l'existence d'offreurs et demandeurs représentatifs. Afin de considérer l'hétérogénéité inter-individuelle des agents économiques responsables des échanges, nous généralisons le modèle à l'échelle des individus dont les interactions se déploient sur un graphe orienté, pondéré et dynamique. Nous développons un modèle déterministe s'appuyant sur des équations différentielles ordinaires ainsi qu'un modèle stochastique basé sur des processus markoviens de saut et nous étudions analytiquement la correspondance entre les deux modèles. Nous paramétrons et explorons numériquement le modèle sur la base des mouvements marchands de bovins entre exploitations en France. En particulier, les données nous permettent d'estimer des paramètres clés de notre modèle, les frictions microscopiques, qui retardent les échanges sans nécessairement modifier les quantités échangées sur une période de temps donnée. Les résultats des simulations, en termes d'étude des équilibres du modèle et d'adéquation qualitative aux données, sont interprétés pour suggérer des mesures de contrôle de la propagation d'agents pathogènes. Nous montrons par le biais d'expériences de percolation inhomogène que les frictions peuvent être manipulées afin d'augmenter le temps critique disponible pour la mise en œuvre des politiques de contrôle.

En pratique, les frictions peuvent êtres augmentées en taxant les transports ou en imposant une taille minimale sur les lots d'animaux transportables. Ce travail se poursuit d'une part avec une étude de sensibilité globale du modèle aux paramètres, tout particulièrement les frictions et les fonctions définissant le réseau d'interactions au niveau de l'individu (en termes de voisinage et de pondération des liens), et d'autre part avec l'intégration des dynamiques épidémiques pour valider les résultats des expériences de percolation.

Remerciements

Je tiens à remercier Elisabeta Vergu qui m'a accompagné et conseillé tout au long du stage. Merci également à Mathieu Moslonka-Lefebvre, pour son soutien et sa disponibilité. Je suis reconnaissant à Patrick Hoscheit de m'avoir apporté son aide sur les différents aspects mathématiques.

Je souhaite aussi exprimer ma gratitude envers Hervé Monod, directeur d'unité, qui m'a accueilli dans son laboratoire. Je remercie également tous les chercheurs du laboratoire de m'avoir intégré au sein de leur équipe.

F.1 Introduction

Le stage se déroule du 31 mars au 15 août 2014 au centre INRA de Jouy-en-Josas, au sein de l'Unité MIA 341 (Mathématiques et Informatique Appliquées). Il est dirigé par Elisabeta Vergu et Mathieu Moslonka-Lefebvre. Le travail s'appuie sur la thèse de ce dernier, intitulée « Vers une épidémiologie des réseaux d'échanges à contraintes : regards croisés issus d'analyses empiriques et de modèles mathématiques. Applications aux maladies infectieuses propagées sur des marchés agricoles et des réseaux de contacts sexuels ». Ce stage est également co-encadré par Patrick Hoscheit pour les aspects mathématiques, notamment liés à la modélisation probabiliste. Par ailleurs, ce travail s'inscrit dans le cadre du projet [Epi-Eco] (Épidémiologie Économique) soutenu par l'Unité MIA-Jouy et le Département MIA de l'INRA. Il bénéficie d'apports de João Filipe (University of Cambridge), de Christopher Gilligan (University of Cambridge) et d'Hervé Monod (Unité MIA) pour la conception du modèle économique microscopique.

L'Unité MIA comprend trois équipes : AnIMod (analyse d'images et modélisation spatio-temporelle en biologie cellulaire), DynEnVie (modélisation des phénomènes dynamiques rencontrés en sciences de la vie et de l'environnement, équipe à laquelle j'ai été rattaché) et MegaDim (exploitation de très grands jeux de données et de modèles complexes développés en biologie). Elle regroupe des compétences diverses de mathématiques appliquées, d'informatique et de biologie. Régulièrement, des séminaires y ont lieu où les chercheurs présentent leurs travaux récents.

F.1.1 Contexte général et objectifs

Des maladies infectieuses peuvent se transmettre sur de nombreux marchés au cours d'échanges commerciaux. C'est par exemple le cas pour les échanges de bovins : une ferme indemne de toute maladie peut voir ses animaux contaminés à la suite d'une transaction avec une ferme dont les animaux sont infectés par un agent pathogène donné, ce qui contribue à la propagation de la maladie. Les épidémies sur des réseaux commerciaux ont été étudiées dans la littérature, mais les structures d'échange associées sont souvent statiques [Rautureau *et al.*, 2011]. Quelques travaux s'appuient sur des réseaux dynamiques, mais l'évolution du réseau commercial à cause de l'épidémie n'y est pas prise en compte. Or, il semble naturel de penser que l'épidémie affecte le comportement des agents et les échanges commerciaux qui en résultent : les agents s'adaptent au risque de contamination et à la disparition de certains de leurs partenaires économiques. La prise en compte de la dynamique et de l'adaptativité du réseau commercial sur lequel se propage l'épidémie, ainsi que des boucles de rétroaction, est l'objet des travaux de Mathieu Moslonka-Lefebvre [Moslonka-Lefebvre *et al.*, 2012b, 2013]. L'étude de l'effet rétroactif de l'épidémie sur les échanges commerciaux se rattache aux problématiques de l'épidémiologie économique [Klein *et al.*, 2007]. Cette discipline vise à développer des modèles épidémiologiques intégrant des processus économiques tels que l'utilisation optimale de budgets de gestion limités. Ces

modèles couplés, souvent plus réalistes que les modèles épidémiologiques isolés, permettent de proposer des politiques publiques efficaces de gestion des épidémies.

Dans [Moslonka-Lefebvre *et al.*, 2013], Mathieu Moslonka-Lefebvre et al. développent un cadre conceptuel pour modéliser conjointement la dynamique économique et la dynamique épidémiologique, grâce à des équations différentielles ordinaires (EDO). Cependant, ce modèle s'applique à un réseau où les agents sont agrégés. Ils sont divisés en trois catégories. Au sein d'une catégorie, les agents sont supposés parfaitement homogènes : ils ont tous la même production et les mêmes volontés d'échanges. Or, il est important de tenir compte de l'hétérogénéité entre agents : chacun d'entre eux a des préférences propres et un comportement particulier. Il est facilement concevable que l'hétérogénéité ait des effets différents sur la propagation de l'épidémie.

L'objectif du stage est de généraliser ce modèle macroscopique à l'échelle microscopique : les agents sont considérés par petites catégories ou au niveau individuel ; leurs interactions s'appuient sur un réseau orienté, pondéré et dynamique dans le temps. Cela permet de rendre compte à la fois de l'hétérogénéité des agents et des effets de rétroaction possibles entre dynamique économique et dynamique épidémiologique.

F.1.2 Contexte de modélisation économique

Le marché est défini comme une institution (que l'on assimile à un réseau) où des biens s'échangent entre agents économiques. Un agent décide d'opérer des transactions en fonction de ses intérêts particuliers. Les économistes développent essentiellement des modèles à équilibration instantanée ou proches de l'équilibre [May *et al.*, 2008; McCauley, 2009]. Or, puisque notre objectif est d'étudier la dynamique des échanges commerciaux, la phase transitoire ne doit pas être négligée. On s'inspire donc de différents modèles pour construire le nôtre.

Le modèle d'équilibre partiel [Mas-Colell *et al.*, 1995] introduit l'offre Q_S (ce que les agents sont prêts à vendre) et la demande Q_D (ce que les agents sont prêts à acheter). Un agent module son offre Q_S et sa demande Q_D en fonction du prix p : par exemple, $Q_S(t) = Q_{S_0} p(t)^{\varepsilon}$ et $Q_D(t) = Q_{D_0} p(t)^{-\eta}$ où ε et η sont des élasticités positives, et où Q_{S_0} et Q_{D_0} sont l'offre et la demande de référence. Si le prix augmente, l'offre augmente et la demande diminue, ce qui est intuitif.

L'évolution du prix s'appuie sur le modèle de Walras-Samuelson [Mas-Colell *et al.*, 1995] : le prix évolue en fonction du déséquilibre $Q_D - Q_S$ entre l'offre et la demande. Une offre élevée par rapport à la demande donnera un prix faible et inversement, une offre faible par rapport à la demande donnera un prix élevé.

La quantité de biens échangés s'inspire du modèle de Fair et Jaffe [Fair & Jaffee, 1972]. Ces derniers supposent que, lors d'une transaction, les agents échangent le maximum possible permis par l'offre et la demande, c'est-à-dire $\min(Q_S, Q_D)$.

Enfin, le modèle de Diamond, Mortensen et Pissarides [Economic Sciences Prize Committee, 2010]

introduit, dans le contexte de l'économie du travail, le paramètre de friction κ. Ce paramètre affecte négativement le taux d'appariement entre offreurs et demandeurs d'emplois. Il agrège les contraintes de recherche d'emploi pour les deux côtés du marché : plus la friction est grande, plus les appariements sont rares. Nous adaptons le concept de friction aux échanges marchands.

F.1.3 Contexte de modélisation épidémiologique

Pour la modélisation épidémiologique, on utilise un modèle classique de type SIRS (Susceptible-Infectious-Removed-Susceptible) présenté dans [Keeling & Rohani, 2008]. Les N agents sont répartis en trois classes : les susceptibles S (sujets sains qui peuvent attraper la maladie), les infectieux I (sujets infectés qui peuvent transmettre la maladie) et les retirés R (sujets ayant été infectés mais qui ne peuvent plus transmettre la maladie). Les agents passent de S à I au taux $\beta\frac{I}{N}$ (infection), de I à R au taux γ (détection) et de R à S au taux ν (perte d'immunité). Les équations régissant l'évolution des effectifs S, I, R sont alors les suivantes :

$$\begin{aligned} \dot{S} &= -\frac{\beta SI}{N} + \nu R \\ \dot{I} &= \frac{\beta SI}{N} - \gamma I \\ \dot{R} &= \gamma I - \nu R. \end{aligned} \tag{F.1}$$

Le terme de transmission est en $\frac{\beta SI}{N}$ car, pour chaque agent dans S qui rencontre un autre agent, la probabilité que l'agent rencontré soit infectieux est $\frac{I}{N}$.

On peut aussi construire un modèle stochastique qui fonctionne selon les mêmes mécanismes, en utilisant des processus markoviens de sauts à durées de séjour exponentielles. Chaque agent peut passer de S à I, de I à R et de R à S. Les taux associés aux transitions de ce processus markovien sont les mêmes que pour le modèle déterministe : $\beta\frac{I}{N}$ pour l'infection, γ pour la détection et ν pour la perte d'immunité.

F.1.4 Démarche scientifique et structuration du rapport

Nous présentons tout d'abord le modèle agrégé d'épidémiologie économique existant, construit à partir des mécanismes de transaction commerciale et de transmission d'agents pathogènes. Puis nous présentons la généralisation à l'échelle microscopique de ce modèle : nous construisons à la fois un modèle déterministe par EDO et un modèle stochastique s'appuyant sur des processus markoviens de saut. Ensuite, nous testons en pratique la partie économique des modèles, en nous appuyant sur une base de données de mouvements de bovins. Nous confrontons les modèles stochastiques et déterministes entre eux et avec les données. Finalement, nous faisons des expériences de percolation pour estimer la propagation de l'épidémie sur le réseau et en déduire des politiques publiques possibles pour la contrer.

Les principales contributions personnelles que j'ai apportées au sujet sont les suivantes :

– analyse détaillée et conjointe des modèles microscopiques (stochastique et EDO) ;
– implémentation optimale des modèles ;
– exploration numérique préliminaire sur la base de données d'échanges commerciaux, ce qui permet de mieux interpréter les modèles et donne des perspectives d'amélioration du scénario de référence pour être en meilleure adéquation avec les données ;
– réalisation d'expériences de percolation inhomogène.

F.2 Modélisation mécaniste des dynamiques marchandes et épidémiques

Dans cette partie, on s'appuie sur des mécanismes marchands et épidémiques pour construire des modèles d'épidémiologie économique. Tout d'abord, on présente le modèle agrégé existant où les agents sont considérés dans leur ensemble. On le généralise ensuite à l'échelle microscopique où les agents sont hétérogènes : on met au point un modèle EDO et un modèle stochastique.

F.2.1 Modèle macroscopique d'épidémiologie économique, formalisme par EDO

On présente ici le modèle macroscopique développé par Mathieu Moslonka-Lefebvre et al. dans [Moslonka-Lefebvre *et al.*, 2013]. Pour la modélisation économique, on se place dans le cadre d'un marché d'échange avec friction, où tous les biens échangés sont identiques et se vendent au même prix $p(t)$. Le modèle est défini par un stock global d'offre $S(t)$ (*supply*) et de demande $D(t)$ (*demand*). Le stock d'offre représente les biens détenus par les vendeurs et disponibles à l'achat, tandis que le stock de demande est virtuel et représente ce que les acheteurs désirent acheter. Au cours du temps, l'évolution du stock est affectée positivement par la création de biens et négativement par les flux échangés. On peut donc écrire

$$\begin{aligned} \dot{S}(t) &= \Sigma(t) - \Phi(t) \\ \dot{D}(t) &= \Delta(t) - \Phi(t) \end{aligned} \tag{F.2}$$

où Σ, Δ sont les taux nets de création et où Φ est le flux échangé instantané. Les taux de création sont donnés par

$$\Sigma(t) = N_S \sigma_0 p(t)^{\varepsilon} \quad ; \quad \Delta(t) = N_D \delta_0 p(t)^{-\eta}. \tag{F.3}$$

Les variables N_S et N_D sont le nombre d'agents respectivement offreurs et demandeurs. Les paramètres σ_0 et δ_0 sont les taux de référence de production de stock d'offre et de demande par agent. Les paramètres ε et η sont des élasticités positives : un prix élevé encourage la production d'offre et décourage la production de demande. Pour des raisons de simplicité, on omet de rajouter dans le taux de création des termes de pertes (dues au vieillissement du stock par exemple) et d'apports extérieurs (dus à l'import ou l'export).

Le flux échangé instantané s'écrit

$$\Phi(t) = \Theta(t) q(t) \tag{F.4}$$

où Θ est le taux de transaction et où q est la quantité moyenne échangée. Les stocks moyens d'offre et de demande par agent sont $S(t)/N_S$ et $D(t)/N_D$. On considère que lorsque deux agents ont décidé

d'effectuer une transaction, ils échangent la quantité maximale possible :

$$q(t) = \min\left(\frac{S(t)}{N_S}, \frac{D(t)}{N_D}\right). \tag{F.5}$$

Le taux de transaction Θ est déterminé par

$$\Theta(t) = \frac{1}{\kappa}\min\left(\Sigma(t), \Delta(t)\right). \tag{F.6}$$

Le terme en min représente la production maximale d'offre et de demande qui peuvent se satisfaire mutuellement (on s'attend à ce que ce soit le taux maximal possible d'échange entre les offreurs et les demandeurs), tandis que la friction κ est un facteur limitant qui tient compte de la recherche d'un partenaire commercial et de la logistique associée à la livraison (par exemple, si on ne peut effectuer que des livraisons de plus de deux biens, on s'attend à ce que $\kappa \geq 2$).

Enfin, la dynamique de prix est déterminée par le déséquilibre entre l'offre et la demande :

$$\dot{p}(t) = \mu(\dot{D}(t) - \dot{S}(t))p(t) \tag{F.7}$$

où μ est le paramètre d'ajustement du prix. On peut exprimer p en fonction de S et D, ce qui donne le système suivant pour le modèle économique :

$$\begin{aligned}
\dot{S}(t) &= \Sigma(t) - \Theta(t)q(t) \\
\dot{D}(t) &= \Delta(t) - \Theta(t)q(t) \\
p(t) &= p(0)\exp\left(\mu(D(t) - S(t) - D(0) + S(0))\right).
\end{aligned} \tag{F.8}$$

On complète ce modèle économique par un modèle épidémiologique de type SIRS [Keeling & Rohani, 2008]. Le nombre total d'agents N se répartit en $N = N^X + N^Y + N^Z$ où X désigne les agents susceptibles, Y les agents infectieux et Z les agents détectés et retirés (mis hors d'état de nuire). Les agents commercialement actifs sont les agents X et Y. On les note avec un exposant XY. On note par un indice S les offreurs (stricts ou non) et par un indice D les demandeurs (stricts ou non). Par extension, $S\backslash D$, $D\backslash S$ et $S \cap D$ désignent respectivement les offreurs stricts, les demandeurs stricts et les offreurs-demandeurs. Par exemple, $N^{XY}_{S\backslash D}$ représente le nombre d'offreurs stricts susceptibles ou infectieux.

L'épidémie peut se transmettre par des voies commerciales (notées tr) et par des voies non commerciales (notées $\overline{\text{tr}}$). La transmission commerciale a lieu dans le sens de la transaction : l'acheteur qui est dans N^X_D peut être contaminé par le vendeur qui est dans N^Y_S. La transmission non commerciale a lieu d'un agent de N^Y à un agent de N^X. Ainsi, le taux Λ d'infection des demandeurs peut s'écrire :

$$\Lambda(t) = \left(\Lambda_{\text{tr}}(t) + \Lambda_{\overline{\text{tr}}}(t)\right) P_{RA}(t) \tag{F.9}$$

où P_{RA} est un facteur d'aversion au risque : plus le nombre de cas détectés est important (ie plus N_Z est grand), plus les agents vont prendre de précautions lors des transactions pour éviter une transmission. On choisit, en accord avec la littérature [Funk *et al.*, 2010],

$$P_{RA}(t) = \left(1 - \frac{N_Z(t)}{N}\right)^{\alpha}. \tag{F.10}$$

Plus α est élevé, plus l'aversion au risque est grande. De plus, l'aversion au risque augmente quand des agents infectieux sont détectés (c'est-à-dire quand $\frac{N_Z(t)}{N}$ augmente).

Le taux $\Lambda_{\text{tr}}(t)$ peut s'exprimer sous la forme $\Lambda_{\text{tr}}(t) = \beta_{\text{tr}}(t)N_S^Y(t)/N_S^{XY}(t)$: $\beta_{\text{tr}}(t)$ est le taux de transmission commerciale sachant que le vendeur est infecté et $N_S^Y(t)/N_S^{XY}(t)$ est la probabilité que le vendeur avec qui l'on échange soit infecté. On peut écrire $\beta_{\text{tr}}(t) = P_{\text{tr}}(q(t))\Theta(t)/N_D^{XY}(t)$ où $\Theta(t)/N_D^{XY}(t)$ est le taux de transaction par demandeur actif et où $P_{\text{tr}}(q(t))$ est la probabilité de devenir infecté si l'on reçoit $q(t)$ unités d'un vendeur infecté. En supposant que chacune des unités échangées a une probabilité d'infection indépendante ϕ, on a $P_{\text{tr}}(q) = 1 - (1-\phi)^q$.

De même, le taux $\Lambda_{\overline{\text{tr}}}$ peut s'écrire $\Lambda_{\overline{\text{tr}}}(t) = \beta_{\overline{\text{tr}}}(t)N^Y(t)/N^{XY}(t)$ où $\beta_{\overline{\text{tr}}}(t)$ est le taux de transmission non-commerciale, que l'on ne spécifie pas.

On peut alors écrire le modèle complet, en tenant compte du fait que les agents infectieux sont retirés à un taux γ et réintroduits au taux ν :

$$\begin{aligned}
\dot{S}(t) &= N_S^{XY}(t)\sigma_0 p(t)^{\varepsilon} - \gamma\frac{N_S^Y(t)}{N_S^{XY}(t)}S(t) - \Theta(t)q(t)\\
\dot{D}(t) &= N_D^{XY}(t)\delta_0 p(t)^{-\eta} - \gamma\frac{N_D^Y(t)}{N_D^{XY}(t)}D(t) - \Theta(t)q(t)\\
\dot{N}^X_{S\cap D}(t) &= \nu N^Z_{S\cap D}(t) - \Lambda(t)N^X_{S\cap D}(t)\\
\dot{N}^Y_{S\cap D}(t) &= \Lambda(t)N^X_{S\cap D}(t) - \gamma N^Y_{S\cap D}(t)\\
\dot{N}^Z_{S\cap D}(t) &= \gamma N^Y_{S\cap D}(t) - \nu N^Z_{S\cap D}(t)
\end{aligned} \tag{F.11}$$

où $q(t)$ et $\Theta(t)$ sont définis comme dans (F.5) et (F.6) avec $N_S^{XY}(t)$ et $N_D^{XY}(t)$ au lieu de N_S et N_D. On a des équations semblables pour $\dot{N}^X_{D\setminus S}, \dot{N}^X_{S\setminus D}, \ldots$

Le terme $-\gamma\frac{N_S^Y(t)}{N_S^{XY}(t)}S(t)$ dans l'expression de $\dot{S}(t)$ est dû au fait que le retrait des agents infectieux diminue les stocks disponibles : par unité de temps, on retire $\gamma N_S^Y(t)$ offreurs infectieux, et chacun de ces offreurs dispose d'un stock d'offre moyen $\frac{S(t)}{N_S^{XY}(t)}$.

Le comportement de ce modèle est étudié en détail dans l'annexe électronique de [Moslonka-Lefebvre *et al.*, 2013]. En particulier, c'est un problème bien posé : quelles que soient les conditions initiales, les solutions existent et sont uniques. Les solutions sont de plus positives si les conditions initiales sont positives. Le flux $\Phi(t)$ et le prix $p(t)$ convergent toujours vers des valeurs de référence Φ^* et p^*. Par contre, les stocks $S(t)$ et $D(t)$ possèdent une infinité d'équilibres instables.

