Aggregate fluctuations and market frictions:The role of firm and job flows

Sophie Osotimehin

Université Paris 1 - Panthéon Sorbonne
U.F.R de Sciences Economiques

Année 2010-2011

Thèse pour le Doctorat de Sciences Economiques

soutenue publiquement par

Sophie Osotimehin

le 19 septembre 2011

Aggregate fluctuations and market frictions: The role of firm and job flows

Directeur de thèse

Monsieur Jean-Olivier Hairault - Professeur à l'Université Paris 1 Panthéon-Sorbonne

Jury

Monsieur Paul Beaudry - Professeur à *British Columbia University*

Monsieur Jean-Olivier Hairault - Professeur à l'Université Paris 1 Panthéon-Sorbonne

Monsieur Jean Imbs - Directeur de recherche CNRS

Monsieur Guy Laroque - Professeur à *University College London*

Monsieur Gianluca Violante - Professeur associé à *New York University*

Remerciements

Je tiens tout d'abord à remercier le professeur Jean-Olivier Hairault pour m'avoir guidée et encouragée tout au long de cette thèse. Cette thèse a considérablement bénéficié de ses critiques et de ses conseils avisés. Je voudrais également lui exprimer ma gratitude de m'avoir associée à un de ses projets de recherche dès le début de ma thèse, me permettant ainsi de me lancer plus facilement dans le travail de recherche.

Je remercie également les professeurs Paul Beaudry, Gianluca Violante, Jean Imbs et Guy Laroque d'avoir accepté de faire partie de mon jury. Leurs remarques me seront très utiles pour améliorer la qualité de ce travail. Je voudrais remercier tout particulièrement Jean Imbs pour ses judicieuses suggestions. Je souhaite également remercier Guy Laroque de m'avoir accueillie au sein du laboratoire de macroéconomie du CREST. Ses encouragements et son soutien m'ont été d'une aide inestimable. L'exigence, ainsi que les nombreux conseils et suggestions de Guy ont permis une amélioration substantielle de cette thèse.

Un chapitre de cette thèse a été réalisé en collaboration avec Jean-Olivier Hairault et François Langot, et un autre avec Francesco Pappadà. Je les remercie tous trois chaleureusement, travailler avec eux fut une expérience très enrichissante.

Cette thèse doit également beaucoup à l'environnement de recherche stimulant et convivial qu'est le CREST. Un grand merci aux membres du CREST qui rendent ce lieu si chaleureux. Je voudrais remercier en particulier Etienne pour sa disponibilité et ses nombreux conseils. Je remercie également Pierre et Jean-Michel pour leurs remarques constructives. Les échanges que j'ai eus avec Isabelle m'ont été très utiles dans le développement du dernier papier, je l'en remercie. Je remercie chaleureusement Julien et Céline pour leurs relectures attentives. Merci à Raul, Marion et Claire de m'avoir initiée au monde obscur du gros système. Merci à Guillaume L. et Giuseppe, toujours partants pour résoudre mes "énigmes" mathématiques. Pour leurs conseils et leur soutien, je voudrais également remercier Jacques M., Francis, Gulten, Dramane, Roland, Linas, Jeremy, Erwan, Stéphane, Manon, Pierre-Alain, Sophie, Guillaume

C., Fanda, Nadine, Catherine L. et Danièle. Je n'oublie pas non plus les anciens: Kenza, Nicolas, Nick, Laura, et Imen.

Je souhaite également remercier ma famille pour leur soutien sans faille. Les dernières relectures de Korede et Sophie R. m'ont été d'une aide précieuse, je les en remercie. Enfin, je voudrais remercier tout particulièrement Julien pour m'avoir épaulée et encouragée pendant toutes ces années, ainsi que pour son amour de la macro micro-fondée.