F.2.2 Modèle microscopique EDO

Dans le modèle macroscopique décrit, on a considéré que les agents étaient répartis en trois classes : les offreurs stricts ($S\backslash D$), les demandeurs stricts ($D\backslash S$) et les autres ($S \cap D$). Au sein de ces classes, le comportement était supposé identique. Nous allons maintenant généraliser le modèle en tenant compte de l'hétérogénéité des agents. Cela permettra de modéliser plus finement les échanges et les interactions.

On considère le cas d'un modèle hétérogène où les agents sont regroupés par catégories : on suppose qu'il y a N_i agents au comportement identique dans la catégorie C_i, caractérisés par σ_0^i, la production d'offre de référence par agent, et par δ_0^i, la production de demande de référence par agent. On fait l'hypothèse que les agents n'échangent pas entre eux au sein d'une catégorie. On écrit $N_i = N_i^X + N_i^Y + N_i^Z$ où X désigne les agents susceptibles, Y les agents infectieux et Z les agents retirés. On suppose que les N_i sont assez grands pour que l'on puisse considérer des effectifs continus. Pour simplifier les notations, on omet dans les expressions les dépendances en temps. Sauf mention contraire, les sommes portent sur l'ensemble des nœuds du réseau.

Par analogie avec le modèle macroscopique, l'évolution du marché avec friction est donnée par :

$$\begin{aligned} \dot{S}_i &= N_i^{XY}\sigma_i - \sum_j \theta_{ij} q_{ij} - \gamma \frac{N_i^Y}{N_i^{XY}} S_i \\ \dot{D}_i &= N_i^{XY}\delta_i - \sum_j \theta_{ji} q_{ji} - \gamma \frac{N_i^Y}{N_i^{XY}} D_i \end{aligned} \tag{F.12}$$

pour tout i, où S_i représente le stock d'offre de C_i et D_i le stock de demande. Les productions d'offre et de demande sont

$$\sigma_i = \sigma_0^i p^{\varepsilon} \quad ; \quad \delta_i = \delta_0^i p^{-\eta} \tag{F.13}$$

où le prix est donné par :

$$p = p(0) \exp\left(\mu \sum_k (D_k - S_k) - \mu \sum_k (D_k(0) - S_k(0)) \right). \tag{F.14}$$

La transaction moyenne entre i et j est

$$q_{ij} = \min\left(\frac{S_i}{N_i^{XY}}, \frac{D_j}{N_j^{XY}} \right). \tag{F.15}$$

Comment définir le taux de transaction θ_{ij} ? Un agent de la catégorie C_i va répartir d'une certaine façon sa production d'offre σ_i et sa production de demande δ_i entre les différentes catégories d'agents. On introduit donc les coefficients de fractionnement f_{ij} et g_{ij}. Le coefficient f_{ij} dénote la fraction de la production d'offre σ_i d'un agent de C_i que ce dernier souhaite vendre à la catégorie C_j. De même, g_{ij} dénote la fraction de la production de demande δ_j d'un agent de C_j que ce dernier souhaite acheter à

la catégorie C_i. En particulier, on a $\sum_j f_{ij} = 1$ pour tout i et $\sum_i g_{ij} = 1$ pour tout j. Les coefficients de fractionnement sont des paramètres très importants du modèle : ce sont eux qui définissent la structure du réseau. Ils déterminent par exemple si deux nœuds i et j sont voisins dans le réseau (suivant que f_{ij} et g_{ij} sont nuls ou non).

Bien sûr, les agents économiques peuvent adapter leurs efforts à la progression de l'épidémie, en choisissant par exemple f_{ij} en fonction de la proportion d'agents encore actifs dans la catégorie C_j. Dans ce cas,

$$f_{ij} = \frac{\frac{N_j^{XY}}{N_j} f_{ij}^{(0)}}{\sum_k \frac{N_k^{XY}}{N_k} f_{ik}^{(0)}} \tag{F.16}$$

et symétriquement

$$g_{ij} = \frac{\frac{N_i^{XY}}{N_i} g_{ij}^{(0)}}{\sum_k \frac{N_k^{XY}}{N_k} g_{kj}^{(0)}}. \tag{F.17}$$

où $f_{ij}^{(0)}$ et $g_{ij}^{(0)}$ sont les valeurs de référence en absence d'épidémie. Des exemples de coefficients de partitionnement sont donnés en parties F.3.2 et F.3.4.

Le taux de transaction θ_{ij} entre les catégories i et j est ensuite le taux maximal permis par les deux agents, modulé par une friction κ_{ij}, c'est-à-dire

$$\theta_{ij} = \frac{1}{\kappa_{ij}} \min\left(N_i^{XY} f_{ij} \sigma_i, N_j^{XY} g_{ij} \delta_j\right). \tag{F.18}$$

L'évolution de l'épidémie suit les équations suivantes :

$$\begin{cases} \dot{N_i^X} = -(\Lambda_{\overline{\text{tr}}} + \Lambda_i) P_{RA}^{(i)} N_i^X + \nu N_i^Z \\ \dot{N_i^Y} = (\Lambda_{\overline{\text{tr}}} + \Lambda_i) P_{RA}^{(i)} N_i^X - \gamma N_i^Y \\ \dot{N_i^Z} = \gamma N_i^Y - \nu N_i^Z \end{cases} \tag{F.19}$$

où γ est le taux de détection et ν le taux de réintroduction. $P_{RA}^{(i)}$ est le facteur d'aversion au risque de i, défini de façon analogue à (F.10) par

$$P_{RA}^{(i)} = \left(1 - \frac{N^Z}{N}\right)^{\alpha_i} \tag{F.20}$$

où α_i caractérise l'aversion au risque de l'agent, qui dépend de sa catégorie. L'état d'avancement de l'épidémie observable par un agent, donné par $\frac{N^Z}{N}$, ne dépend pas de la catégorie à laquelle appartient l'agent.

Le taux $\Lambda_{\overline{\text{tr}}}$ d'infection non-commerciale s'écrit

$$\Lambda_{\overline{\text{tr}}} = \beta_{\overline{\text{tr}}} \frac{N^Y}{N^{XY}} \tag{F.21}$$

et le taux Λ_i d'infection due aux échanges pour i est donné par

$$\Lambda_i = \sum_j P_{\text{tr}}(q_{ji}) \frac{\theta_{ji}}{N_i^{XY}} \frac{N_j^Y}{N_j^{XY}} \tag{F.22}$$

où $P_{\text{tr}}(q) = 1 - (1-\phi)^q$ est défini de la même façon que pour le modèle macroscopique.

F.2.3 Modèle microscopique stochastique et algorithme de simulation

Les mécanismes décrits dans le modèle précédent par des EDO peuvent aussi être formalisés dans un cadre stochastique. Les événements aléatoires tels que les transactions, les retraits et les réintroductions sont modélisés comme des processus markoviens de saut (avec des durées de séjour exponentielles). La production de stock a lieu continûment. On travaille ici au niveau individuel : chaque agent est considéré indépendamment (les agents ne sont plus regroupés au sein de catégories : $N_i = 1$ pour tout i). On répartit les agents en trois classes X, Y, Z. On modélise chaque agent séparément. L'appartenance d'un agent i à une certaine classe est décrite par une variable binaire : par exemple, $X_i = 1$ si i est susceptible et $X_i = 0$ sinon (cette représentation est dictée par l'implémentation). Ainsi, pour tout i, on a un espace d'états de la forme $(S_i, D_i, X_i, Y_i, Z_i) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \{0,1\}^3$ où $X_i + Y_i + Z_i = 1$. Comme précédemment, un agent i est caractérisé par ses productions de stock σ_0^i et δ_0^i, et ses préférences $f_{ij}^{(0)}$ et $g_{ij}^{(0)}$. Il possède un stock d'offre S_i et un stock de demande D_i.

Les équations (F.13) restent valables mais dans le calcul du prix, il ne faut tenir compte que des stocks des agents encore actifs (un agent est considéré actif s'il appartient à X ou Y, c'est-à-dire s'il n'est pas retiré) :

$$p(t) = p(0) \exp\left(\mu \sum_{k \in X \cup Y} (D_k(t) - S_k(t)) - \mu \sum_{k \in X(0) \cup Y(0)} (D_k(0) - S_k(0)) \right). \tag{F.23}$$

Les transactions commerciales entre i et j ont lieu au taux θ_{ij} où

$$\theta_{ij}(t) = \frac{1}{\kappa_{ij}} \min\left(f_{ij}\sigma_i, g_{ij}\delta_j\right) \tag{F.24}$$

à condition que i et j soient actifs. Si une transaction a lieu, on considère que la quantité échangée est la plus grande possible :

$$q_{ij}(t) = \min\left(S_i, D_j\right). \tag{F.25}$$

Là encore, les agents ne consacrent leurs efforts qu'à leurs voisins actifs :

$$f_{ij}(t) = \frac{\mathbb{1}_{j\in X\cup Y} f_{ij}^{(0)}}{\sum_k \mathbb{1}_{k\in X\cup Y} f_{ik}^{(0)}} \quad \text{et} \quad g_{ij}(t) = \frac{\mathbb{1}_{i\in X\cup Y} g_{ij}^{(0)}}{\sum_k \mathbb{1}_{k\in X\cup Y} g_{kj}^{(0)}}. \tag{F.26}$$

Finalement, on a donc les événements aléatoires suivants :

Événement	Agents	Taux	Transitions
Transaction (transmission potentielle)	$i \in X \cup Y$ $j \in X \cup Y$	$\theta_{ij}(t)$	$S_i \leftarrow S_i - q_{ij}$, $D_j \leftarrow D_j - q_{ij}$ et avec probabilité $\mathbb{1}_{j\in X, i\in Y} P_{\mathrm{tr}}(q_{ij}) P_{RA}^{(j)}$: $X_j \leftarrow 0$, $Y_j \leftarrow 1$
Transmission $\overline{\mathrm{tr}}$	$i \in X$	$\Lambda_{\overline{\mathrm{tr}}}(t) P_{RA}^{(i)}(t)$	$X_i \leftarrow 0$, $Y_i \leftarrow 1$
Retrait	$i \in Y$	γ	$Y_i \leftarrow 0$, $Z_i \leftarrow 1$
Réintroduction	$i \in Z$	ν	$Z_i \leftarrow 0$, $X_i \leftarrow 1$ $S_i \leftarrow 0$, $D_i \leftarrow 0$

où les expressions de $P_{RA}^{(i)}$ et P_{tr} sont adaptées des équations (F.20) et (F.21) :

$$P_{RA}^{(i)}(t) = \left(1 - \frac{|Z|}{|X| + |Y| + |Z|}\right)^{\alpha_i} \quad \text{et} \quad \Lambda_{\overline{\mathrm{tr}}}(t) = \beta_{\overline{\mathrm{tr}}} \frac{|Y|}{|X| + |Y|}. \tag{F.27}$$

En dehors de ces événements ponctuels, les stocks d'offre et de demande sont alimentés continûment par la production σ_i et δ_i :

$$\dot{S}_i = \sigma_i \ \text{ et } \ \dot{D}_i = \delta_i \ \text{ en dehors des sauts.} \tag{F.28}$$

On définit ainsi un processus de Markov à temps continu et à espace d'états $E = \mathbb{R}^{+N} \times \mathbb{R}^{+N} \times \{0,1\}^N \times \{0,1\}^N \times \{0,1\}^N$ où N est le nombre total d'agents.

Afin de simuler les événements, on peut utiliser la méthode directe de Gillespie [Keeling & Rohani, 2008], qui consiste à déterminer l'instant où le prochain événement a lieu (la durée jusqu'au prochain saut suit une loi exponentielle de paramètre la somme des taux de transition, car toutes les durées de séjour sont exponentielles) :

1. Calculer le taux total de transition :

$$R_{\mathrm{total}} = \sum_{i\in X\cup Y} \sum_{j\in X\cup Y} \theta_{ij}(t) + \sum_{i\in X} \Lambda_{\overline{\mathrm{tr}}}(t) P_{RA}^{(i)}(t) + |Y|\gamma + |Z|\nu \tag{F.29}$$

2. Simuler le temps avant le prochain événement :

$$\Delta t = \frac{-\log(U)}{R_{\text{total}}} \tag{F.30}$$

où U est une variable aléatoire uniforme sur $[0;1]$.

3. Déterminer quel est le type d'événement qui a lieu et quel(s) agent(s) il concerne. Pour ce faire, on sélectionne l'événement aléatoirement, proportionnellement à son taux : l'événement de taux R est choisi avec probabilité $\frac{R}{R_{\text{total}}}$.
4. Augmenter le temps de Δt, appliquer les changements dus à l'événement choisi, mettre à jour le prix puis retourner à l'étape 1.

Ce procédé n'est valable que si les taux de transition restent constants entre deux événements. Or le taux de transaction θ_{ij} dépend indirectement du prix. Pour appliquer cet algorithme, on supposera donc que le prix est constant entre deux transactions, ce qui est raisonnable si les transactions arrivent à des intervalles de temps très rapprochés.

F.3 Paramétrisation du modèle économique, lien avec les données, implémentation

Avant de pouvoir étudier le comportement du modèle d'épidémiologie économique à l'échelle microscopique, il est nécessaire de comprendre les mécanismes sous-jacents aux échanges commerciaux. On se restreint donc à l'étude de la partie économique du modèle. Dans un premier temps, on présente les caractéristiques de la base de données. Puis, on s'intéresse aux équilibres du modèle EDO afin de le paramétrer à partir des données. On présente ensuite une implémentation efficace du modèle économique stochastique. Enfin, on définit un scénario de référence qui servira par la suite dans toutes les simulations (stochastiques ou EDO).

F.3.1 Base de données étudiée

Afin d'étudier des exemples concrets d'échanges commerciaux, on s'appuie sur les données fournies par la direction générale de l'alimentation (DGAL) du ministère de l'agriculture, qui garde un historique de la vie de chaque bovin sur le territoire français, depuis sa naissance jusqu'à sa mort (en accord avec des directives européennes mises en place suite à l'épidémie d'ESB). En particulier, sur la période 2005-2009, on dispose pour chaque mouvement d'un animal de :

- la date du mouvement ;
- l'identifiant de l'animal ;
- l'identifiant de l'exploitation source et de la destination ;
- la commune source et la commune destination.

Ces informations permettent de reconstruire les transactions entre exploitations, une transaction étant définie comme un mouvement d'une ou plusieurs bêtes entre deux exploitations à un instant donné. Ces données peuvent donc être représentées par un réseau orienté (les situations de vendeur et d'acheteur ne sont pas symétriques : le flux d'animaux va du vendeur vers l'acheteur), pondéré (une transaction peut concerner une ou plusieurs bêtes) et dynamique (les transactions ont lieu à un moment déterminé dans le temps) : les exploitations sont les nœuds et les couples d'exploitations qui échangent des animaux sont les arêtes. Ces arêtes sont orientées dans le sens du transfert d'animaux, et pondérées par le nombre d'animaux échangés. Les caractéristiques détaillées de ce réseau sont étudiées dans [Lal Dutta *et al.*, 2014]. Annuellement, la base de données est constituée d'environ 240000 fermes qui effectuent 8.6 millions de mouvements, et 5.5 millions d'animaux sont concernés. La taille du réseau tend à décroître au cours du temps ; cela est essentiellement dû aux petites fermes individuelles qui fusionnent ou qui déposent le bilan. De plus, on observe que le réseau est fortement interconnecté : environ 40% des fermes et 80% des communes sont interconnectées (i.e. sont dans la plus grande composante fortement connexe du réseau agrégé annuel).

Pour limiter la taille des données à considérer, on se place à l'échelle d'un département : le Finistère. Les raisons de ce choix sont les suivantes : ce département fait partie de la Bretagne, qui est une région avec une forte densité d'élevages bovins ; de plus, on dispose pour le Finistère d'informations supplémentaires qui sont susceptibles d'être utiles dans de futures études de modèle d'épidémiologie économique (on dispose des coordonnées géographiques des exploitations laitières et de certaines données épidémiologiques sur différentes maladies).

Ainsi, la base de données étudiée recense tous les mouvements de bovins dont la source ou la destination se trouve dans le Finistère (29), entre 2005 et 2009. Elle a les caractéristiques suivantes :

- 25802 nœuds ;
- 995618 mouvements (un mouvement correspondant à une bête) ;
- 222838 transactions (une transaction étant un lot d'une ou plusieurs bêtes échangées entre deux mêmes exploitations au cours de la même journée) ;
- 173748 arêtes orientées actives (une arête orientée (i, j) est considérée active s'il y a eu au moins un mouvement de i vers j) ;
- 5696 nœuds dans la composante fortement connexe du graphe (soit 22 % du nombre total de nœuds).

Afin d'étudier spécifiquement la situation dans le Finistère, on pourra définir une base de données réduite en ne gardant que les mouvements dont la source et la destination sont toutes les deux dans le Finistère. Ce sous-ensemble a les caractéristiques suivantes :

- 5798 nœuds ;
- 456420 mouvements ;
- 79751 transactions ;
- 55470 arêtes orientées actives ;
- 3252 nœuds dans la composante fortement connexe du graphe (soit 56 % du nombre total de nœuds).

Les distributions des degrés entrants/sortants de cette base de données sont présentées en figure F.1. On remarque que la distribution des degrés entrants est, en échelle logarithmique, proche d'une droite : elle est donc proche d'une distribution à invariance d'échelle Newman [2003], ce qui n'est pas la cas de la distribution des degrés sortants. Il y a 1723 offreurs stricts (nœuds sans voisin entrant) et 398 demandeurs stricts (nœuds sans voisin sortant).

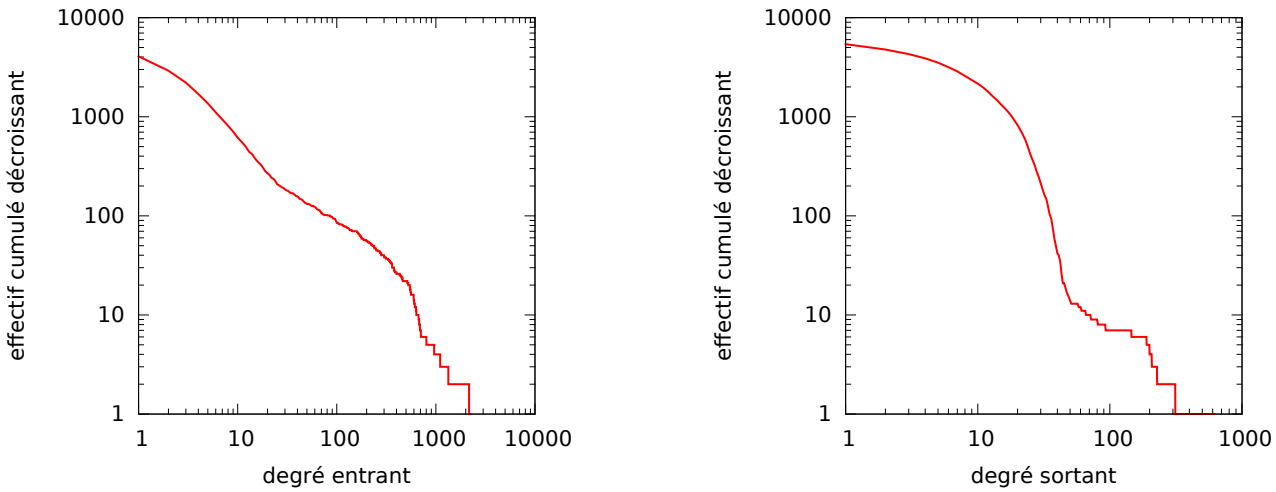


FIGURE F.1 – Distribution des degrés entrants/sortants

F.3.2 Modèle économique EDO

Le modèle économique microscopique par EDO où chaque individu est modélisé séparément s'écrit

$$\begin{aligned} \dot{S}_i &= \sigma_i - \sum_j \theta_{ij} q_{ij} \\ \dot{D}_i &= \delta_i - \sum_j \theta_{ji} q_{ji} \\ \dot{p} &= \mu p \sum_k (\delta_k - \sigma_k) \end{aligned} \tag{F.31}$$

où les coefficients de fractionnement f_{ij} et g_{ij} intervenant dans le calcul des taux de transaction $\theta_{ij}(t) = \frac{1}{\kappa_{ij}} \min\left(f_{ij}\sigma_i, g_{ij}\delta_j\right)$ sont constants au cours du temps.

Comment paramétrer le modèle à partir de la base de données d'échanges d'animaux ?

En notant par des tildes les valeurs à l'équilibre (i.e. satisfaisant $\dot{S}_i = 0$, $\dot{D}_i = 0$, $\dot{p} = 0$), le système (F.31) permet de déduire que $\tilde{\sigma}_i = \tilde{\Phi}_i^{\text{out}}$ et $\tilde{\delta}_i = \tilde{\Phi}_i^{\text{in}}$ où l'on note $\Phi_i^{\text{out}} = \sum_j \theta_{ij} q_{ij}$ et $\Phi_i^{\text{in}} = \sum_j \theta_{ji} q_{ji}$. D'après la façon dont le modèle a été construit, σ_0^i et δ_0^i sont les productions de stock à l'équilibre : on a $\tilde{\sigma}_i = \sigma_0^i$ et $\tilde{\delta}_i = \delta_0^i$ (ce qui équivaut à $\tilde{p} = 1$). On a donc

$$\sigma_0^i = \tilde{\Phi}_i^{\text{out}} \quad \text{et} \quad \delta_0^i = \tilde{\Phi}_i^{\text{in}} \tag{F.32}$$

où l'on peut estimer $\tilde{\Phi}_i^{\text{out}}$ et $\tilde{\Phi}_i^{\text{in}}$ par le nombre de bêtes vendues/achetées par i ramené à l'unité de temps. Le taux de transaction à l'équilibre est donné par

$$\tilde{\theta}_{ij} = \frac{1}{\kappa_{ij}} \min\left(\tilde{\Phi}_i^{\text{out}} f_{ij}, \tilde{\Phi}_j^{\text{in}} g_{ij}\right). \tag{F.33}$$

ce qui permet de calculer κ_{ij} à partir de la base de données puisque $\tilde{\theta}_{ij}$ peut être estimé par le nombre de transactions de i vers j ramené à l'unité de temps.

Exemple d'un réseau complètement interconnecté

Plaçons-nous dans le cas où les fonctions de fractionnement sont proportionnelles au taux de production de référence, c'est-à-dire :

$$f_{ij} = \frac{\delta_0^j}{\sum_{k\neq i} \delta_0^k} \quad \text{et} \quad g_{ij} = \frac{\sigma_0^i}{\sum_{k\neq j} \sigma_0^k} \tag{F.34}$$

si $i \neq j$ et $f_{ii} = g_{jj} = 0$. Cela traduit le fait que chaque nœud est connecté à tous les autres (à condition que ce soit un producteur), sauf à lui-même. En négligeant la production du nœud par rapport à la production totale, on a alors

$$f_{ij} \approx \frac{\tilde{\Phi}_j^{\text{in}}}{\tilde{\Phi}} \quad \text{et} \quad g_{ij} \approx \frac{\tilde{\Phi}_i^{\text{out}}}{\tilde{\Phi}} \tag{F.35}$$

où l'on note Φ le flux total : $\Phi = \sum_k \Phi_k^{\text{out}} = \sum_k \Phi_k^{\text{in}}$. Les équations (F.33) et (F.35) permettent ensuite d'obtenir l'approximation suivante de κ_{ij} :

$$\kappa_{ij} \approx \frac{\tilde{\Phi}_j^{\text{in}} \tilde{\Phi}_i^{\text{out}}}{\tilde{\Phi}\tilde{\theta}_{ij}} \tag{F.36}$$

où $\tilde{\theta}_{ij}$ peut être estimé par le nombre de transactions de i vers j ramené à l'unité de temps.

Tentons maintenant de trouver des valeurs possibles de S_i et D_i à l'équilibre sous cette approximation. Ces valeurs doivent vérifier :

$$\begin{aligned} \tilde{\Phi}_i^{\text{out}} = \tilde{\sigma}_i = \sum_j \tilde{\theta}_{ij}\tilde{q}_{ij} = \sum_j \frac{\tilde{\Phi}_j^{\text{in}} \tilde{\Phi}_i^{\text{out}}}{\tilde{\Phi}\kappa_{ij}} \min(S_i, D_j) \\ \tilde{\Phi}_i^{\text{in}} = \tilde{\delta}_i = \sum_j \tilde{\theta}_{ji}\tilde{q}_{ji} = \sum_j \frac{\tilde{\Phi}_i^{\text{in}} \tilde{\Phi}_j^{\text{out}}}{\tilde{\Phi}\kappa_{ji}} \min(S_j, D_i) \end{aligned} \tag{F.37}$$

ce qui se ramène à

$$\begin{aligned} \tilde{\Phi} = \sum_j \frac{\tilde{\Phi}_j^{\text{in}}}{\kappa_{ij}} \min(S_i, D_j) \\ \tilde{\Phi} = \sum_j \frac{\tilde{\Phi}_j^{\text{out}}}{\kappa_{ji}} \min(S_j, D_i) \end{aligned} \tag{F.38}$$

On remarque qu'en particulier, le cas où S et D sont choisis tels que $\kappa_{ij} = \min(S_i, D_j)$ pour tout i et j correspond à un équilibre. Mais de telles valeurs de S et D n'existent pas forcément. Les κ qui permettent un tel équilibre sont caractérisés de la façon suivante :

Proposition. *Soit $(\kappa_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}}$ des réels. Les propositions suivantes sont équivalentes :*

(i). Il existe $S \in \mathbb{R}^n$ et $D \in \mathbb{R}^n$ tels que $\forall i, \forall j, \kappa_{ij} = \min(S_i, D_j)$.

(ii). $\forall i, \forall j, \kappa_{ij} = \min\left(\max_{1\le k\le n} \kappa_{ik}, \max_{1\le k\le n} \kappa_{kj}\right)$.

(iii). $\forall i, \forall j, \left(\kappa_{ij} = \max_{1\le k\le n} \kappa_{ik}$ ou $\kappa_{ij} = \max_{1\le k\le n} \kappa_{kj}\right)$.

Démonstration. $(ii) \Rightarrow (iii)$: trivial.

$(iii) \Rightarrow (ii)$: comme $\max_{1\le k\le n} \kappa_{ik}$ et $\max_{1\le k\le n} \kappa_{kj}$ sont tous les deux supérieurs à κ_{ij}, si l'un des deux termes vaut κ_{ij}, le min vaut aussi κ_{ij}.

$(ii) \Rightarrow (i)$: il suffit de prendre $S_i = \max_{1\le k\le n} \kappa_{ik}$ et $D_j = \max_{1\le k\le n} \kappa_{kj}$.

$(i) \Rightarrow (ii)$: $\forall i, \forall j, S_i \ge \min(S_i, D_j) = \kappa_{ij}$ donc pour tout i, $S_i \ge \max_{1\le k\le n} \kappa_{ik}$. De même, pour tout j, $D_j \ge \max_{1\le k\le n} \kappa_{kj}$. On en déduit que :

$$\kappa_{ij} = \min(S_i, D_j) \ge \min\left(\max_{1\le k\le n} \kappa_{ik}, \max_{1\le k\le n} \kappa_{kj}\right)$$

Puisque $\max_{1\leq k\leq n} \kappa_{ik} \geq \kappa_{ij}$ et $\max_{1\leq k\leq n} \kappa_{kj} \geq \kappa_{ij}$, on a aussi l'inégalité dans l'autre sens

$$\min\left(\max_{1\leq k\leq n} \kappa_{ik}, \max_{1\leq k\leq n} \kappa_{kj}\right) \geq \kappa_{ij}$$

ce qui permet de montrer (ii). □

En pratique, les κ_{ij} déterminés expérimentalement ont peu de chance de satisfaire ces contraintes. Les conditions présentées ci-dessus sont en effet assez fortes : elles imposent entre autres que la matrice κ comporte une ligne ou une colonne constante (égale à $\min_{i,j} \kappa_{ij}$). Il est donc peu réaliste de trouver un équilibre de cette façon.

Remarquons que la proposition énoncée ci-dessus est indépendante du modèle. On peut donc l'appliquer à $q_{ij}(t) = \min(S_i(t), D_j(t))$. Ainsi, la matrice des quantités échangées q_{ij} dans le modèle EDO a une forme bien particulière : elle n'est donc pas forcément en mesure de représenter correctement une structure d'échange quelconque observée sur des données. C'est une limite d'expressivité du modèle, qui découle de la façon dont il a été construit.

Exemple d'un réseau qui reproduit les données

Si l'on veut rester au plus près des données, il est possible de définir les coefficients de fractionnement de la façon suivante :

$$f_{ij} = \frac{\tilde{\Phi}_{ij}}{\tilde{\Phi}_i^{\text{out}}} \quad \text{et} \quad g_{ij} = \frac{\tilde{\Phi}_{ij}}{\tilde{\Phi}_j^{\text{in}}}. \tag{F.39}$$

Dans ce cas, on a $\theta_{ij}(t) = \dfrac{\tilde{\Phi}_{ij}}{\kappa_{ij}}$, ce qui implique que $\tilde{q}_{ij} = \kappa_{ij}$: la friction s'interprète alors comme la quantité échangée à l'équilibre.

F.3.3 Implémentation efficace du modèle économique stochastique

Lors de la simulation par la méthode de Gillespie (partie F.2.3), les taux θ_{ij} changent entre deux événements (en effet, ils dépendent tous du prix). Recalculer à chaque itération tous les θ_{ij} est très coûteux en temps de calcul, surtout s'il y a beaucoup d'arêtes ou si les événements considérés sont fréquents (fort taux de transition). Afin d'économiser en temps de calcul, on classe les arêtes en deux catégories au sein desquelles le comportement est semblable. Pour ce faire, on tient compte de la forme particulière de l'expression de θ_{ij} :

$$\theta_{ij}(t) = \frac{1}{\kappa_{ij}} \min\left(f_{ij}\sigma_0^i p(t)^{\varepsilon}, g_{ij}\delta_0^j p(t)^{-\eta}\right). \tag{F.40}$$

Le seul élément dépendant du temps dans cette expression est le prix p. On va distinguer, selon le prix, quel est l'argument du min. Ainsi, on définit le prix limite pour l'arête (i,j) par

$$p_{ij}^{\lim} = \left(\frac{g_{ij}\delta_0^j}{f_{ij}\sigma_0^i} \right)^{\frac{1}{\varepsilon+\eta}} \tag{F.41}$$

avec la convention $p_{ij}^{\lim} = +\infty$ si $f_{ij}\sigma_0^i = 0$. Grâce à ce prix limite, on peut partitionner l'ensemble des arêtes en deux catégories F_t et G_t que l'on définit de la façon suivante :

$$\begin{aligned} (i,j) \in F_t &\Leftrightarrow p(t) \leq p_{ij}^{\lim} \\ (i,j) \in G_t &\Leftrightarrow p(t) > p_{ij}^{\lim} \end{aligned} \tag{F.42}$$

Bien sûr, les catégories évoluent au cours du temps puisque le prix p varie. Mais tant que l'arête (i,j) reste dans une des deux catégories, le taux de transaction θ_{ij} évolue de façon plus régulière :

$$\theta_{ij}(t) = \begin{cases} \theta_{ij}^F p(t)^{\varepsilon} & \text{si } (i,j) \in F_t \\ \theta_{ij}^G p(t)^{-\eta} & \text{si } (i,j) \in G_t \end{cases} \tag{F.43}$$

où l'on note θ_{ij}^F et θ_{ij}^G (ne dépendant pas du temps) les taux de transaction de référence du point de vue du vendeur et de l'acheteur :

$$\theta_{ij}^F = \frac{f_{ij}\sigma_0^i}{\kappa_{ij}} \quad \text{et} \quad \theta_{ij}^G = \frac{g_{ij}\delta_0^j}{\kappa_{ij}}. \tag{F.44}$$

Ainsi, le taux total de transition R_{total} s'écrit (premier terme de l'équation (F.29))

$$R_{\text{total}}(t) = \sum_{(i,j)} \theta_{ij}(t) = p(t)^{\varepsilon} \sum_{(i,j)\in F_t} \theta_{ij}^F + p(t)^{-\eta} \sum_{(i,j)\in G_t} \theta_{ij}^G \tag{F.45}$$

ce que l'on peut calculer sans avoir à parcourir toutes les arêtes, pourvu que l'on garde en mémoire les sommes sur F et sur G des taux de transactions de référence, qui sont constantes au cours du temps tant que F et G ne changent pas.

Finalement, on propose l'algorithme suivant :

1. Initialisation : $p_{\text{inf}} \leftarrow -1$, $p_{\text{sup}} \leftarrow -1$, $t \leftarrow 0$.
2. Calculer $p(t)$. Si $p(t) \in]p_{\text{inf}}; p_{\text{sup}}]$, F et G n'ont pas changé : aller à l'étape 4. Sinon continuer à l'étape 3.
3. Changement de F et G : recalculer $\sum_{(i,j)\in F} \theta_{ij}^F$ et $\sum_{(i,j)\in G} \theta_{ij}^G$, ainsi que les bornes de prix : $p_{\text{inf}} = \max_{p_{ij}^{\lim} < p(t)} p_{ij}^{\lim}$ et $p_{\text{sup}} = \min_{p_{ij}^{\lim} \geq p(t)} p_{ij}^{\lim}$.
4. Générer le temps Δt avant le prochain événement en utilisant une loi exponentielle de paramètre $R_{\text{total}}(t)$, que l'on calcule grâce à (F.45).

5. Déterminer quelle arête est concernée, proportionnellement à son taux. Pour ce faire, on tire une variable U uniforme sur $]0;1]$. Si $R_{\text{total}}(t)U \leq p(t)^{\varepsilon} \sum_{(i,j)\in F} \theta_{ij}^F$, l'arête est choisie dans F ; en supposant que l'on a fixé un ordre complet sur les arêtes, on choisit l'arête (i_0, j_0) telle que :

$$\sum_{\substack{(i,j)\in F \\ (i,j)<(i_0,j_0)}} \theta_{ij}^F < R_{\text{total}}(t)Up(t)^{-\varepsilon} \leq \sum_{\substack{(i,j)\in F \\ (i,j)\leq(i_0,j_0)}} \theta_{ij}^F. \tag{F.46}$$

   Si $R_{\text{total}}(t)U > p(t)^{\varepsilon} \sum_{(i,j)\in F} \theta_{ij}^F$, l'arête est choisie dans G suivant une règle similaire. Pour trouver efficacement l'arête, on peut procéder par dichotomie, à condition d'avoir mémorisé les sommes cumulées partielles des θ_{ij}^F et des θ_{ij}^G lors de l'étape 3.

6. Augmenter le temps de Δt, simuler l'échange entre i_0 et j_0, puis retourner à l'étape 2.

Afin de disposer efficacement des sommes cumulées partielles des θ_{ij}^F et des θ_{ij}^G, il suffit de trier initialement les arêtes par ordre de $p_{ij}^{\lim}$.

F.3.4 Définition du scénario de simulation de référence

Dans toute la suite, que ce soit dans les simulations du modèle EDO ou du modèle stochastique, on se place dans le même scénario de référence. On définit ce scénario à l'aide de paramètres calculés à partir de la base de données.

On fixe la production d'offre de référence d'un nœud i au flux moyen sortant de i dans la base de données (nombre total de bêtes vendues par i divisé par la durée d'observation qui est de 5 ans) : $\sigma_0^i = \tilde{\Phi}_i^{\text{out}}$. De même, on fixe la production de demande de référence d'un nœud i au flux moyen entrant en i dans la base de données (nombre total de bêtes achetées par i divisé par la durée d'observation) : $\delta_0^i = \tilde{\Phi}_i^{\text{in}}$.

Afin de faciliter la simulation, on associe à chaque nœud i du réseau des voisinages V_i^{in} et V_i^{out} de petite taille : ceux observés dans la base de données. Le nœud j est dans V_i^{in} s'il y a au moins une transaction de j vers i dans la base de données. De même, le nœud j est dans V_i^{out} s'il y a au moins une transaction de i vers j dans la base de données. On se donne aussi les coefficients de fractionnement suivants :

$$\begin{aligned} f_{ij} &= \frac{\delta_0^j}{\sum_{k\in V_i^{\text{out}}} \delta_0^k} \quad \text{si } j \in V_i^{\text{out}}, f_{ij} = 0 \text{ sinon} \\ g_{ij} &= \frac{\sigma_0^i}{\sum_{k\in V_j^{\text{in}}} \sigma_0^k} \quad \text{si } i \in V_j^{\text{in}}, g_{ij} = 0 \text{ sinon} \end{aligned} \tag{F.47}$$

On fixe la friction de telle sorte que le taux de transaction à l'équilibre donné par (F.33) corresponde

à celui observé dans la base de données. Ainsi, on prend

$$\kappa_{ij} = \frac{1}{\tilde{\theta}_{ij}} \min\left(\tilde{\Phi}_i^{\text{out}} f_{ij}, \tilde{\Phi}_j^{\text{in}} g_{ij}\right) \tag{F.48}$$

où $\tilde{\theta}_{ij}$ est le nombre de transactions de i vers j divisé par la durée d'observation.

Les stocks initiaux sont pris nuls. La simulation est réalisée sur 5000 jours.

Enfin, on fixe le prix initial à 1. Comme les σ_0^i et δ_0^i ont été construits de telle sorte que $\sum_i \sigma_0^i = \sum_i \delta_0^i$ et comme le prix suit l'équation $\dot{p} = \mu p \sum_i (\delta_0^i p^{-\eta} - \sigma_0^i p^{\varepsilon})$, cela implique que le prix reste constant, égal à 1 : il n'y a pas de dynamique de prix, que le modèle soit stochastique ou déterministe.

F.4 Analyse et simulation du modèle économique

Afin d'interpréter les modèles microscopiques précédemment décrits, on étudie d'abord le lien théorique entre le modèle stochastique et le modèle EDO. Puis on simule ces deux modèles, en les confrontant entre eux et avec les données. Enfin, on procède à des expériences de percolation pour estimer la propagation de l'épidémie sur le réseau.

F.4.1 Lien entre modèle stochastique et équations différentielles

Afin de rendre compte de la proximité entre les EDO et le processus stochastique, nous nous intéressons à l'espérance de ce dernier. Il est possible d'exprimer l'évolution des moments du modèle stochastique par des équations différentielles. Ainsi, dans [Daley *et al.*, 2001], il est montré à l'aide de fonctions génératrices que pour un modèle épidémique stochastique simple à deux états (l'effectif des susceptibles est noté $X(t)$ et l'effectif des infectés est noté $Y(t)$) et à une seule transition (les susceptibles deviennent infectés au taux $\beta Y(t)$), l'espérance de $X(t)$ vérifie l'équation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(X(t)) = -\beta\mathbb{E}(X(t))\,(N - \mathbb{E}(X(t))) + \operatorname{Var} X(t) \tag{F.49}$$

où on a fixé les conditions initiales $X(0) = N - I$ et $Y(0) = I$. En effet, on est ramené à l'étude d'un processus de Markov à temps continu et espace d'états discrets. À un terme de variance près, on obtient donc la même équation que le modèle déterministe équivalent, qui est $\dot{X}(t) = -\beta X(t)(N - X(t))$. Dans cette section, on montre des résultats analogues pour le modèle économique à l'échelle microscopique.

Considérons tout d'abord le modèle économique où les transactions et la production sont stochastiques : pour tout agent i, les événements $S_i \to S_i + 1$ et $D_i \to D_i + 1$ arrivent avec les taux respectifs $\sigma_i(t)$ et $\delta_i(t)$; pour toute arête (i, j), la transaction entre i et j se fait au taux $\theta_{ij}(t)$. Même si ce n'est pas le modèle que nous avons étudié dans les simulations, il présente l'avantage de ne permettre que des valeurs entières positives de stocks : l'espace des états possibles du processus de Markov est donc discret. Les taux σ_i,δ_i et θ_{ij} dépendent du prix qui est une fonction de S et D, on a donc des transitions dépendantes de l'état. Les stocks S et D sont assimilés à des vecteurs de $\mathbb{N}^K$ où K est le nombre d'agents. On note e_i le vecteur dont toutes les coordonnées sont nulles sauf la i-ème qui vaut 1.

Ainsi, si les valeurs de stocks à t sont (S, D), les transitions possibles jusqu'en $t + \mathrm{d}t$ sont les suivantes, à $o(\mathrm{d}t)$ près :

- Production d'une unité d'offre i : $(S, D) \to (S + e_i, D)$ avec probabilité $\sigma_i(S, D)\mathrm{d}t$.
- Production d'une unité de demande i : $(S, D) \to (S, D + e_i)$ avec probabilité $\delta_i(S, D)\mathrm{d}t$.
- Transaction entre i et j : $(S, D) \to (S - \min(S_i, D_j)e_i, D - \min(S_i, D_j)e_j)$ avec probabilité $\theta_{ij}(S, D)\mathrm{d}t$.

– Pas d'événement : $(S,D) \to (S,D)$ avec probabilité $1 - \left(\sum_i \sigma_i + \sum_i \delta_i + \sum_{ij} \theta_{ij}\right) \mathrm{d}t$.

Réciproquement, les transitions entre t et $t + \mathrm{d}t$ qui permettent d'arriver à l'état (S,D) sont les suivantes :

– $(S - e_i, D) \to (S,D)$ si $S_i \geq 1$ avec probabilité $\sigma_i(S - e_i, D)\mathrm{d}t$.
– $(S, D - e_i) \to (S,D)$ si $D_i \geq 1$ avec probabilité $\delta_i(S, D - e_i)\mathrm{d}t$.
– $(S + ke_i, D + ke_j) \to (S,D)$ pour tout $k \in \mathbb{N}$ si $S_i = 0$ ou $D_j = 0$ avec probabilité $\theta_{ij}(S + ke_i, D + ke_j)\mathrm{d}t$.
– $(S,D) \to (S,D)$ avec probabilité $1 - \left(\sum_i \sigma_i + \sum_i \delta_i + \sum_{ij} \theta_{ij}\right) \mathrm{d}t$, où les taux sont évalués en (S,D).