Contents

Introduction générale

Amorcée aux Etats-Unis en 2007, sur le marché des subprimes[1], la crise financière a débouché fin 2008 sur la plus grande récession qu'aient connue les pays développés depuis la crise de 1929. En 2009, le PIB français a reculé de 2.7% et cette contraction de l'activité s'est traduite par une dégradation du marché du travail : le taux de chômage s'élève à 9.1% en 2009 contre 7.4% en 2008. Bien qu'exceptionnelle par son ampleur, cette récession partage les principales caractéristiques des récessions précédentes. La crise financière a en effet occasionné une baisse conjointe de la production, de l'investissement et de l'emploi, ainsi qu'un ralentissement de la consommation. Ces épisodes de récession sont suivis de phases de reprise économique, pendant lesquelles la croissance s'accélère et le chômage recule. Le système économique est ainsi soumis à des cycles conjoncturels ; ces cycles sont de périodicité et d'amplitude irrégulières, mais se caractérisent par des régularités statistiques liées à la volatilité et aux covariations des principaux agrégats macroéconomiques. Quels sont les mécanismes à l'origine des cycles conjoncturels ? Quel est l'impact des fluctuations conjoncturelles sur le bien-être des ménages ?

Ces questions, qui constituent un thème central de la macroéconomie, sont traditionnellement étudiées à partir de modèles à agents représentatifs : les fluctuations conjoncturelles sont analysées au regard des décisions d'investissement et d'emploi d'une firme et d'un travailleur représentatifs. Les fluctuations du nombre d'entreprises et du nombre d'employés, ainsi que celles des flux de création et de destruction associés, ont pendant longtemps été largement

[1] prêts hypothécaires à risque.

ignorées de l'analyse macroéconomique. Elles représentent pourtant une facette importante des fluctuations : suite à la crise, le nombre de défaillances d'entreprises s'est accru de 12% entre 2008 et 2009 et le rythme des créations d'entreprises s'élevait en 2008 à 1.8% seulement (contre 12,5% en 2007). Du côté du marché du travail, c'est avant tout la baisse des créations d'emploi qui explique le recul de l'emploi pendant la crise. Le taux de création d'emploi trimestriel a reculé d'un point de pourcentage entre 2008 et 2009 tandis que le taux de destruction d'emploi est resté stable durant cette période. Tenir compte de ces flux de création et de destruction permet une analyse plus fine du cycle économique. Cette approche est également susceptible de conduire à l'identification de nouveaux mécanismes de propagation des chocs conjoncturels.L'approche en termes de flux a récemment contribué à renouveler l'analyse du marché du travail. Depuis les travaux de Diamond (1982), Mortensen (1982) et Pissarides (1985), les flux de création et de destruction d'emploi occupent une place de plus en plus centrale dans l'analyse du marché du travail. L'emploi et le chômage sont analysés à partir des flux bruts d'emploi sous-jacents et non directement à partir des variations du stock d'emploi. Ce changement de perspective a été favorisé par le développement, ces vingt dernières années, de bases de données microéconomiques longitudinales. Ces données ont permis de mettre en évidence l'importance des flux bruts d'emploi, ainsi qu'une forte hétérogénéité de ces flux entre entreprises d'un même secteur : chaque année, certaines entreprises créent des emplois pendant que d'autres en détruisent ; de nouvelles entreprises se créent alors que d'autres disparaissent. Les facteurs de production sont ainsi sans cesse réalloués entre entreprises.

L'analyse en termes de flux n'a cependant été que peu explorée au-delà du marché du travail. En particulier, le rôle des réallocations de capital productif et celui des flux de création et de destruction d'entreprises sont encore mal connus. L'objectif de cette thèse est d'étudier le rôle de ces flux dans la dynamique du cycle conjoncturel. Il s'agit alors d'analyser les flux de réallocation et les déterminants microéconomiques sous-tendant les évolutions macroéconomiques. Plus particulièrement, ce travail s'attache à montrer les effets de ces flux de réallocation en présence de frictions sur les marchés du travail et du capital. Nous nous intéresserons en particulier aux frictions qui distordent l'allocation des facteurs de production telles que les contraintes de crédit, les coûts d'ajustement et les coûts de recherche d'emploi.