En notant $P_{S,D}(t)$ la probabilité d'être dans l'état (S,D) à la date t, on obtient l'égalité suivante :

$$\begin{aligned} P_{S,D}(t + \mathrm{d}t) = &\sum_i \mathbb{1}_{S_i \geq 1} \sigma_i(S - e_i, D, t)\mathrm{d}t P_{S-e_i,D}(t) + \sum_i \mathbb{1}_{D_i \geq 1} \sigma_i(S, D - e_i, t)\mathrm{d}t P_{S,D-e_i}(t) \\ &+ \sum_{i,j} \mathbb{1}_{S_i = 0 \text{ ou } D_j = 0} \sum_{k \in \mathbb{N}} \theta_{ij}(S + ke_i, D + ke_j, t)\mathrm{d}t P_{S+ke_i, D+ke_j}(t) \\ &+ \left(1 - \left(\sum_i \sigma_i(S,D,t) + \sum_i \delta_i(S,D,t) + \sum_{ij} \theta_{ij}(S,D,t)\right)\mathrm{d}t\right) P_{S,D}(t) \end{aligned}$$

Passer à la limite $\mathrm{d}t \to 0$ permet de trouver l'équation :

$$\begin{aligned} \dot{P}_{S,D}(t) = &\sum_i \mathbb{1}_{S_i \geq 1} \sigma_i(S - e_i, D, t) P_{S-e_i,D}(t) + \sum_i \mathbb{1}_{D_i \geq 1} \sigma_i(S, D - e_i, t) P_{S,D-e_i}(t) \\ &+ \sum_{i,j} \mathbb{1}_{S_i = 0 \text{ ou } D_j = 0} \sum_{k \in \mathbb{N}} \theta_{ij}(S + ke_i, D + ke_j, t) P_{S+ke_i, D+ke_j}(t) \\ &- \left(\sum_i \sigma_i(S,D,t) + \sum_i \delta_i(S,D,t) + \sum_{ij} \theta_{ij}(S,D,t)\right) P_{S,D}(t) \end{aligned}$$

Définissons maintenant φ, la fonction génératrice de (S,D) :

$$\varphi(z,t) = \mathbb{E}\left(\left(\prod_{i=1}^K z_i^{S_i(t)}\right)\left(\prod_{j=1}^K z_{K+j}^{D_j(t)}\right)\right) = \sum_{(S,D) \in \mathbb{N}^{2K}} \left(\prod_{i=1}^K z_i^{S_i}\right)\left(\prod_{j=1}^K z_{K+j}^{D_j}\right) P_{S,D}(t). \tag{F.50}$$

Cette fonction vérifie $\varphi(1,t) = 1$, $\dfrac{\partial \varphi}{\partial z_i}(1,t) = \mathbb{E}(S_i(t))$ et $\dfrac{\partial \varphi}{\partial z_{K+i}}(1,t) = \mathbb{E}(D_i(t))$ pour $1 \leq i \leq K$. Grâce aux équations précédentes, on peut trouver une expression de $\dfrac{\partial \varphi}{\partial t}$. Cette expression peut être simplifiée en utilisant, pour i, j fixés, le changement de variable

$$\psi : \left\{ \begin{array}{ccc} \{(S,D) \in \mathbb{N}^{2K} | S_i = 0 \text{ ou } D_j = 0\} \times \mathbb{N} & \to & \mathbb{N}^{2K} \\ (S,D,k) & \mapsto & (S + ke_i, D + ke_j) \end{array} \right.$$

de réciproque

$$\psi^{-1} : \begin{cases} \mathbb{N}^{2K} & \to & \{(S,D) \in \mathbb{N}^{2K} | S_i = 0 \text{ ou } D_j = 0\} \times \mathbb{N} \\ (S,D) & \mapsto & (S - \min(S_i, D_j)e_i, D - \min(S_i, D_j)e_j, \min(S_i, D_j)) \end{cases}$$

qui permet d'écrire pour toute fonction f :

$$\sum_{(S,D)\in\mathbb{N}^{2K}} \mathbb{1}_{S_i=0 \text{ ou } D_j=0} \sum_{k\in\mathbb{N}} f(S,D,k) = \sum_{(S,D)\in\mathbb{N}^{2K}} f(\psi^{-1}(S,D)).$$

De même

$$\sum_{(S,D)\in\mathbb{N}^{2K}} \mathbb{1}_{S_i\geq 1} f(S,D) = \sum_{(S,D)\in\mathbb{N}^{2K}} f(S+e_i, D)$$

et

$$\sum_{(S,D)\in\mathbb{N}^{2K}} \mathbb{1}_{D_i\geq 1} f(S,D) = \sum_{(S,D)\in\mathbb{N}^{2K}} f(S, D+e_i).$$

Finalement, on obtient :

$$\frac{\partial\varphi}{\partial t} = \sum_{(S,D)\in\mathbb{N}^{2K}} \left(\prod_{i=1}^{K} z_i^{S_i}\right)\left(\prod_{j=1}^{K} z_{K+j}^{D_j}\right) P_{S,D}(t) \left(\sum_i (z_i - 1)\sigma_i(S,D,t) + \sum_i (z_{K+i} - 1)\delta_i(S,D,t) + \sum_{i,j}\left(\frac{1}{(z_i z_{K+j})^{\min(S_i,D_j)}} - 1\right)\theta_{ij}(S,D,t)\right).$$

Dériver par rapport à z_i ou z_{K+i} puis évaluer en $z = 1$ permet d'obtenir :

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(S_i(t)) &= \mathbb{E}(\sigma_i(t)) - \sum_j \mathbb{E}(\theta_{ij}(t)\min(S_i(t), D_j(t))) \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(D_i(t)) &= \mathbb{E}(\delta_i(t)) - \sum_j \mathbb{E}(\theta_{ji}(t)\min(S_i(t), D_j(t))). \end{aligned} \tag{F.51}$$

Si les taux σ, δ, θ sont déterministes (c'est-à-dire ne dépendent pas de S et D), on a des équations de la forme $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(S_i(t)) = \sigma_i(t) - \sum_j \theta_{ij}(t)\mathbb{E}(\min(S_i(t), D_j(t)))$.

Étudions maintenant le cas où la production est déterministe et où seules les transactions sont stochastiques. C'est le modèle que nous avons utilisé dans les simulations. Dans ce cas, le prix p et les taux σ, δ, θ sont déterministes. Cette situation est plus difficile car les stocks prennent des valeurs continues et non plus discrètes. Si les valeurs de stocks en t sont (S, D), les transitions possibles jusqu'en $t + \mathrm{d}t$ sont les suivantes, à $o(\mathrm{d}t)$ près :

- Transaction : $(S,D) \to (S + \sigma\mathrm{d}t - \min(S_i, D_j)e_i, D + \delta\mathrm{d}t - \min(S_i, D_j)e_j)$ avec probabilité $\theta_{ij}(t)\mathrm{d}t$.
- Pas de transaction : $(S,D) \to (S + \sigma\mathrm{d}t, D + \delta\mathrm{d}t)$ avec probabilité $1 - \sum_{ij}\theta_{ij}(t)\mathrm{d}t$.

Pour toute fonction ϕ intégrable, on a :

$$\mathbb{E}\left[\phi((S,D)(t+\mathrm{d}t))\right] = \mathbb{E}\Big(\mathbb{E}\left[\phi((S,D)(t+\mathrm{d}t))|S(t),D(t)\right]\Big) \tag{F.52}$$

où

$$\begin{aligned}\mathbb{E}\left[\phi((S,D)(t+\mathrm{d}t))|S(t),D(t)\right] &= \phi(S(t)+\sigma(t)\mathrm{d}t, D(t)+\delta(t)\mathrm{d}t)\left(1-\sum_{ij}\theta_{ij}(t)\mathrm{d}t\right)\\ &+\sum_{ij}\phi(S(t)+\sigma(t)\mathrm{d}t-\min(S_i(t),D_j(t))e_i, D(t)+\delta(t)\mathrm{d}t-\min(S_i(t),D_j(t))e_j)\theta_{ij}(t)\mathrm{d}t.\end{aligned}$$

De plus, à condition que ϕ soit dérivable :

$$\begin{aligned}\lim_{\mathrm{d}t\to 0}\frac{1}{\mathrm{d}t}&\left(\mathbb{E}\left[\phi(S(t)+\sigma(t)\mathrm{d}t, D(t)+\delta(t)\mathrm{d}t)\right]-\mathbb{E}\left[\phi(S(t),D(t))\right]\right)\\ &= \mathbb{E}\left[\frac{\partial\phi}{\partial S}(S(t),D(t))\sigma(t)+\frac{\partial\phi}{\partial D}(S(t),D(t))\delta(t)\right].\end{aligned}$$

Ainsi, $\mathbb{E}\left[\phi(S(t),D(t))\right]$ vérifie l'équation suivante :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\phi(S(t),D(t))\right] &= \mathbb{E}\left[\frac{\partial\phi}{\partial S}(S,D)\right]\sigma(t)+\mathbb{E}\left[\frac{\partial\phi}{\partial D}(S,D)\right]\delta(t)-\sum_{i,j}\theta_{ij}(t)\mathbb{E}\left[\phi(S,D)\right]\\ &+\sum_{i,j}\theta_{ij}(t)\mathbb{E}\left[\phi(S-\min(S_i,D_j)e_i, D-\min(S_i,D_j)e_j)\right].\end{aligned}$$

On aurait aussi pu retrouver ce résultat en utilisant l'équation de Kolmogorov.

En prenant $\phi(S,D)=S_i$ ou $\phi(S,D)=D_i$, on obtient les équations suivantes :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(S_i(t)) &= \sigma_i(t)-\sum_j\theta_{ij}(t)\mathbb{E}(\min(S_i(t),D_j(t)))\\ \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(D_i(t)) &= \delta_i(t)-\sum_j\theta_{ji}(t)\mathbb{E}(\min(S_i(t),D_j(t)))\end{aligned} \tag{F.53}$$

qui sont proches des équations que l'on utilise dans le modèle EDO, qui sont les suivantes :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}S_i(t) &= \sigma_i(t)-\sum_j\theta_{ij}(t)\min(S_i(t),D_j(t))\\ \frac{\mathrm{d}}{\mathrm{d}t}D_i(t) &= \delta_i(t)-\sum_j\theta_{ji}(t)\min(S_i(t),D_j(t)).\end{aligned} \tag{F.54}$$

Ainsi, l'espérance des stocks dans le modèle stochastique vérifie les équations du modèle EDO à un terme $\sum_j\theta_{ij}(t)\left(\min(\mathbb{E}(S_i(t)),\mathbb{E}(D_j(t)))-\mathbb{E}(\min(S_i(t),D_j(t)))\right)$ près. Ce terme, qui traduit la variabilité des S_i et D_j, peut être mis en parallèle avec le terme $\mathrm{Var}\,X(t)$ obtenu dans l'équation (F.49) pour le modèle simple.

F.4.2 Simulations stochastiques

Dans cette section, on simule la partie économique du modèle stochastique, avec les paramètres du scénario de référence défini en partie F.3.4. On travaille donc sur la fenêtre de temps $[0;T]$ où $T = 5000$ jours. La simulation est implémentée en C++ par l'algorithme décrit en partie F.3.3.

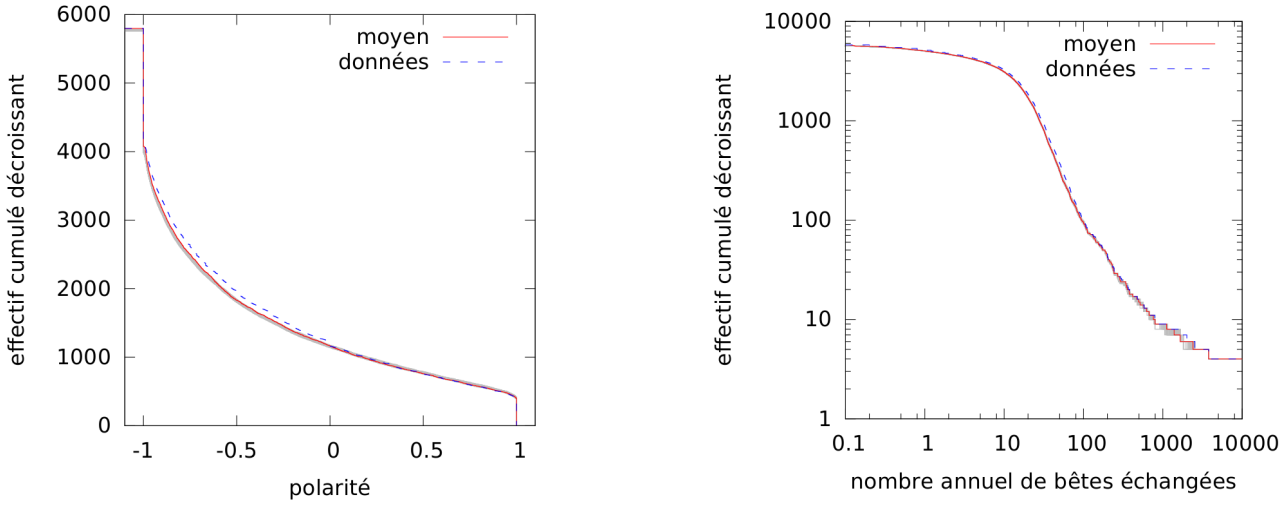
Afin d'étudier la correspondance du modèle stochastique avec les données, on s'intéresse à deux variables, la polarité

$$\rho_i = \frac{\tilde{\phi}_i^{\text{in}} - \tilde{\phi}_i^{\text{out}}}{\tilde{\phi}_i^{\text{in}} + \tilde{\phi}_i^{\text{out}}} \tag{F.55}$$

et le nombre de bêtes échangées par unité de temps

$$m_i = \tilde{\phi}_i^{\text{in}} + \tilde{\phi}_i^{\text{out}} \tag{F.56}$$

où $\tilde{\phi}_i^{\text{in}}$ (respectivement $\tilde{\phi}_i^{\text{out}}$) est le flux moyen entrant (respectivement sortant) obtenu en divisant le nombre total de bêtes achetées par i (respectivement vendues par i) pendant toute la simulation par la durée de simulation. La polarité, comprise entre -1 et 1, caractérise à quel point le nœud est offreur ou demandeur : une polarité de -1 correspond à un offreur strict, une polarité de 1 correspond à un demandeur strict. Le nombre de bêtes échangées quantifie l'importance de la production du nœud et donne une idée de sa part de marché. On peut remarquer que la connaissance de la polarité et du nombre de bêtes échangées permet de retrouver les flux moyens entrants et sortants.



(a) Polarité

(b) Nombre annuel de bêtes échangées

FIGURE F.2 – Distribution de deux grandeurs : la polarité et le nombre de bêtes échangées. En gris, les distributions de 200 simulations du scénario de référence. En rouge, les distributions des moyennes sur ces simulations. En pointillés bleus, les distributions calculées sur la base de données.

En figure F.2, on a tracé les distributions de la polarité et du nombre annuel de bêtes échangées pour 200 simulations (en gris). On a également représenté les distributions des moyennes sur toutes

les simulations de ces grandeurs (en rouge), et les distributions calculées sur la base de données (en pointillés bleus). On remarque que la variabilité est assez faible (les simulations sont très proches de la moyenne). Les polarités simulées ainsi que le nombre annuel de bêtes échangées respectent bien les valeurs observées sur la base de données.

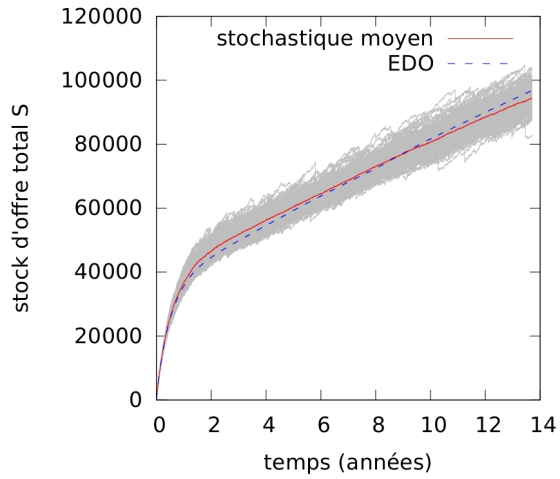


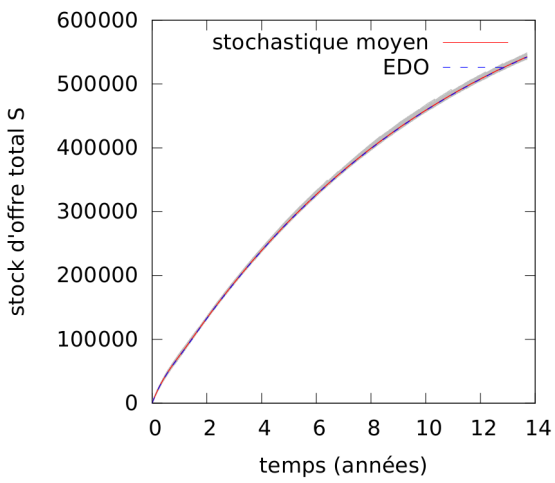
FIGURE F.3 – Évolution du stock total d'offre S lors de 200 simulations stochastiques du scénario de référence (en gris), moyenne de ce stock total sur les 200 simulations (en rouge) et stock total donné par le modèle EDO (en pointillés bleus)

Afin de corroborer les résultats trouvés dans la partie précédente, on compare l'évolution des stocks totaux dans les modèles stochastiques et EDO. En figure F.3, on a représenté l'évolution du stock total d'offre S lors de 200 simulations stochastiques du scénario de référence (en gris), sa moyenne (en rouge) et l'évolution donnée par le modèle EDO (en pointillés bleus). Le stock total de demande D n'est pas représenté car son évolution est ici exactement la même que celle de S : en effet, dans le scénario de référence, il n'y a pas de dynamique de prix, ce qui impose l'égalité des stocks totaux d'offre et de demande. On constate que la variabilité autour de la moyenne est notable, mais que la moyenne du modèle stochastique est très proche du modèle EDO.

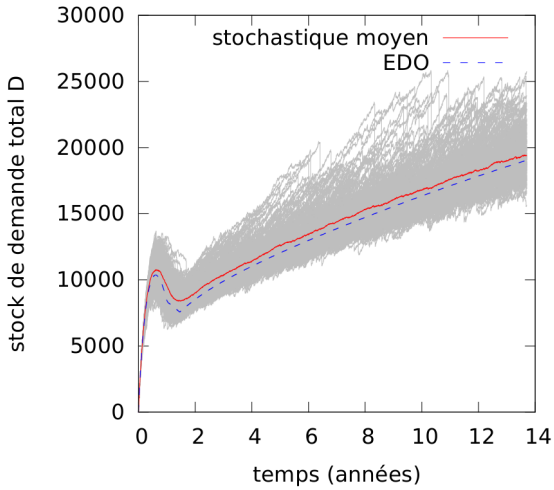
On a représenté en figure F.4 ce qui se passe quand on n'est plus dans le scénario de référence : on se place dans un scénario identique au scénario de référence à la différence près que l'on fixe le prix initial à 2. On observe que le stock d'offre prend de grandes valeurs (ce qui est intuitif : le prix fort incite les agents à produire de l'offre), ce qui diminue l'importance relative de sa variabilité. Dans les deux cas, la moyenne du modèle stochastique reste proche du modèle EDO.

F.4.3 Simulations par EDO

Dans cette section, on simule le modèle économique EDO, en le paramétrant selon le scénario de référence décrit en partie F.3.4. On prend donc toujours $p(t = 0) = 1$, qui correspond à un équilibre : il n'y a donc pas de dynamique de prix. On effectue la simulation sur 5000 jours, soit environ 14 années.



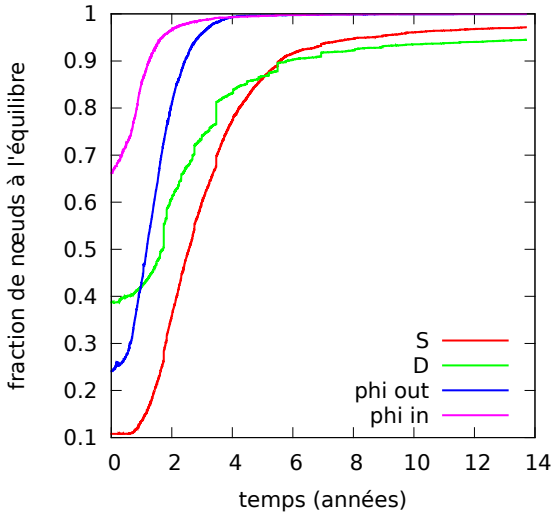
(a) Stock total d'offre S



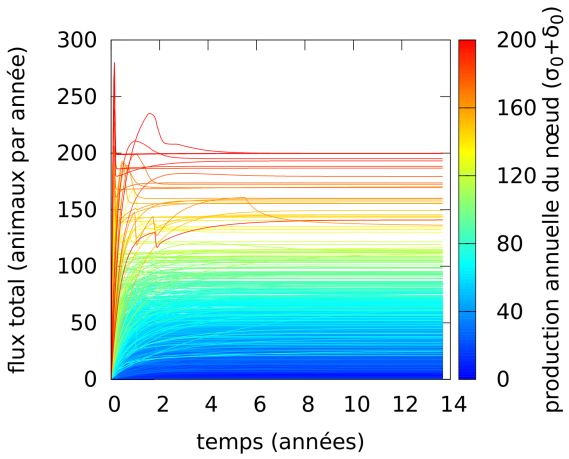
(b) Stock total de demande D

FIGURE F.4 – Évolution du stock total lors de 200 simulations stochastiques avec dynamique de prix (en gris), moyenne de ce stock total sur les 200 simulations (en rouge) et stock total donné par le modèle EDO (en pointillés bleus)

La simulation est implémentée en C++ par la méthode d'Euler.