Dans le premier chapitre, écrit en collaboration avec F. Pappadà, nous étudions l'impact d'un choc conjoncturel sur la production agrégée lorsque les entreprises font face à des contraintes de crédit. Depuis les travaux de Bernanke and Gertler (1989) et Kiyotaki and Moore (1997), il est bien établi que les imperfections sur le marché du crédit contribuent à amplifier et à propager les chocs macroéconomiques en exacerbant les variations des capacités d'emprunt et d'investissement des entreprises. Ce mécanisme d'accélérateur financier est mis en évidence à partir des décisions d'investissement d'une firme représentative. En tenant compte de l'effet des frictions financières et des chocs macroéconomiques sur le nombre d'entreprises, le modèle théorique que nous proposons met en évidence un nouveau canal d'amplification. Nous montrons que les contraintes de crédit rendent les entreprises plus vulnérables à un choc conjoncturel, ce qui accroît le nombre d'entreprises conduites à liquider leur activité pendant la récession. La récession occasionne alors une plus forte baisse de production en présence de contraintes financières. Le modèle prédit également que les entreprises liquidées pendant la récession étaient en moyenne moins productives que celles qui survivent. Ce résultat suggère que les variations du taux de sortie des entreprises tendent à accroître la productivité agrégée en éliminant davantage d'entreprises pendant la récession. Le second chapitre étudie cette question et analyse les déterminants microéconomiques de l'évolution de la productivité agrégée dans le cycle. Pour ce faire, une nouvelle méthode de décomposition de la productivité agrégée est développée. Contrairement aux décompositions usuelles, la méthode que nous proposons s'appuie sur une fonction de production construite à partir de l'agrégation de fonctions de production microéconomiques hétérogènes. Cette approche permet de donner un fondement théorique à la mesure des effets des flux de réallocation, ainsi que ceux des créations et destructions d'entreprises sur l'efficacité du système productif. Estimée sur données françaises sur la période 1991-2006, cette décomposition indique que les créations et destructions d'entreprises jouent un rôle négligeable dans la dynamique de la croissance de la productivité agrégée. En revanche, l'efficacité de l'allocation des ressources entre entreprises apparaît contracyclique et contribue significativement à la volatilité de la productivité agrégée. Ainsi, ces résultats suggèrent que les récessions contribuent à améliorer la productivité agrégée, non via une plus forte destruction d'entreprises, mais à travers une amélioration de l'allocation des ressources. Le dernier chapitre, fruit d'une collaboration avec J-O Hairault et F. Langot, s'intéresse aux conséquences des fluctuations économiques sur le bien-être des agents économiques. La stabilisation des cycles économiques est un objectif central de politique économique. Des études ont par ailleurs montré que la volatilité liée aux cycles conjoncturels réduisait le bien-être des ménages (Wolfers, 2003). Pourtant, les tentatives formelles d'évaluation du coût des fluctuations suggèrent que ce dernier serait négligeable. Dans ce chapitre, nous montrons que tenir compte des flux d'emploi et des frictions liés aux réallocations d'emploi permet de rendre compte d'un coût en bien-être des fluctuations conjoncturelles non négligeable. En présence de frictions d'appariement, les fluctuations économiques augmentent le niveau moyen du chômage et réduisent ainsi le niveau de consommation des ménages. Les fluctuations conjoncturelles sont alors coûteuses, non seulement du fait de l'aversion des ménages pour l'incertitude macroéconomique, mais également du fait des conséquences indirectes de cette volatilité sur le niveau moyen de consommation. La suite de cette introduction présente tout d'abord le cadre théorique dans lequel s'inscrit cette thèse, pour ensuite présenter plus en détail les principales contributions de ces trois travaux de recherche.

L'analyse des fluctuations économiques : de l'agent représentatif aux flux de réallocation entre agents hétérogènes

Les cycles économiques sont appréhendés depuis la fin des années 1980 comme résultants de chocs exogènes, les sources d'impulsions, qui se propagent dans le système économique en donnant lieu à une suite récurrente de phases de croissance et de récession[2]. L'analyse des cycles économiques passe alors par l'identification des sources d'impulsion et de leurs mécan-

[2]Les premières représentations du schéma impulsion-propagation datent des années 1930 (Frisch, 1933; Slutsky, 1937).

ismes de propagation. La théorie des cycles réels (Kydland and Prescott, 1982; Long and Plosser, 1983), qui s'inscrit dans ce cadre impulsion-propagation, a modifié le paradigme de la modélisation des fluctuations conjoncturelles. Cette théorie reprend le projet défini par Lucas (1977), dont l'objectif est de faire reposer l'analyse des fluctuations économiques sur des fondements microéconomiques. Elle met le modèle néoclassique de croissance optimale au centre de l'analyse des fluctuations conjoncturelles. Les fluctuations sont ainsi analysées comme résultant des réponses optimales des agents économiques à des chocs technologiques exogènes. Plus précisément, il s'agit d'étudier les interactions entre un ménage et une entreprise représentatifs dans un cadre où les marchés sont parfaitement flexibles et concurrentiels. Dans cet environnement où aucune imperfection de marché ne distord l'allocation des ressources, les fluctuations conjoncturelles sont optimales. Le cycle économique n'est alors plus envisagé comme un écart au produit potentiel (« naturel ») mais comme les fluctuations mêmes du produit naturel. Cette vision du cycle économique laisse donc peu de place à l'intervention publique et aux politiques de stabilisation. L'explication des fluctuations conjoncturelles par des chocs technologiques est loin de faire consensus dans la littérature. En particulier, le courant de la nouvelle macroéconomie keynésienne souligne l'importance des chocs de demande (d'origine monétaire ou budgétaire) et des imperfections de marché pour expliquer les fluctuations conjoncturelles. Au-delà du débat sur les sources possibles d'impulsion[3] , la méthodologie et le cadre d'analyse proposés par la théorie des cycles réels sont devenus le cadre de référence de l'analyse conjoncturelle. Dès lors, les fluctuations économiques sont expliquées à partir de modèles d'équilibre général dynamiques micro-fondés, dont les prédictions sont ensuite confrontées aux faits observés. A ce titre, le modèle des cycles réels permet de reproduire la volatilité relative de la consommation et de l'investissement, ainsi que leur corrélation avec le produit. Toutefois, ce modèle ne génère pas suffisamment d'amplification pour expliquer les fluctuations à partir des seuls chocs technologiques. D'autres canaux d'amplification doivent donc être trouvés. En outre, le modèle des cycles réels possède également de faibles mécanismes

[3]Ce débat se situe également sur le terrain empirique. En particulier, différentes études utilisant la méthodologie VAR (Vector Autoreregressive), indiquent que les chocs technologiques expliquent tout au plus un quart des fluctuations du produit.

de propagation. Cogley and Nason (1995) font remarquer que, dans ce modèle, la persistance du produit et de l'emploi reflète exactement la dynamique du choc technologique. Ce modèle ne permet donc pas de rendre compte de l'existence de phases récurrentes d'expansion et de récession à partir de chocs non persistants.

Face à l'échec du modèle des cycles réels pour reproduire certaines dimensions des fluctuations observées, la littérature s'est tournée vers les imperfections de marché. Tout en conservant le cadre d'analyse développé par le courant des cycles réels, Merz (1995) et Andolfatto (1996) soulignent le rôle des frictions d'appariement sur le marché du travail. Outre qu'elle permette de modéliser l'évolution du chômage (et non seulement des heures travaillées), l'introduction des frictions de recherche d'emploi renforce les mécanismes de propagation et d'amplification des chocs technologiques.[4] Certains travaux s'intéressent quant à eux aux imperfections des marchés financiers. Bernanke and Gertler (1989), Bernanke et al. (1999), Carlstrom and Fuerst (1997), ou encore Kiyotaki and Moore (1997) ont étudié la propagation des chocs conjoncturels lorsque les entreprises font face à des contraintes de crédit. Ces travaux montrent que les fluctuations des conditions de financement des entreprises constituent un mécanisme important de propagation et d'amplification des chocs conjoncturels. Ici encore, le cadre d'analyse retenu est celui d'un modèle à agents représentatifs. Plus récemment, Chari et al. (2007) ont proposé un cadre quasi-comptable qui permet d'évaluer quantitativement l'importance relative des différentes frictions pour les fluctuations macroéconomiques. Ils modélisent les frictions de façon générique, par l'écart aux conditions optimales gouvernant les décisions des agents représentatifs dans un modèle néoclassique sans frictions. La plupart des modèles avec frictions pouvant s'écrire sous cette forme, cette approche permet d'évaluer indirectement les performances empiriques de ces différents modèles. Les résultats de Chari et al. (2007) indiquent que les frictions qui se manifestent comme une distorsion sur l'arbitrage consommation-loisir de l'agent représentatif et celles qui se manifestent comme un choc d'efficacité sur la fonction