(a) Fraction des nœuds à l'équilibre pour différentes grandeurs (nœud considéré à l'équilibre si la dérivée de la grandeur est inférieure à un seuil arbitraire fixé)

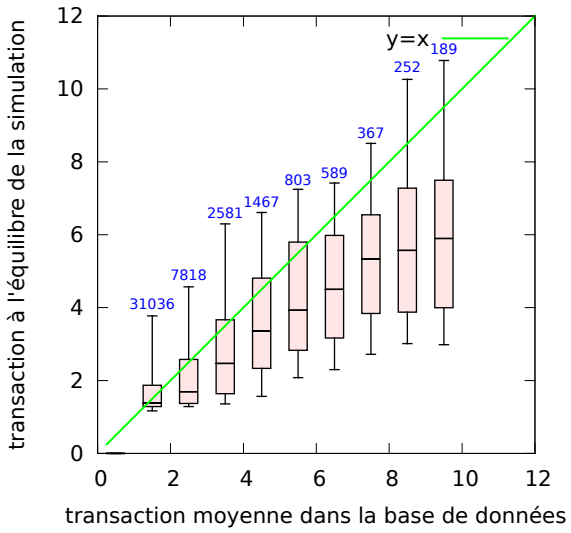


(b) Évolution du flux total $\phi_i^{\text{in}} + \phi_i^{\text{out}}$ (couleur en fonction de la production de référence $\sigma_0^i + \delta_0^i$) pour tous les nœuds de production annuelle inférieure à 200 animaux (43 nœuds omis)

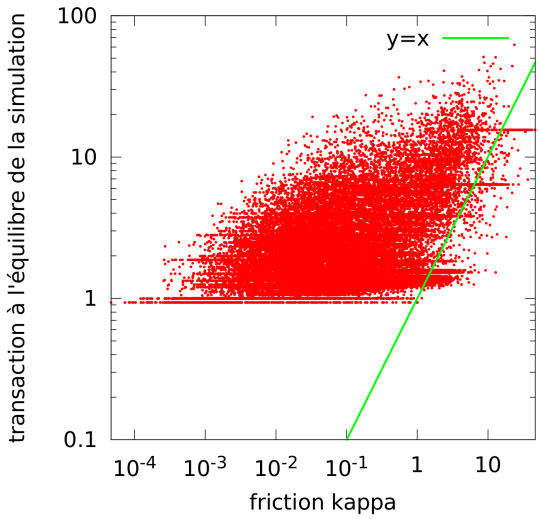
FIGURE F.5 – Étude de l'équilibre du modèle EDO

Intéressons-nous à l'équilibre du système. La figure F.5a présente la fraction des nœuds qui sont à l'équilibre au cours du temps. Un nœud est considéré comme étant à l'équilibre pour une certaine grandeur si la dérivée de cette grandeur est inférieure à un certain seuil empirique. Pour l'équilibre en S ou D, on a pris un seuil de 0.2 animal/année. Pour l'équilibre en flux entrant ou sortant, on a pris un seuil de 10^{-5} animaux/jours^2, soit environ 1 animal/année^2. On remarque que l'équilibration n'est pas immédiate mais qu'en fin de simulation, la grande majorité des nœuds est à l'équilibre. La figure F.5b, qui représente les flux sortants en fonction du temps, permet de constater que l'équilibre est atteint assez rapidement.

Étudions maintenant plus en détail les valeurs à l'équilibre : sont-elles cohérentes avec les données ? Dans toute la suite, on définit les valeurs à l'équilibre comme étant les valeurs au temps final (t = 5000 jours) de la simulation, à condition que les seuils de dérivées définis précédemment soient respectés : on ne représente dans les graphiques que les valeurs correspondant aux nœuds et arêtes que l'on considère à l'équilibre. La figure F.6 permet de confronter les valeurs à l'équilibre des q_{ij} avec les valeurs des données et les frictions. Comme il n'y a pas de dynamique des prix dans le scénario simulé, les taux de transaction θ_{ij} sont constants au cours du temps, et leurs valeurs correspondent à celles de la base de données par construction. Ainsi, il est équivalent d'étudier l'adéquation de q_{ij} aux données et l'adéquation de ϕ_{ij} aux données puisque $\phi_{ij}(t) = \theta_{ij} q_{ij}(t)$. On constate sur la figure F.6a que les transactions à l'équilibre correspondent assez bien aux données. Sur la figure F.6b, on peut voir que les frictions correspondent en un certain sens à une borne inférieure de la transaction, ce qui est assez intuitif : on a vu plus tôt que la friction caractérise les contraintes liées à la logistique de la transaction



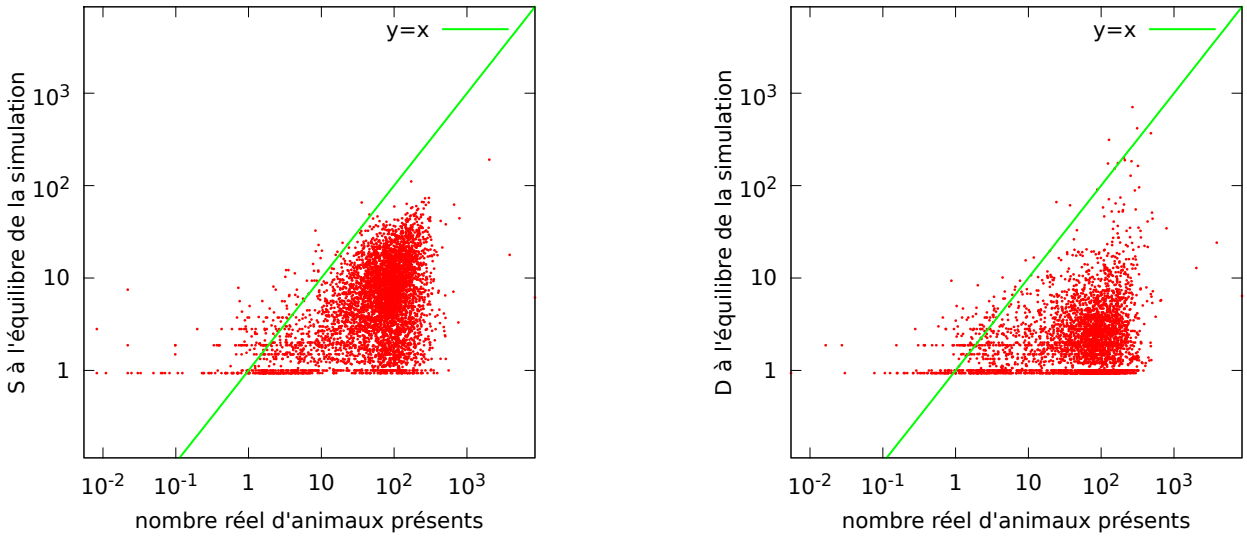
(a) Les valeurs des transactions moyennes ont été discrétisées par intervalles de taille 1 (les valeurs supérieures à 10 ont été omises ; elles représentent 2 % des arêtes). Les effectifs correspondant à chaque boîte sont indiqués en bleu.



(b) Chaque point représente une arête à l'équilibre ($95,8$ % des arêtes sont au-dessus de la droite $y = x$)

FIGURE F.6 – Comparaison des valeurs à l'équilibre des q_{ij} avec les valeurs des données (a) et avec les frictions κ_{ij} (b)

(par exemple, si les transactions doivent toujours concerner au moins 2 animaux, on s'attend à ce que $\kappa \geq 2$).



(a) Stocks d'offre

(b) Stocks de demande

FIGURE F.7 – Comparaison des stocks à l'équilibre avec le nombre d'animaux physiquement présents dans les fermes (fourni par la base de données)

La base de données fournit pour chaque exploitation le nombre d'animaux physiquement présents. Il est donc intéressant de comparer ce nombre aux stocks d'offre et de demande simulés. La figure F.7 présente ces résultats. On remarque que les stocks S et D sont inférieurs au nombre d'animaux présents. C'est cohérent avec le modèle : il est naturel que les stocks disponibles à la vente soient inférieurs aux stocks physiques.

F.4.4 Étude des propriétés de propagation du réseau simulé par une expérience de percolation

L'objectif ultime de l'étude est d'étudier la propagation d'épidémie sur le réseau. Pour ce faire, on simule les échanges commerciaux de façon stochastique, puis on utilise une approximation courante dans la littérature [Newman, 2003] : on agrège le réseau sur une fenêtre de temps, et on calcule la taille de la plus grande composante fortement connexe (GSCC). Cela donne un ordre de grandeur du nombre de nœuds qui seraient infectés en cas d'épidémie. On commet cependant une double erreur : on néglige la dynamique du réseau (le fait que les arêtes du réseau soient ordonnées temporellement peut bloquer l'épidémie, ce qui n'est pas le cas pour le réseau agrégé) et la dynamique de l'épidémie (même sur un réseau statique qui reste fixe au cours du temps, l'épidémie peut toucher une zone plus petite ou plus grande que la GSCC). De plus, la proportion des nœuds du graphe entier que contient la GSCC permet d'avoir une idée de la probabilité qu'une épidémie d'importance ait lieu.

Pour chaque arête (i, j), le taux de transaction θ_{ij} est déterministe et tend vers le taux de transaction observé dans la base de données. Sur une fenêtre de temps assez grande, le réseau simulé agrégé sera donc le même que le réseau observé dans la base de données. Ainsi, on définira le temps critique de la façon suivante : c'est la taille de la fenêtre de temps nécessaire pour que la taille de la GSCC du réseau simulé atteigne la moitié de la taille de la GSCC du réseau de la base de données. Bien sûr, pour éliminer la variabilité due à l'aspect stochastique, on moyenne la taille de la GSCC sur 200 simulations.

En pratique, il n'y a pas besoin de simuler toutes les transactions pour générer le réseau agrégé : les taux de transactions θ_{ij} étant déterministes, et les instants des transactions étant indépendants (mais pas la quantité échangée), la probabilité $p_{ij}(t_1, t_2)$ qu'il y ait au moins une transaction entre i et j sur la fenêtre de temps $[t_1; t_2]$ est donnée par

$$p_{ij}(t_1, t_2) = 1 - \exp\left(-\int_{t_1}^{t_2} \theta_{ij}(t)\mathrm{d}t\right). \tag{F.57}$$

On peut donc générer un réseau agrégé sur $[t_1, t_2]$ simplement en sélectionnant chaque arête de façon indépendante avec une probabilité $p_{ij}(t_1, t_2)$.

Grâce au temps critique, on peut évaluer l'efficacité de certaines politiques publiques : l'objectif du régulateur public est d'augmenter le temps critique, afin de disposer de plus de temps pour enrayer

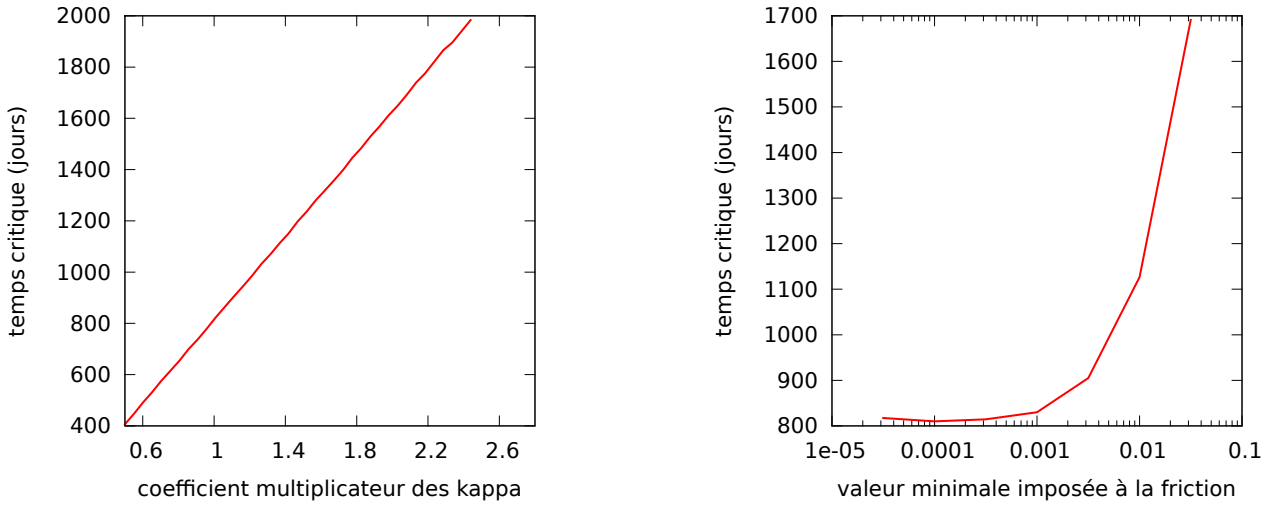
l'épidémie. Dans la section précédente, nous avons vu que la friction κ_{ij} correspondait dans le modèle à une borne inférieure de la quantité échangée lors d'une transaction. Or, il est possible pour un régulateur public d'altérer la friction. Nous avons donc tracé l'évolution du temps critique quand on modifie les κ.

Dans un premier cas (figure F.8a), on a multiplié toutes les frictions de référence par le même facteur. Cela correspond par exemple au cas où un régulateur public taxe les transactions, ce qui incite à faire moins de transactions, mais à échanger plus de bêtes à chaque fois. On observe sur les simulations que le temps critique évolue linéairement en fonction de ce facteur multiplicatif. Ce résultat est dû au fait que dans le scénario de simulation choisi, il n'y a pas de dynamique de prix et donc les taux de transaction θ_{ij} sont constants. De plus, après les changements sur la friction, on aura $\theta_{ij} = \theta_{ij}^{\text{ref}} \frac{\kappa_{ij}^{\text{ref}}}{\kappa_{ij}}$. La probabilité de présence d'une arête devient donc

$$p_{ij}(t_1, t_2) = 1 - \exp\left((t_1 - t_2) \theta_{ij}^{\text{ref}} \frac{\kappa_{ij}^{\text{ref}}}{\kappa_{ij}} \right) \tag{F.58}$$

ce qui explique la relation linéaire observée.

Dans un deuxième cas (figure F.8b), on impose une valeur minimale $\kappa_{\min}$ aux frictions. Ainsi, on change la friction de référence κ_{ij}^{ref} en $\max(\kappa_{ij}^{\text{ref}}, \kappa_{\min})$. On peut remarquer que pour $\kappa_{\min} = 0.03$, ce qui est une valeur assez faible, le temps critique est doublé par rapport à la situation de référence. Un régulateur public peut donc facilement ralentir l'épidémie en agissant sur la friction, par exemple en imposant un nombre minimal d'animaux par transaction. Si l'on trouve ici des valeurs intéressantes de $\kappa_{\min}$ plus petites que 1, c'est parce que les frictions de référence sont souvent inférieures à 1. C'est un inconvénient du modèle que l'on pourra essayer de corriger, par exemple en changeant les coefficients de fractionnement.



(a) Dilatation des κ_{ij}

(b) Imposition d'une valeur minimale de κ_{ij}

FIGURE F.8 – Évolution du temps critique (temps pour que la GSCC du réseau simulé agrégé atteigne la moitié de sa valeur maximale) lorsque l'on fait varier les frictions κ_{ij}

F.5 Conclusion et perspectives

Grâce à un modèle macroscopique d'épidémiologie économique, nous avons développé deux modèles microscopiques qui tiennent compte de l'hétérogénéité du réseau : un modèle déterministe à équations différentielles et un modèle stochastique. Nous avons simulé la partie économique de ces modèles en nous appuyant sur une base de données de bovins. Cela nous a permis de trouver des propriétés intéressantes sur les valeurs à l'équilibre. Enfin, nous avons pu proposer des leviers d'action pour contrer la propagation de l'épidémie.

Cependant, il y a des pistes d'approfondissement. Tout d'abord, nous sommes dans cette étude restés assez près de la base de données ; en particulier, nous avons conservé pour chaque nœud le voisinage qu'il avait dans la base de données.

Il serait intéressant d'étudier le modèle sur des réseaux générés plus aléatoirement, où le voisinage simulé d'un nœud n'est pas le même que dans la base de données. Cela reviendrait à étudier d'autres coefficients de fractionnement.

Plus généralement, on pourrait mener une étude de sensibilité aux paramètres tels que les productions de référence ou les frictions. Il faudrait également examiner en détail les effets de la dynamique de prix. Ces différentes analyses permettraient de définir un meilleur scénario de référence, qui simule un réseau commercial dynamique de façon réaliste.

Enfin, nous n'avons évalué l'impact de l'épidémie que par des calculs de composante connexe sur le réseau économique agrégé. Il conviendrait de tester la robustesse de cette approximation, et de simuler le modèle complet d'épidémiologie économique et non pas seulement le modèle économique comme nous l'avons fait.

F.6 Compléments

F.6.1 Étude des équilibres pour différents coefficients de fractionnement

Nous avons étudié les valeurs finales de simulation et la mise en place de l'équilibre pour différents modèles de coefficients de fractionnement (et des frictions associées). En plus du modèle utilisé dans le rapport de stage, quatre autres modèles ont été envisagés. Ils sont notés FGbis, FGter, FGqua et FGadapt. Leurs descriptions se trouvent ci-dessous (tous ces modèles imposent aux nœuds le voisinage qu'ils ont dans les données). Toutes les simulations ont été faites grâce au modèle EDO, sur un temps de 15000 jours. Le prix initial est 1, qui est un équilibre, il n'y a donc pas de dynamique de prix.

Modèle utilisé dans le rapport

Dans le rapport, les coefficients de fractionnement f_{ij} et g_{ij} sont proportionnels aux productions de référence δ_j et σ_i. Ces productions sont prises égales aux flux entrants et sortants observés dans la base de données : $\delta_j = \hat{\phi}_j^{\text{in}}$ et $\sigma_i = \hat{\phi}_i^{\text{out}}$. Ainsi les fonctions de fractionnement s'écrivent

$$f_{ij}^{\text{rapport}} = \frac{\delta_j}{\sum_k \delta_k} = \frac{\hat{\phi}_j^{\text{in}}}{\sum_k \hat{\phi}_k^{\text{in}}} \quad \text{et} \quad g_{ij}^{\text{rapport}} = \frac{\sigma_i}{\sum_k \sigma_k} = \frac{\hat{\phi}_i^{\text{out}}}{\sum_k \hat{\phi}_k^{\text{out}}} \tag{F.59}$$

où les sommes sont sur le voisinage observé dans les données.

Les frictions sont calibrées de sorte que le taux de transaction θ_{ij} soit celui observé dans les données $\hat{\theta}_{ij}$. On prend donc

$$\kappa_{ij} = \frac{1}{\hat{\theta}_{ij}} \min\left(f_{ij}\hat{\phi}_i^{\text{out}}, g_{ij}\hat{\phi}_j^{\text{in}}\right). \tag{F.60}$$

Modèle FGbis

Dans le modèle FGbis, afin de se rapprocher des données, les coefficients de fractionnements f_{ij} et g_{ij} sont pris proportionnels aux flux $\hat{\phi}_{ij}$ observés dans les données :

$$f_{ij}^{\text{bis}} = \frac{\hat{\phi}_{ij}}{\sum_k \hat{\phi}_{ik}} \quad \text{et} \quad g_{ij}^{\text{bis}} = \frac{\hat{\phi}_{ij}}{\sum_k \hat{\phi}_{kj}}. \tag{F.61}$$

Les frictions κ sont prises égales aux quantités moyennes échangées dans les données : $\kappa_{ij} = \hat{q}_{ij}$.

Que vaut alors le taux de transaction θ_{ij} ? Un calcul montre que si $p = 1$, alors

$$\theta_{ij} = \frac{1}{\kappa_{ij}} \min\left(f_{ij}\hat{\phi}_i^{\text{out}}, g_{ij}\hat{\phi}_j^{\text{in}}\right) = \frac{\hat{\phi}_{ij}}{\hat{q}_{ij}} = \hat{\theta}_{ij}.$$

Ainsi, en l'absence de dynamique de prix, le modèle FGbis et le modèle du rapport sont rigoureusement identiques.

On considère, ici ainsi que dans toute la suite, quatre scénarios de stocks initiaux, dont on rappelle

la définition :

- A : S et D nuls ;
- B : S et D égaux à la taille de l'exploitation (nbpres moyen) ;
- C : S et D égaux au flux observé annuel : $S_i = D_i = (1 \text{ an}) \times (\hat{\phi}_i^{\text{in}} + \hat{\phi}_i^{\text{out}})$;
- D : S et D asymétriques : $S_i = \text{constante} \times \hat{\phi}_i^{\text{out}}$, $D_i = \text{constante} \times \hat{\phi}_i^{\text{in}}$.

La figure F.9 montre que quels que soient les stocks initiaux, on atteint approximativement le même équilibre (il peut y avoir quelques disparités car dans les scénarios B et C, on attribue initialement des stocks non nuls de demande à des offreurs stricts ; ces agents ne pouvant pas acheter, leur stock de demande reste constant à leur valeur initiale ; néanmoins, ces points ne sont pas représentés sur la figure puisque l'échelle logarithmique est utilisée et que le stock de demande final de ces points dans le scénario A est nul).

La figure F.10 montre cependant que l'équilibre n'est pas atteint à la même vitesse. Il est par exemple atteint beaucoup plus lentement dans le scénario B que dans le scénario A. Cela peut s'expliquer par le fait que les valeurs d'équilibre de S et D sont plus proches de zéro (ce qui correspond à l'état initial du scénario A) que de nbpres (ce qui correspond à l'état initial du scénario B), comme on peut le constater sur la figure 7 du rapport.