[4]Toutefois, rendre de compte de la volatilité du chômage avec ce type de modèle reste problématique. Shimer (2005) montre que le modèle d'appariement standard ne permet pas de reproduire la volatilité du chômage à partir de chocs de productivité.

de production de la firme représentative jouent un rôle central dans les fluctuations conjoncturelles. Utilisant un modèle néokeynesien, Justiniano et al. (2010) trouvent au contraire que les chocs sur la décision d'investissement sont le facteur principal à l'origine des fluctuations du produit et des heures travaillées. En explorant le rôle des frictions pour expliquer les fluctuations conjoncturelles, cette thèse s'inscrit en prolongement de cette ligne de recherche. Nous nous intéressons en particulier au rôle des contraintes de crédit (chapitre 1) et des frictions liées aux réallocations de ressources entre entreprises (chapitre 2) comme source d'amplification et de propagation des variations du produit et de la productivité agrégés. Nous analysons également l'effet des frictions d'appariement sur le coût en bien-être des fluctuations conjoncturelles (chapitre 3).

Ce travail se distingue du reste de la littérature par la place qu'il accorde aux flux de création et de destruction, non seulement d'emploi, mais également de capital productif et d'entreprises. Depuis les travaux de Diamond (1982), Mortensen (1982) et Pissarides (1985), l'analyse des frictions sur le marché du travail est étroitement liée au processus de création et de destruction d'emploi. Il n'en est cependant pas de même pour les frictions associées au capital productif, le plus souvent étudiées dans un cadre à agent représentatif. Par ailleurs, les flux de création et de destruction d'entreprises restent encore largement ignorés de l'analyse du cycle économique. Pourtant, les bases de données microéconomiques longitudinales qui ont permis de mettre en évidence l'importance des flux bruts d'emploi, ont également révélé d'important flux de création et de destruction d'entreprises : chaque année, environ 10% des emplois sont créés et détruits, dans le même temps 8% des entreprises sont créées et presque autant sont détruites (Davis and Haltiwanger, 1990; Dunne et al., 1989)[5]. Les études empiriques soulignent également une disparité importante de ces flux entre entreprises. Au sein d'un même secteur, certaines entreprises créent des emplois pendant que d'autres en détruisent ; de nouvelles entreprises se créent alors que d'autres disparaissent. Les facteurs de produc-

[5]Il est plus difficile d'obtenir des données pour étudier les réallocations de capital existant. A partir de données de ventes et d'acquisition d'équipements, d'usines et de bâtiments, Eisfeldt and Rampini (2006) trouvent que les réallocations de capital existant représentent environ un quart de l'investissement annuel et entre 1.4 et 5.5% du stock de capital.

tion sont ainsi sans cesse réalloués entre entreprises. Tenir compte de ces flux de réallocation est d'autant plus important qu'il existe une forte hétérogénéité entre entreprises, notamment en termes de productivité. Les deux premiers chapitres de cette thèse explorent les conséquences des flux de réallocation et des flux de création et destruction d'entreprises. Cette dimension microéconomique permet d'approfondir la compréhension de l'effet des frictions sur les fluctuations du produit et de la productivité agrégée. Cette approche trouve écho dans une série de travaux récents qui montrent les limites du modèle à agent représentatif pour analyser les implications macroéconomiques des frictions. Buera and Moll (2011) montrent qu'utiliser l'approche quasi-comptable de Chari et al. (2007) pour évaluer l'importance relative des différentes frictions dans un modèle à agent représentatif peut conduire à des conclusions erronées. En effet, selon la nature de l'hétérogénéité sous-jacente (sur les technologies d'investissement, de production, ou de recrutement), Buera and Moll (2011) montrent qu'un resserrement du crédit peut se manifester dans le modèle à agent représentatif comme une distorsion sur la décision d'investissement, d'offre de travail, ou de productivité. L'analyse des frictions dans le modèle à agent représentatif donne ainsi une vision biaisée du rôle des frictions dans les fluctuations conjoncturelles. De même, Chang et al. (2010) montrent dans le cadre d'un modèle à agents hétérogènes avec des frictions financières comment l'inférence à partir du modèle à agent représentatif conduit à conclure, à tort, que les fluctuations sont dues à des frictions sur la décision d'offre de travail.

Chapitre 1 - Un nouveau canal d'amplification des chocs macroéconomiques : contraintes de crédit et cessation d'activité

Depuis la crise financière de 2008, l'interaction entre les marchés financiers et l'économie non-financière a suscité un regain d'intérêt. De nombreuses études empiriques ont mis en évidence les conséquences des conditions financières sur la durée et l'amplitude des récessions. Ainsi, les récessions associées à un resserrement du crédit seraient en moyenne plus longues et plus profondes que les autres (Claessens et al., 2009). D'autre part, Braun and Larrain (2005) ont montré que pendant les récessions, les entreprises dans les secteurs caractérisés par un plus fort besoin en financement externe sont également les plus touchées par la récession. Il s'agit en particulier de secteurs dans lesquels la taille des investissements est élevée par rapport au cash flow ou bien des secteurs dans lesquels les actifs sont incorporels et peuvent par conséquent difficilement être utilisés comme collatéral. Braun and Larrain (2005) montrent également que le nombre d'établissements a davantage diminué dans ces secteurs financièrement vulnérables. Ce résultat suggère que, suite à un choc conjoncturel, les frictions sur les marchés financiers n'affectent pas seulement les décisions d'investissement des entreprises (la marge intensive), mais également les fluctuations du nombre d'établissements ou d'entreprises (la marge extensive). Les fluctuations à la marge extensive sont, par ailleurs, une dimension importante des fluctuations conjoncturelles. Les données américaines publiées par le *Bureau of Labor Statistics* indiquent qu'entre 1994 et 2010, les créations et destructions d'établissements expliquent environ 20% de la volatilité de la croissance de l'emploi. Pourtant, c'est une dimension qui n'a été que très peu explorée dans l'analyse du cycle économique. En particulier, la littérature macroéconomique sur les contraintes de crédit (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997) ne s'intéresse qu'à l'effet des frictions financières sur la marge intensive et ignore leurs effets sur la marge extensive. Cette littérature met en évidence un mécanisme d'accélérateur financier : les chocs conjoncturels seraient amplifiés par leurs conséquences sur la capacité d'emprunt des entreprises. L'objectif de ce chapitre est d'analyser l'impact des contraintes de crédit sur les destructions d'entreprises. A l'aide d'un modèle théorique, nous montrons comment les contraintes de crédit conduisent à amplifier l'effet des récessions sur la production agrégée en fragilisant les entreprises. Les fluctuations du taux de sortie sont exacerbées : la récession conduit des firmes qui auraient survécu en l'absence de contraintes de crédit à liquider leur activité. Ce modèle prédit que les pays présentant des marchés financiers peu développés sont caractérisés par des taux de destruction d'entreprises moyens plus élevés mais également par une plus grande volatilité du taux de destruction. Contrairement au canal traditionnel de l'accélérateur financier, le mécanisme d'amplification que nous mettons en évidence dans ce chapitre ne dépend pas de la sensibilité de la capacité d'emprunt des firmes au choc agrégé.