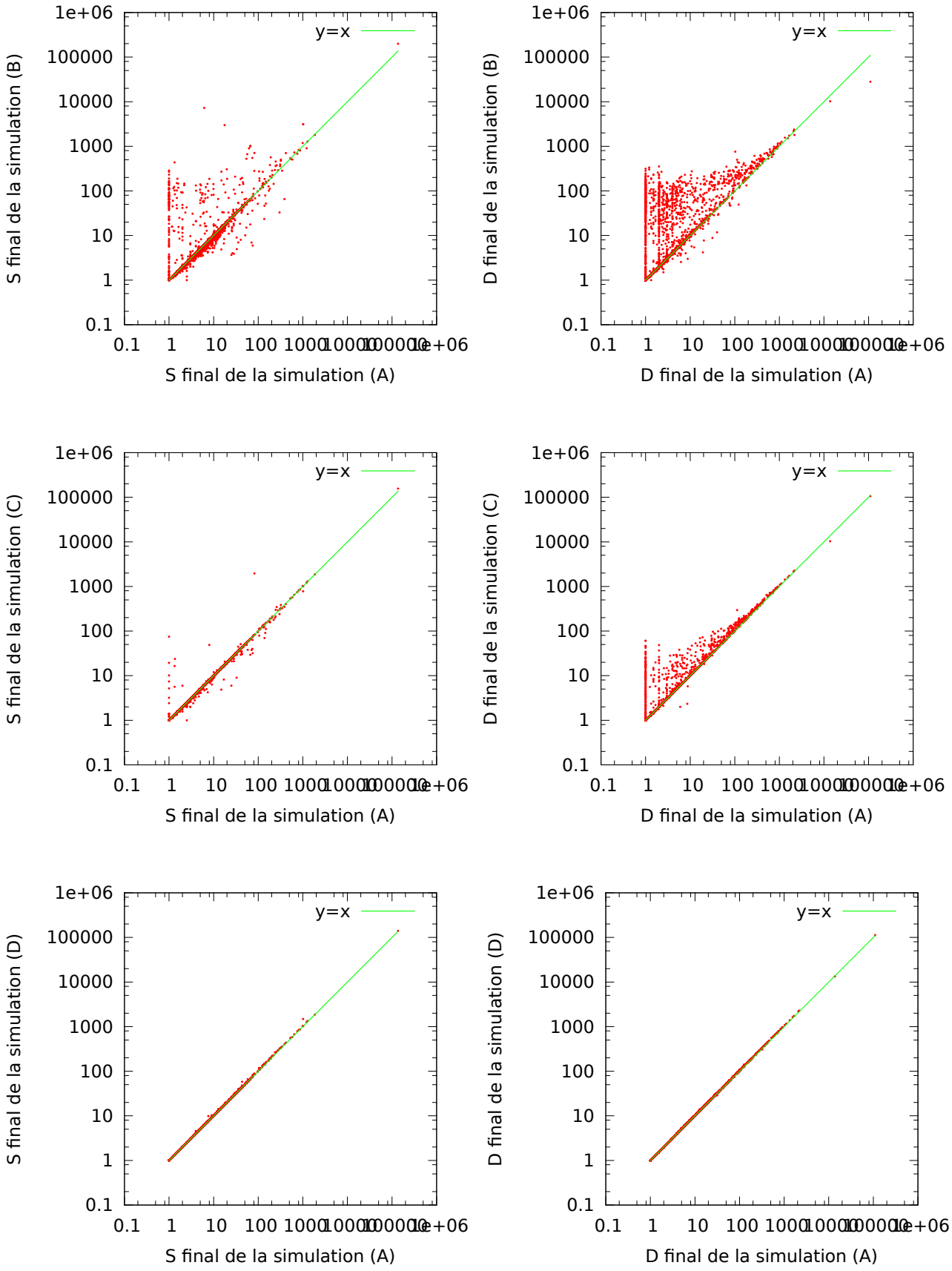
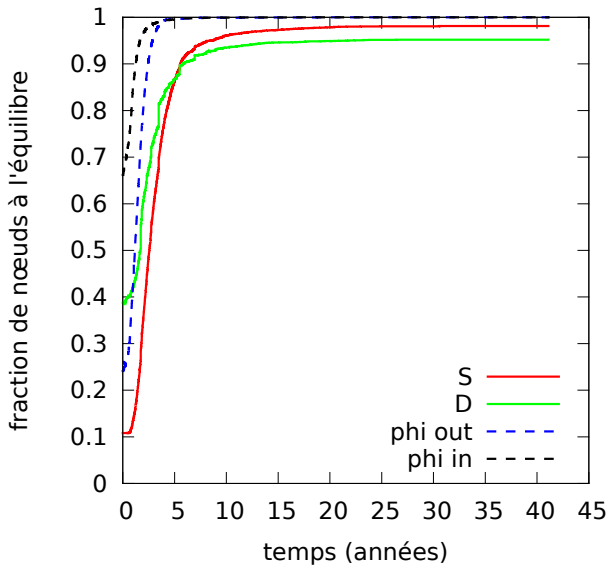
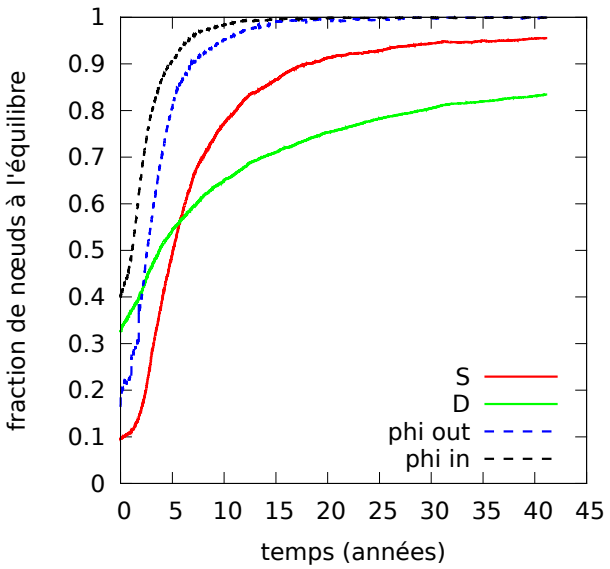


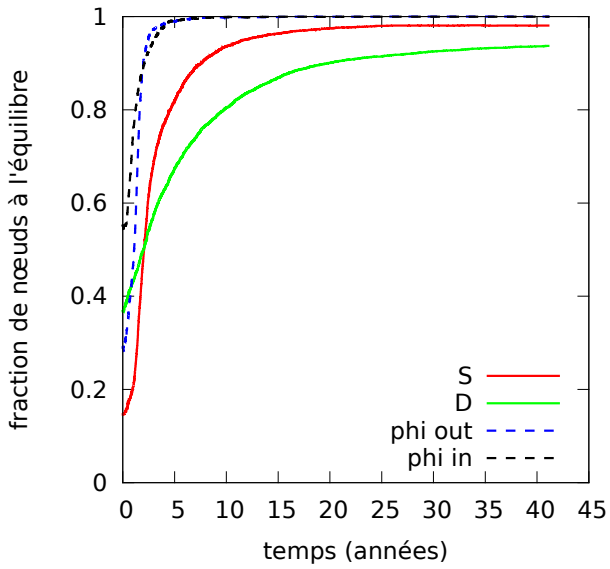
FIGURE F.9 – Comparaison des valeurs finales de stocks S_i et D_i entre les différents scénarios dans le modèle FGbis



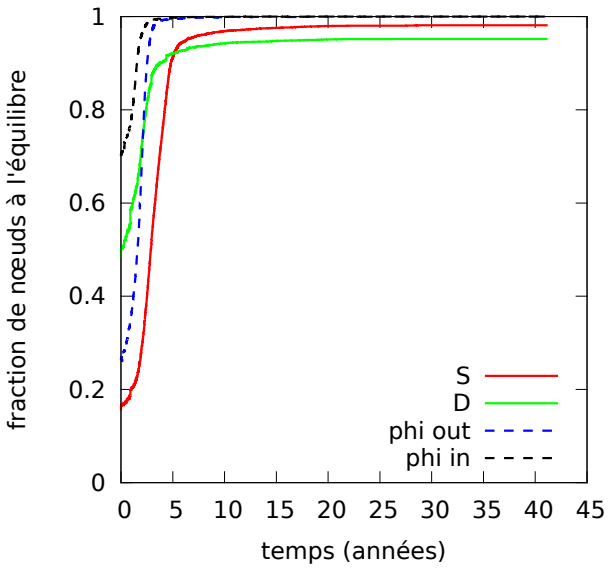
(a) Scénario A



(b) Scénario B



(c) Scénario C



(d) Scénario D

FIGURE F.10 – Évolution temporelle de l'équilibre pour différents scénarios de stocks initiaux dans le modèle FGbis

Modèle FGter

Le modèle FGter vise à être plus général. Les coefficients de fractionnement dépendent de paramètres α et β et sont de la forme

$$f_{ij}^{\text{ter}}(\alpha_i) = \frac{\delta_j^{\alpha_i}}{\sum_k \delta_k^{\alpha_i}} \quad \text{et} \quad g_{ij}^{\text{ter}}(\beta_j) = \frac{\sigma_i^{\beta_j}}{\sum_k \sigma_k^{\beta_j}} \tag{F.62}$$

où les sommes sont sur le voisinage observé dans les données. On a toujours $\delta_j = \hat{\phi}_j^{\text{in}}$ et $\sigma_i = \hat{\phi}_i^{\text{out}}$.

On prend encore $\kappa_{ij} = \hat{q}_{ij}$.

Les paramètres α et β choisis sont ceux qui permettent de refléter le mieux les données. En pratique, on minimise donc la somme des carrés des écarts entre f^{ter} et f^{bis} :

$$\alpha_i = \underset{\alpha}{\operatorname{argmin}} \sum_j \left(f_{ij}^{\text{ter}}(\alpha) - f_{ij}^{\text{bis}} \right)^2 \tag{F.63}$$

où la somme se fait sur les voisins sortants de i. De même

$$\beta_j = \underset{\beta}{\operatorname{argmin}} \sum_i \left(g_{ij}^{\text{ter}}(\beta) - g_{ij}^{\text{bis}} \right)^2 \tag{F.64}$$

où la somme se fait sur les voisins entrants de j.

La figure F.11 montre la corrélation des α et β obtenus avec la polarité et le nombre de bêtes échangées.

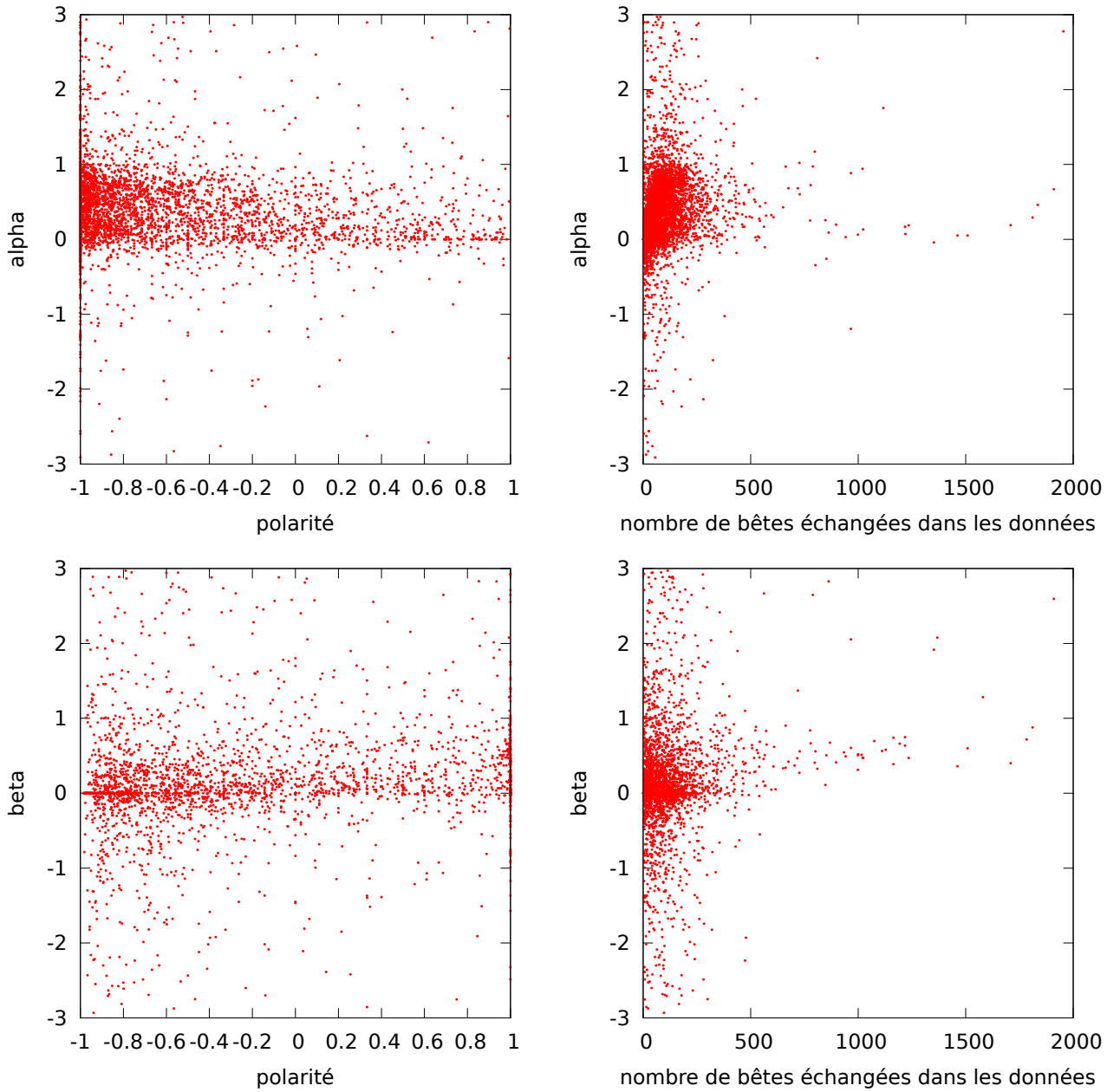


FIGURE F.11 – Lien entre α, β, la polarité et le nombre de bêtes échangées

Modèle FGqua

Afin d'évaluer l'intérêt d'ajuster aux données des paramètres α et β, le modèle FGqua reprend les coefficients du modèle FGter, sauf que tous les α et β sont pris égaux à 1 :

$$f_{ij}^{\text{qua}} = \frac{\delta_j}{\sum_k \delta_k} \quad \text{et} \quad g_{ij}^{\text{qua}} = \frac{\sigma_i}{\sum_k \sigma_k}. \tag{F.65}$$

On retrouve les coefficients de fractionnement du rapport. La différence de ce modèle avec celui du rapport se fait dans les frictions : on prend $\kappa_{ij} = \hat{q}_{ij}$ comme dans FGbis et FGter.

Modèle FGadapt

Le dernier modèle envisagé, FGadapt, est une tentative de maximiser les échanges étant donné les productions individuelles (σ et δ) et les voisinages. On utilise pour cela un procédé itératif. À chaque itération, les agents répartissent leur production en fonction de ce qu'ils ont reçu de leurs voisins à l'itération précédente :

$$f_{ij}^{[n+1]} = \frac{g_{ij}^{[n]}\delta_j}{\sum_k g_{ik}^{[n]}\delta_k} \quad \text{et} \quad g_{ij}^{[n+1]} = \frac{f_{ij}^{[n]}\sigma_i}{\sum_k f_{kj}^{[n]}\sigma_k} \tag{F.66}$$

où les sommes se font sur les voisinages observés dans les données. On prend comme état initial le partitionnement uniforme : $f_{ij}^{[0]} = \dfrac{1}{N_i^{\text{voisins out}}}$ et $g_{ij}^{[0]} = \dfrac{1}{N_j^{\text{voisins in}}}$.

Le modèle FGadapt est finalement défini par $f_{ij}^{\text{adapt}} = f_{ij}^{[3000]}$, $g_{ij}^{\text{adapt}} = g_{ij}^{[3000]}$ et $\kappa_{ij} = \hat{q}_{ij}$.

La convergence de ce procédé est illustrée en figure F.12 par la variation relative (moyenne et maximale) des coefficients f et g au cours de chaque itération.

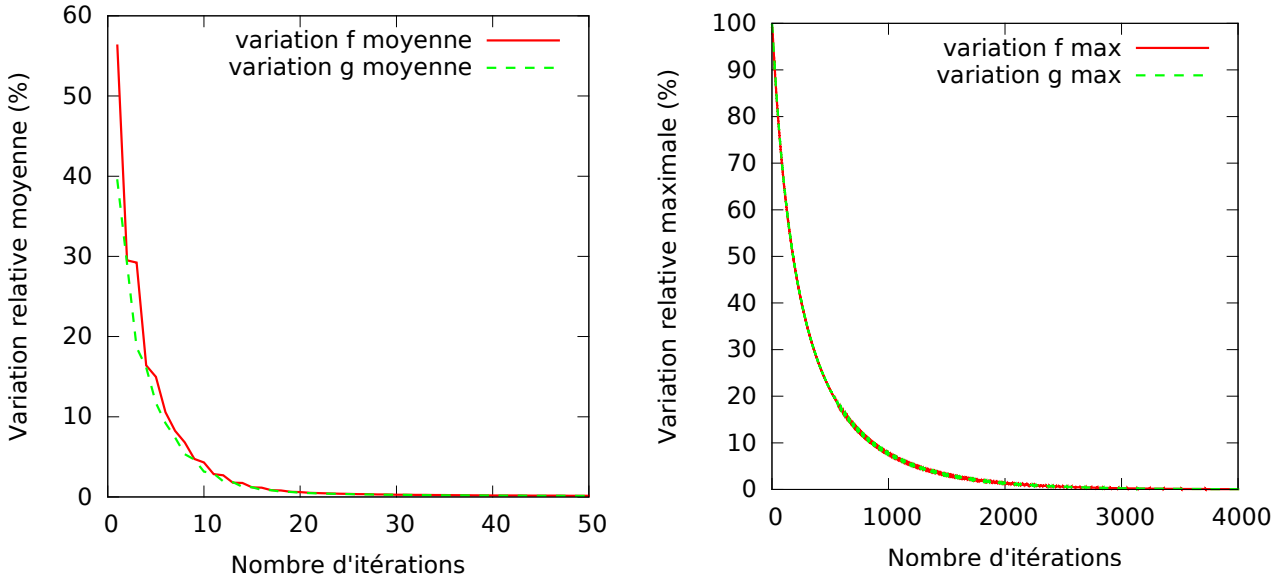


FIGURE F.12 – Convergence du procédé itératif

Comparaison des quatre modèles

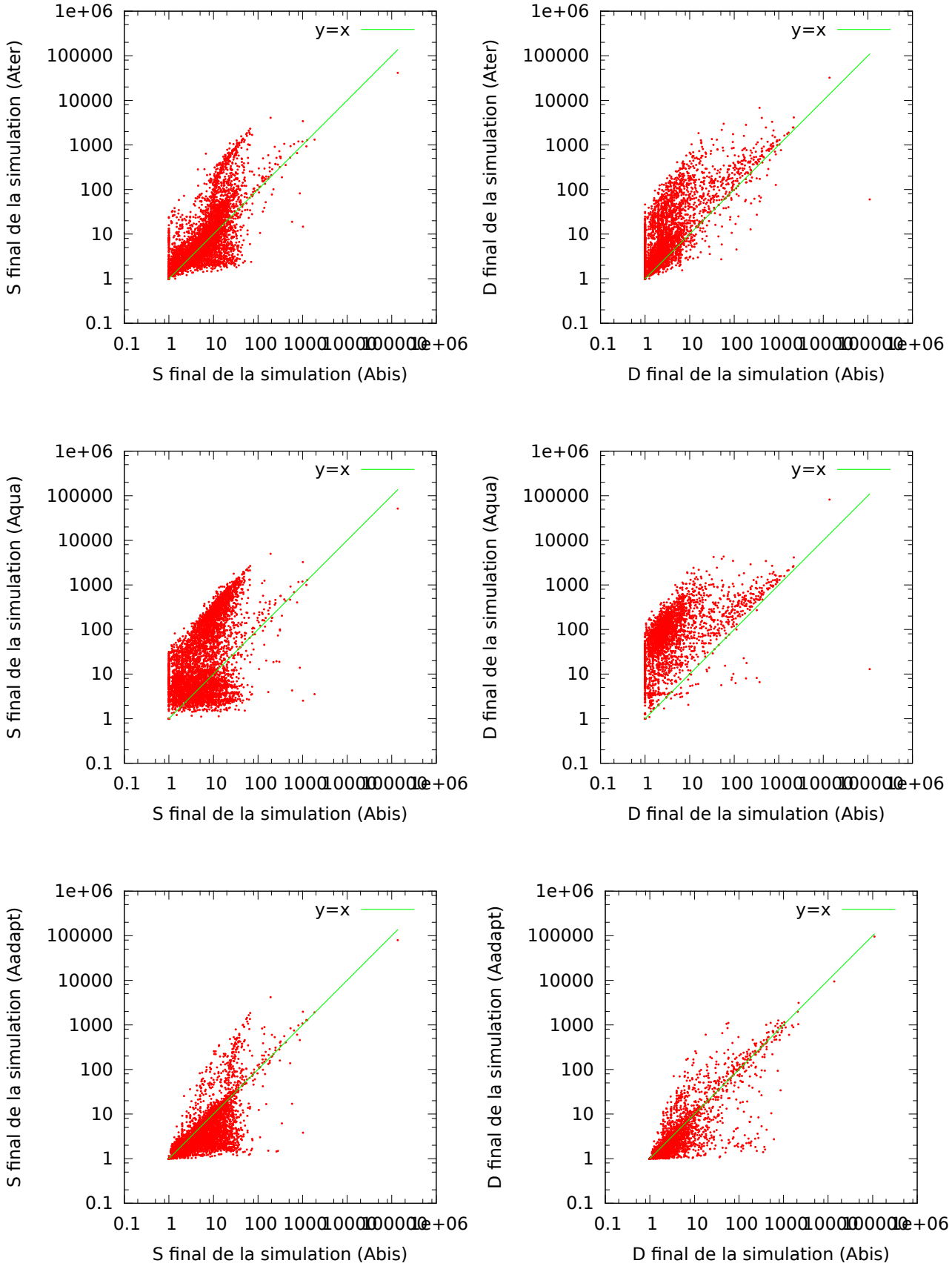
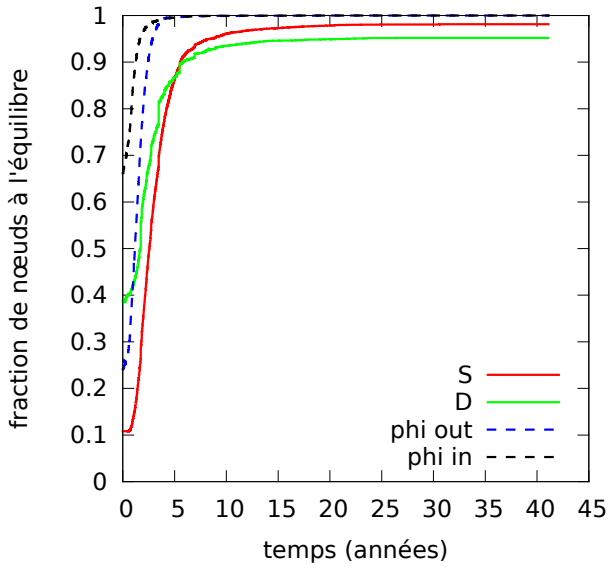
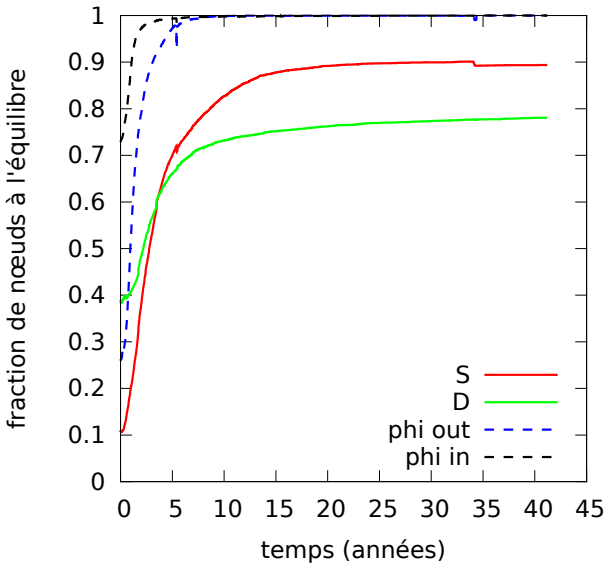


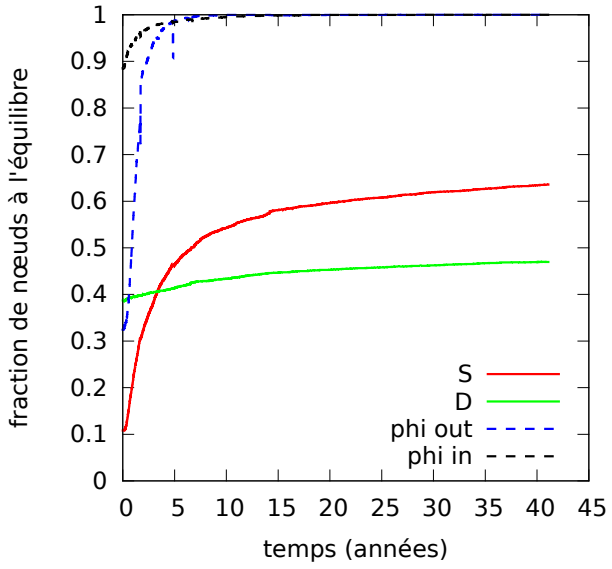
FIGURE F.13 – Comparaison des valeurs finales de stocks S_i et D_i entre différents modèles avec des stocks initiaux nuls (scénario A)



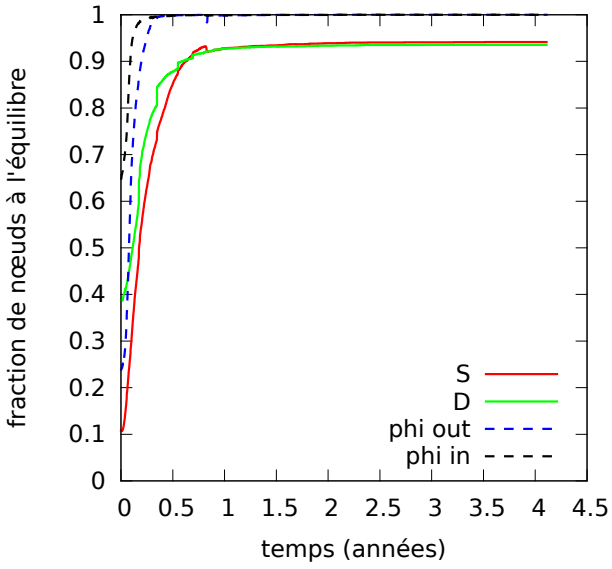
(a) Modèle FGbis



(b) Modèle FGter

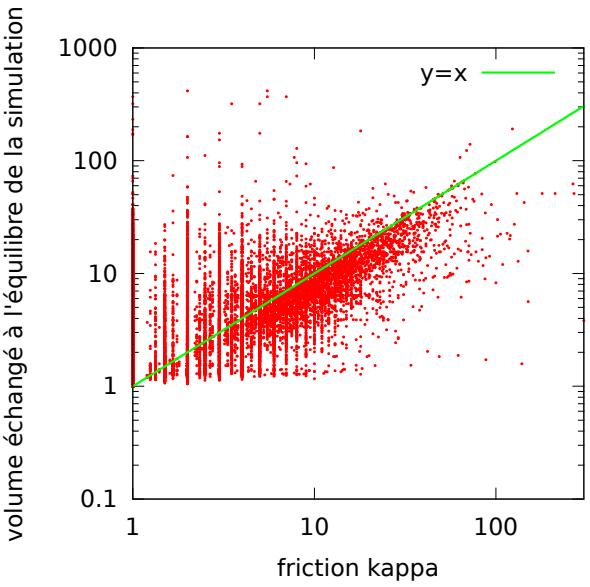


(c) Modèle FGqua

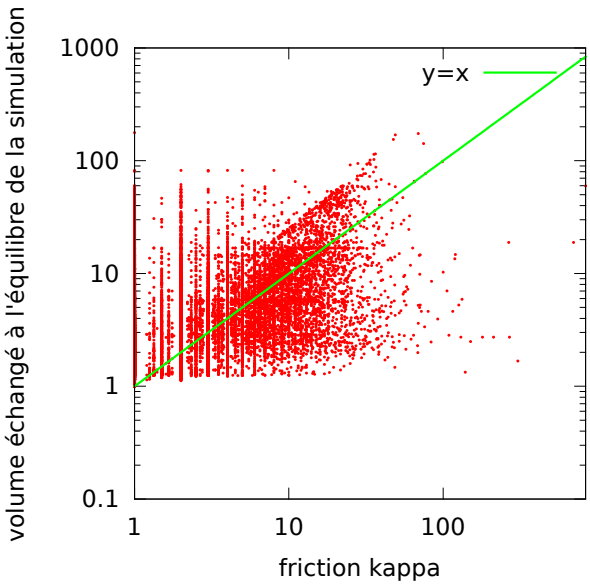


(d) Modèle FGadapt

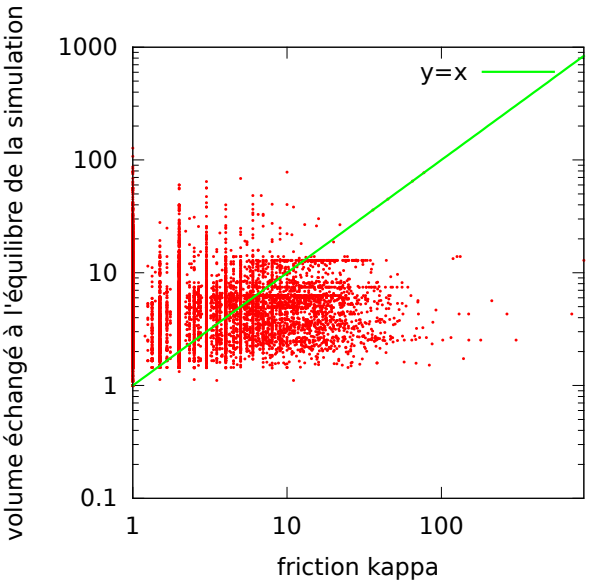
FIGURE F.14 – Évolution temporelle de l'équilibre pour différents modèles avec des stocks initiaux nuls (scénario A)



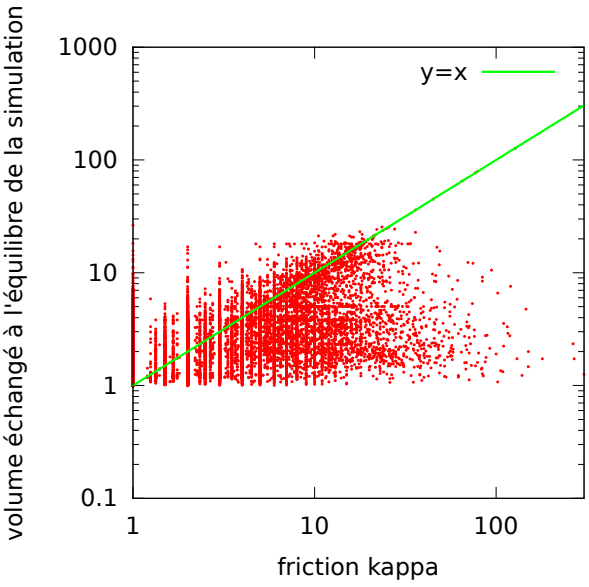
(a) Modèle FGbis



(b) Modèle FGter



(c) Modèle FGqua



(d) Modèle FGadapt

FIGURE F.15 – Valeurs d'équilibre de q_{ij} en fonction de la friction κ_{ij} (qui est égale à $\hat{q}_{ij}$) pour différents modèles avec des stocks initiaux nuls (scénario A). Seules les arêtes à l'équilibre sont représentées (c'est-à-dire celles telles que S_i et D_j sont tous les deux à l'équilibre).

En figures F.13, F.14 et F.15 se trouvent différentes comparaisons des modèles évoqués précédemment. Voici les coefficients R^2 ajustés des régressions linéaires des valeurs finales de q_{ij} en fonction de $\hat{q}_i j$:

	y=ax+b	y=ax
FGbis	0.540	0.564
FGter	0.372	0.383
FGqua	0.059	0.087
FGadapt	0.515	0.536

On voit que le modèle FGbis est celui qui est le plus proche des données (ce qui est cohérent avec la façon dont il a été construit). Le modèle FGter s'éloigne un peu plus des données, et le modèle FGqua s'en éloigne encore plus, ce qui est cohérent également. Enfin, le modèle FGadapt semble bien approcher les données, mais pas tout à fait aussi bien que FGbis.

F.6.2 Nouveau modèle de quantité échangée

Afin d'obtenir des simulations plus proches des données, on envisage un nouveau modèle de quantité échangée : $q_{ij} = \min(f_{ij}S_i, g_{ij}D_j)$. Ce nouveau modèle de quantité échangée est noté q2, le modèle classique est noté q1. Les simulations de q2 ont été faites pour les modèles FGbis et FGadapt, avec des stocks initiaux nuls (scénario A).

Différents résultats sont présentés en figures F.16, F.17 et F.18. Les coefficients R^2 ajustés des régressions linéaires des valeurs finales de q_{ij} en fonction de $\hat{q}_i j$ sont les suivants :

	y=ax+b	y=ax
FGbis, q2	0.971	0.972
FGadapt, q2	0.446	0.465

On voit ainsi que la correspondance est très bonne pour FGbis (de plus, le coefficient a trouvé par la régression linéaire est dans ce cas très proche de 1).

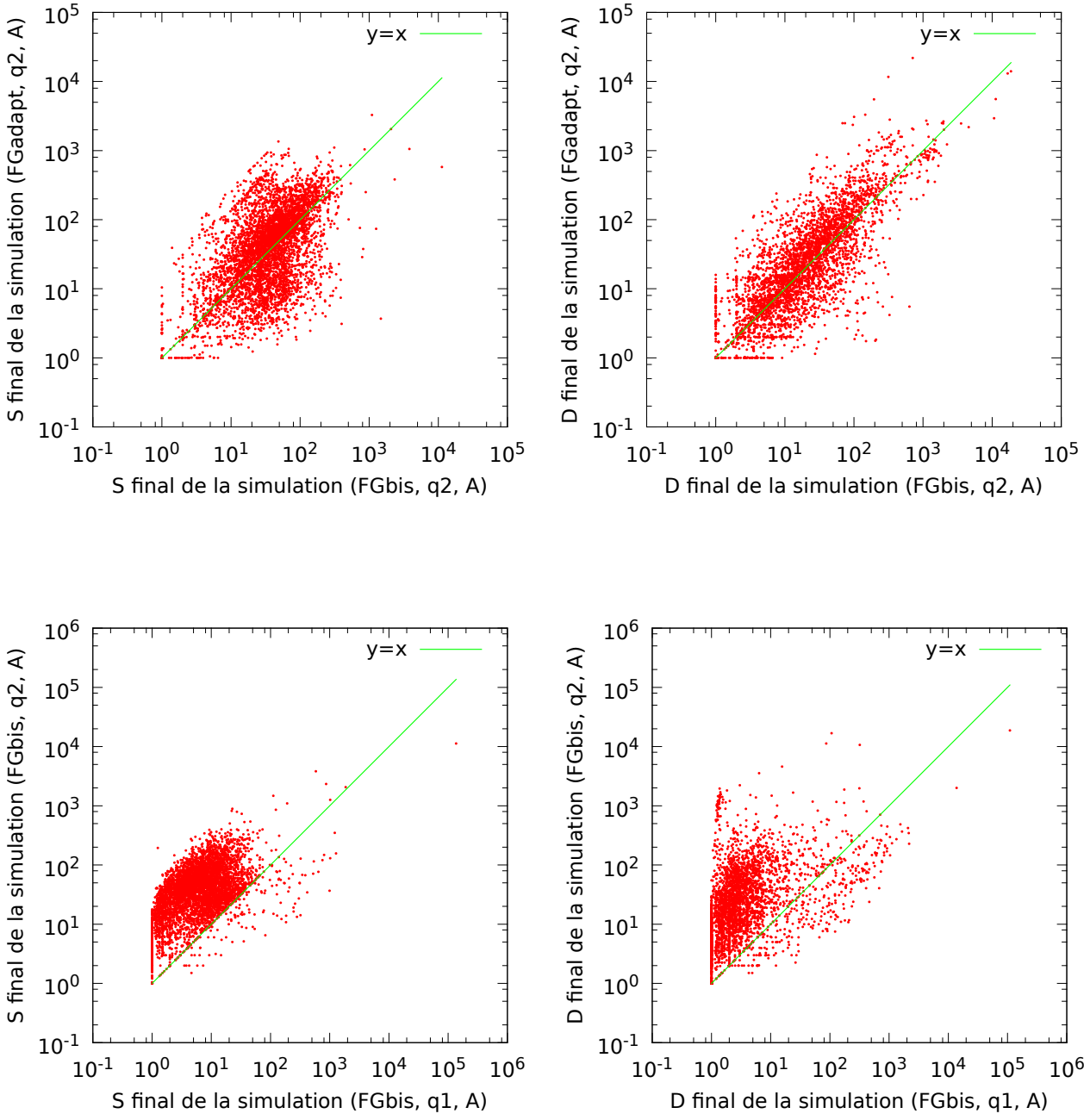
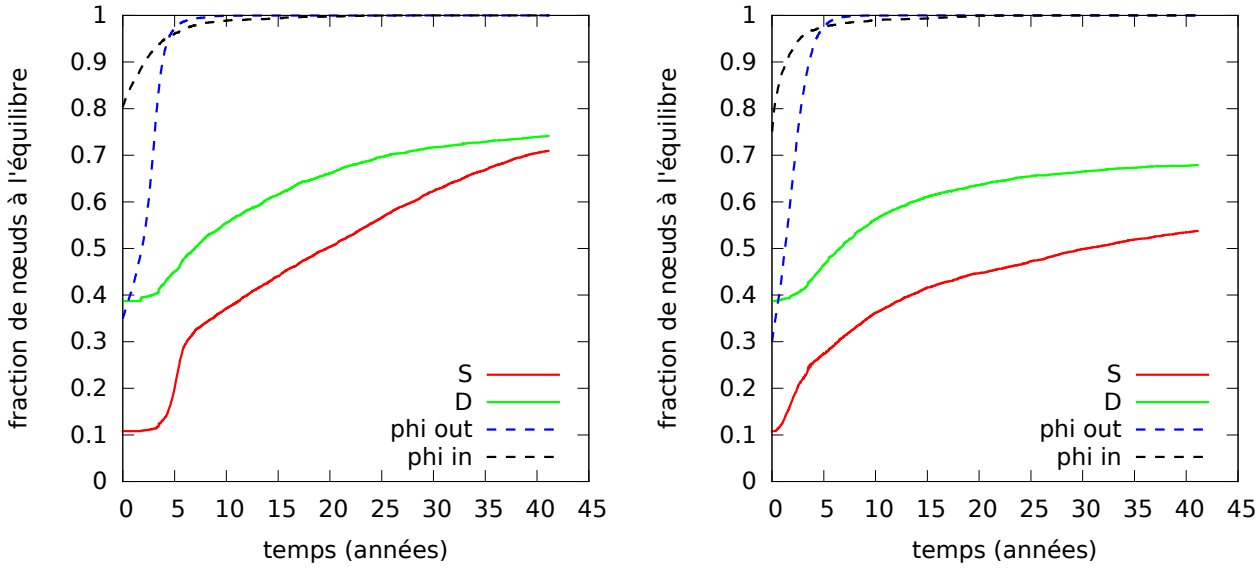


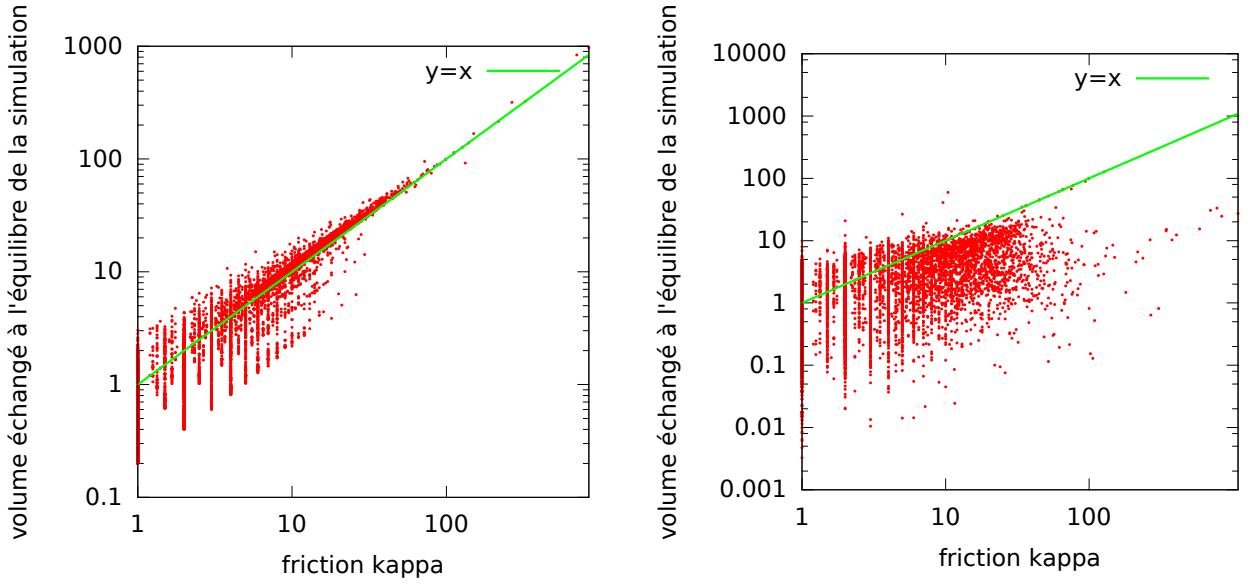
FIGURE F.16 – Comparaison des valeurs finales de stocks S_i et D_i entre différents modèles de f, g et q avec des stocks initiaux nuls (scénario A)



(a) Modèle FGbis, q2

(b) Modèle FGadapt, q2

FIGURE F.17 – Évolution temporelle de l'équilibre pour différents modèles avec des stocks initiaux nuls (scénario A)



(a) Modèle FGbis, q2

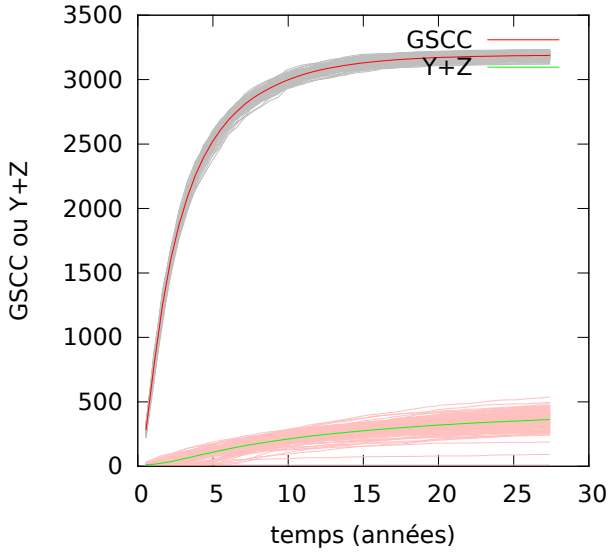
(b) Modèle FGadapt, q2

FIGURE F.18 – Valeurs d'équilibre de q_{ij} en fonction de la friction κ_{ij} (qui est égale à $\hat{q}_{ij}$) pour différents modèles avec des stocks initiaux nuls (scénario A). Seules les arêtes à l'équilibre sont représentées (c'est-à-dire celles telles que S_i et D_j sont tous les deux à l'équilibre).

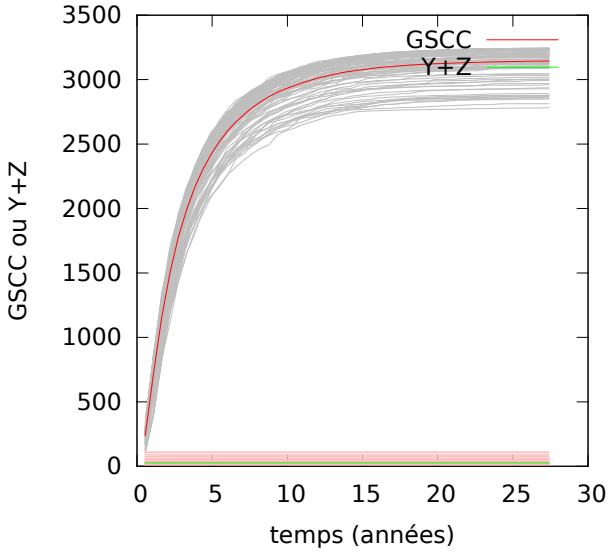
F.6.3 Percolation épi-éco

Afin de réaliser des études de percolation, on a calculé la taille moyenne de la GSCC du réseau agrégé et la valeur moyenne de $Y + Z$ au cours du temps dans le modèle épi-éco. Les stocks initiaux sont les valeurs finales d'une simulation sans épidémie sur une durée de 5000 jours. Le modèle de coefficients de fractionnement et de frictions utilisé est le modèle FGbis. Les coefficients de fractionnement s'adaptent à la présence ou à l'absence de certains voisins. À chaque expérience, 100 simulations ont été réalisées. Le coefficient d'ajustement des prix μ est choisi à 10^{-6}. On suppose qu'il n'y a pas de réintroduction : $\nu = 0$.

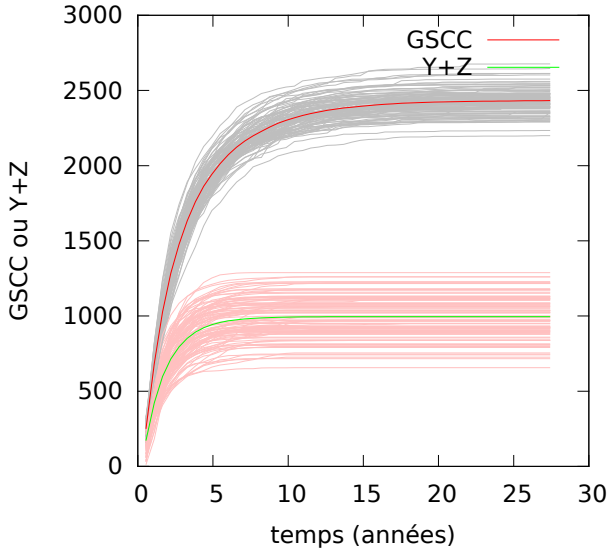
Différents couples $(\omega, \frac{1}{\gamma})$ ont été testés, ainsi que différentes dilatations des frictions κ. Les résultats sont présentés en figures F.19, F.20 et F.21.



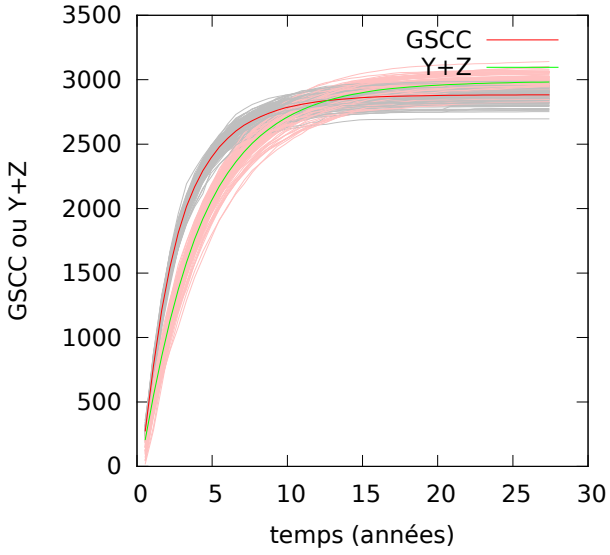
(a) $\omega = 0.01$, $\frac{1}{\gamma} = 10$ ans



(b) $\omega = 1$, $\frac{1}{\gamma} = 1$ mois

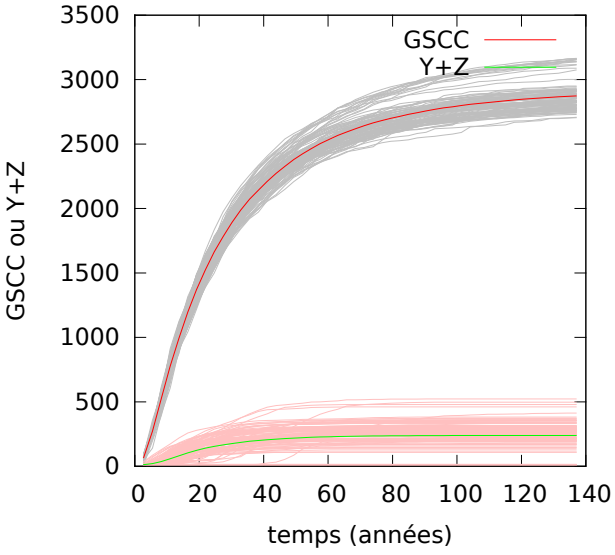


(c) $\omega = 1$, $\frac{1}{\gamma} = 1$ an

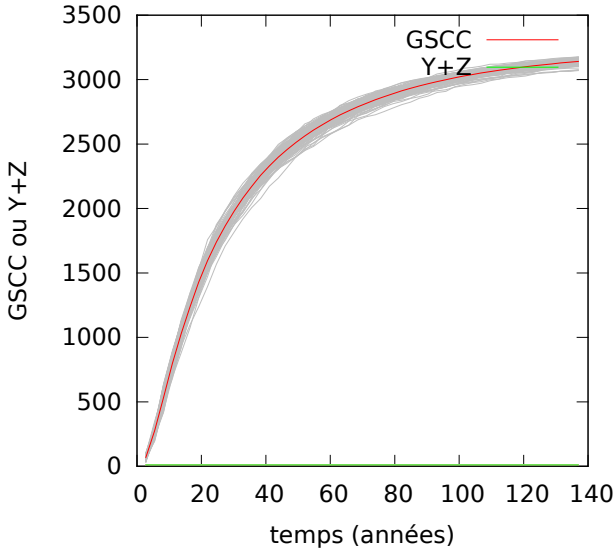


(d) $\omega = 1$, $\frac{1}{\gamma} = 10$ ans

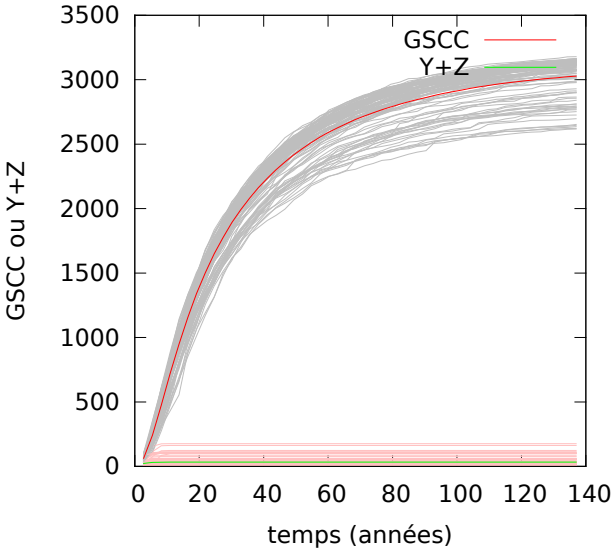
FIGURE F.19 – Évolution de la taille de la GSCC du réseau agrégé simulé et de $Y + Z$ au cours du temps, sans dilatation des frictions



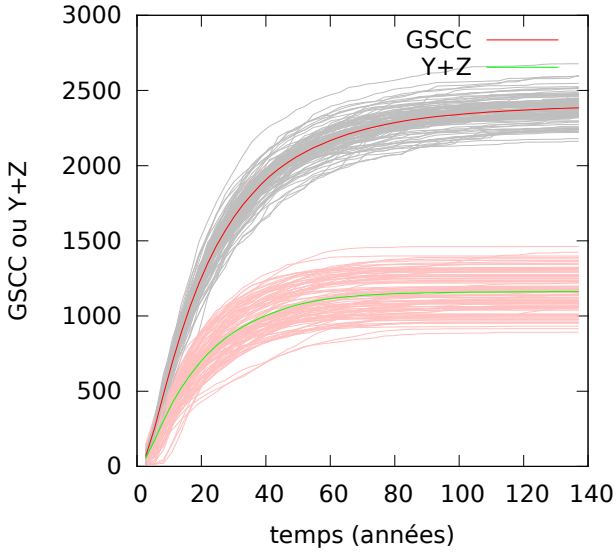
(a) $\omega = 0.01$, $\frac{1}{\gamma} = 10$ ans



(b) $\omega = 1$, $\frac{1}{\gamma} = 1$ mois

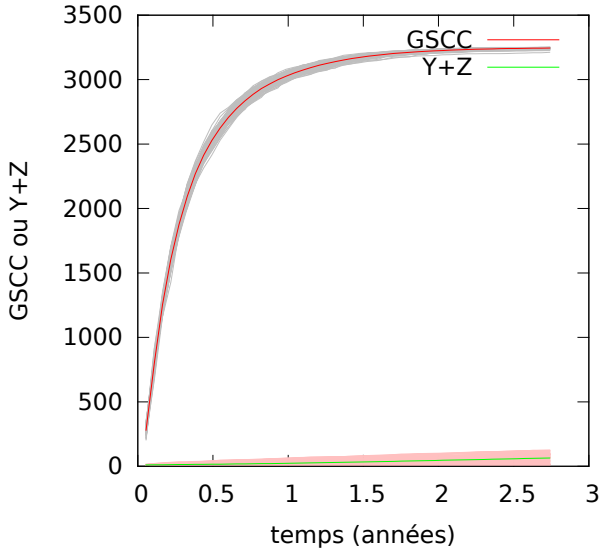


(c) $\omega = 1$, $\frac{1}{\gamma} = 1$ an

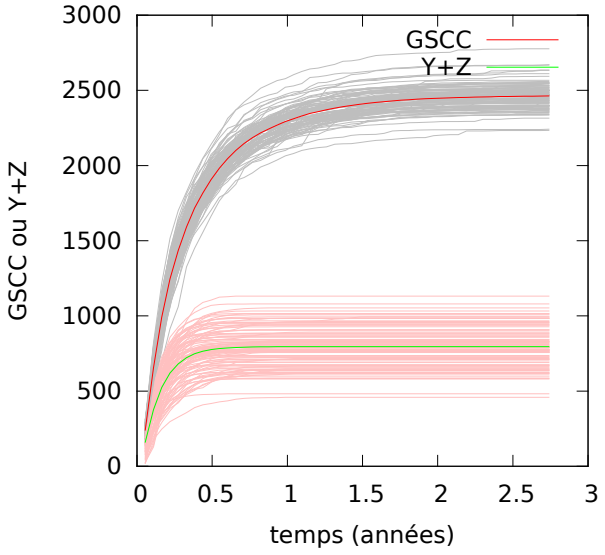


(d) $\omega = 1$, $\frac{1}{\gamma} = 10$ ans

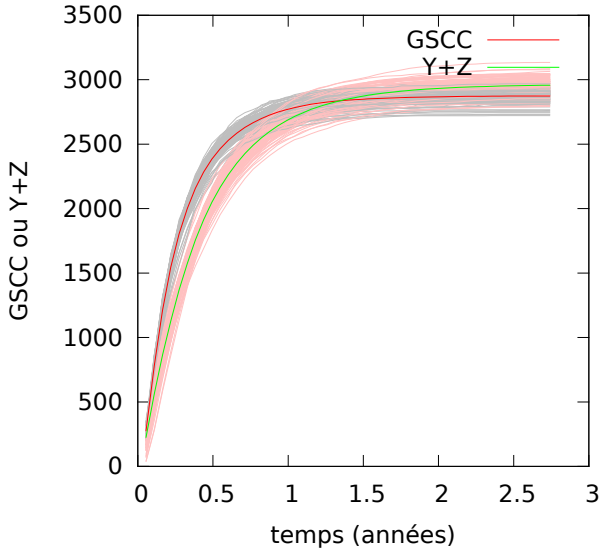
FIGURE F.20 – Évolution de la taille de la GSCC du réseau agrégé simulé et de $Y + Z$ au cours du temps, avec des frictions dilatées d'un facteur 10



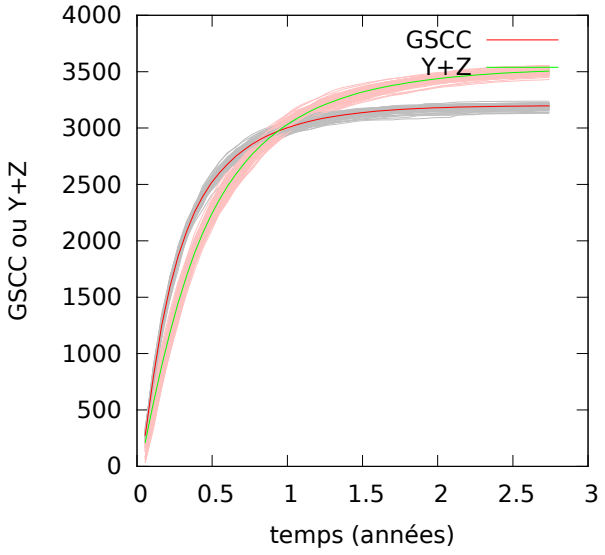
(a) $\omega = 0.01$, $\frac{1}{\gamma} = 10$ ans



(b) $\omega = 1$, $\frac{1}{\gamma} = 1$ mois



(c) $\omega = 1$, $\frac{1}{\gamma} = 1$ an



(d) $\omega = 1$, $\frac{1}{\gamma} = 10$ ans

FIGURE F.21 – Évolution de la taille de la GSCC du réseau agrégé simulé et de $Y + Z$ au cours du temps, avec des frictions dilatées d'un facteur 0.1

Résumé

Les échanges entre agents tels que les individus ou les entreprises couvrent de nombreux besoins tels que la reproduction et la recherche de profits, mais peuvent dans le même temps faciliter la transmission des maladies infectieuses. En conséquence, les épidémies sont susceptibles de causer des dommages au sein des populations biologiques et d'altérer ces mêmes échanges qui véhiculent les infections. Les modèles épidémiologiques intègrent de plus en plus les comportements d'aversion au risque en réaction aux épidémies. Ces comportements induisent une réduction des contacts infectieux entre agents. Toutefois, les comportements qui sous-tendent les échanges ne se réduisent pas à l'aversion au risque. Les échanges sont conditionnés par des motivations différentes et sont contraints en raison de leurs coûts et du caractère limité des ressources. Face à une émergence épidémique, les comportements adaptatifs se traduisent par des ajustements dans les actions des agents qui peuvent augmenter ou réduire leur exposition au risque d'infection. De multiples mécanismes adaptatifs liés à des processus sociaux et économiques contribuent à expliquer ces ajustements. Les contraintes d'interaction limitent la capacité des agents à échanger de diverses manières. Elles peuvent réduire la capacité d'un agent à interagir avec ses alter ego (contrainte de creux), le taux d'échange (contrainte de pondération), la direction des échanges (contrainte de direction) ou la fréquence de rencontre entre agents (contrainte de friction). Si la contrainte de creux a été largement étudiée, les implications épidémiologiques des trois autres contraintes restent mal comprises. Dans cette recherche, nous combinons analyses empiriques et explorations de modèles mathématiques pour étudier l'influence des contraintes d'interaction et des comportements adaptatifs sur la dynamique conjointe des échanges et de l'infection. Nous pouvons ainsi proposer des politiques de prévention et de maîtrise des épidémies véhiculées par les échanges. Nos études de cas incluent la dynamique des infections sexuellement transmissibles dans des réseaux de contacts sexuels et les épidémies propagées dans des marchés d'échanges d'animaux ou de plantes. Dans un premier temps, nous montrons que la superposition de contraintes d'interaction engendre une grande diversité de structures d'échanges et de dynamiques épidémiques. Nous identifions des conditions analytiques pour lesquelles les contraintes de pondération et de direction limitent la probabilité d'émergence et la sévérité des maladies infectieuses, et traduisons ces conditions théoriques en matière de politiques de santé. Ces résultats sont obtenus en supposant que les agents sont passifs. Dans un second temps, nous prenons également en considération les comportements adaptatifs rencontrés dans les marchés qui véhiculent des infections, et suggérons que la dynamique conjointe de l'infection et des échanges est limitée par la friction marchande. La contrainte de friction engendre un compromis entre fréquence et intensité des transactions commerciales, et peut amoindrir les épidémies de manière plus importante que l'aversion au risque. Nous évoquons finalement des mesures pratiques pour limiter la transmission des maladies infectieuses dans les marchés tout en minimisant les effets délétères sur le commerce. Notre thèse démontre l'importance d'adopter des approches interdisciplinaires pour relever les défis des émergences épidémiques imputables aux échanges. Les idées et outils développés peuvent être transposés à d'autres systèmes, par exemple en écologie ou en économie.

Mots clés : épidémiologie économique ; prévention des maladies infectieuses ; réponse comportementale ; réseaux complexes ; R_0

Abstract

Exchanges among agents, typically individuals or companies, fulfil various needs such as reproduction and economic profit, but can also support infectious disease transmission, impacting biological populations and potentially altering the disease-conducive exchanges. Epidemiological models increasingly account for reductions in infectious contact, such as risk-aversion behaviour in response to pathogen outbreaks. However, behavioural responses in exchange dynamics are not limited to risk-aversion ; they are driven by different motivation and are constrained because resources are limited and exchanges are costly. Adaptive behaviour refers to change in agent behaviour, and potentially in exposure to risk, in response to disruption such as disease outbreak ; the change may be through adaptation of social or economic mechanisms. Interaction constraints limit in different ways the capacity of agents to interact. They can limit the interaction of a given agent with everybody else (sparseness constraint), the rate and direction of exchanges (weighting and directional constraints), or the rate of encounter between agents (frictional constraint). While the sparseness constraint has been exhaustively studied, the epidemiological consequences of the three other constraints are poorly understood. Here, we use a combination of empirical analyses and mathematical modelling approaches to explore the influence of interaction constraints and adaptive behaviour on the combined dynamics of exchanges and infection. We can hence suggest relevant policies to prevent and mitigate exchange-driven epidemics. The examples investigated are sexually transmitted infection dynamics in sexual-contact networks, and epidemics in markets of animal livestock or ornamental plants. First, we show that differing patterns in agent contact structures and in epidemic dynamics arise when the networks are subject to combinations of interaction constraints. We identify analytical conditions when weighting and directional constraints limit the occurrence and severity of epidemic outbreaks, and translate these threshold conditions into disease-control strategies. These results hold in the case when agents are passive. Second, we account for adaptive behaviour encountered in markets that propagate infections and propose that the joint dynamics of markets and disease spread are limited by trade friction. This specific constraint creates a trade-off between the frequency and intensity of market transactions that can influence epidemics more strongly than risk-aversion. We finally suggest policy for limiting disease contagion in markets and minimise its adverse impact on trade. Our work demonstrates that the integration of differing standpoints at the crossroad of natural and social sciences is important in tackling the challenges posed by the emergence of exchange-driven epidemics. We believe our general approach can be transposed to other systems where agents exhibit adaptive behaviour and face interaction constraints, e.g. in ecology or in economics.

Keywords : behavioural response ; complex networks ; infectious disease prevention ; economic epidemiology ; R_0