Le modèle décrit dans un cadre dynamique les décisions d'investissement et de cessation d'activité des entreprises en présence de contraintes de crédit. Nous caractérisons analytiquement la décision de sortie des entreprises, contribuant ainsi à la littérature initiée par McDonald and Siegel (1985), Dixit (1989) et Hopenhayn (1992). Comme chez Bernanke and Gertler (1989) et Carlstrom and Fuerst (1997), les contraintes de crédit proviennent de l'existence d'une asymétrie d'information entre l'entreprise et la banque, ainsi que du coût qu'occasionne la procédure de défaut pour les banques. Le coût associé au défaut rend le financement externe plus coûteux que l'autofinancement pour l'entreprise. Le contrat d'endettement fait dépendre le montant d'emprunt et le taux auquel les entreprises peuvent emprunter des fonds propres détenus par l'entreprise. Le montant de fonds propres et donc les conditions d'accès au crédit des entrerprises fluctuent en raison des chocs agrégés et idiosyncratiques de productivité auxquels sont exposées les entreprises. A l'hétérogénéité exogène en termes de productivité, s'ajoute alors une hétérogénéité endogène en termes de capacité d'emprunt. Suite à un choc de productivité négatif, l'entreprise peut être amenée à faire défaut. Nous montrons que dans la plupart des cas, ces entreprises en défaut de paiement sont ensuite liquidées. Parce qu'elles ne détiennent plus de fonds propres, beaucoup de ces entreprisees sont exclues du marché bancaire. Le défaut n'est toutefois pas la seule cause de cessation d'activité. Les études empiriques indiquent en effet que la majeure partie des cessations d'activité n'est pas liée à une procédure de défaillance. Notre modèle reproduit ce fait stylisé : les entreprises liquident leur activité lorsqu'elles ne font plus suffisamment de profit et non uniquement en cas de défaut. En présence de contraintes de crédit, les entreprises en cessation d'activité ne sont pas nécessairement les moins productives ; les entreprises qui font face à des coûts d'emprunt élevés ont également une probabilité de sortie plus élevées que la moyenne. Les contraintes de crédit affectent ainsi la composition des flux de destruction d'entreprises. Cependant, malgré cette corrélation imparfaite entre probabilité de sortie et productivité, il s'avère que les destructions d'entreprises contribuent davantage à la hausse de la productivité agrégée en présence de contraintes de crédit. Bien que les entreprises liquidées ne soient pas nécessairement les moins productives, elles sont malgré tout moins productives en moyenne que les entreprises restantes. En amplifiant l'effet de la récession sur les destructions d'entreprises, les contraintes de crédit conduisent donc à améliorer la productivité agrégée pendant la récession. Contrairement à l'intuition, les contraintes de crédit semblent donc exacerber l'effet de *cleansing* des récessions.

Par ailleurs, ce travail éclaire le débat sur l'importance de la marge extensive dans l'analyse des cycles macroéconomiques. A l'aide d'un modèle théorique simulé, Samaniego (2008) montre que les taux de sortie et les taux d'entrée des entreprises ne sont pas affectés par les chocs de productivité agrégée. De leur côté, Bilbiie et al. (2007) et Clementi and Palazzo (2010) mettent en avant le rôle de la marge extensive dans la propagation des chocs agrégés. En particulier, ces deux articles soulignent le rôle joué par les variations du nombre de créations d'entreprises. Ce chapitre se distingue de cette littérature émergente en mettant l'accent sur les fluctuations du taux de destruction d'entreprises plutôt que sur les créations. Nos résultats suggèrent que la sortie est une dimension importante des fluctuations macroéconomiques, en particulier lorsque les entreprises font face à des contraintes financières.

Chapitre 2 - Variation cyclique de la productivité agrégée : quelle est la contribution de l'efficacité allocative ?

L'idée selon laquelle les récessions permettraient d'améliorer la productivité agrégée est très répandue. Selon cette vision Schumpeterienne du cycle, les récessions permettraient de liquider les entreprises les moins productives et d'ainsi libérer les ressources en faveurs d'entreprises plus efficaces. C'est ce qu'on appelle l'effet de *cleansing* des récessions. Formalisé par Caballero and Hammour (1994), l'effet de *cleansing* suggère que les récessions améliorent l'efficacité dans l'allocation des ressources. La littérature s'est concentrée sur le rôle de la marge extensive (les entrées et les sorties d'entreprises) mais la récession pourrait également affecter la productivité agrégée en modifiant l'allocation des ressources entre firmes existantes. En particulier, si les coûts d'ajustement auxquels font face les entreprises sont asymétriques, l'efficacité dans l'allocation des ressources entre entreprises est susceptible de varier dans le cycle. Quelle est la contribution de l'efficacité allocative aux fluctuations de la productivité agrégée ? Pour répondre à cette question, nous proposons dans ce chapitre une nouvelle méthode de décomposition de la productivité agrégée qui permet d'isoler les effets des changements de productivité au sein de chaque firme, des changements d'allocation entre firmes existantes, et des changements liés aux flux de création et de destruction d'entreprises. Cette décomposition permet de décrire le lien entre la dynamique microéconomique et les évolutions de la productivité agrégée. Nous tiendrons tout particulièrement compte de l'hétérogénéité entre entreprises et des frictions qui distordent l'allocation des ressources entre entreprises. Cette approche étend ainsi la méthode de décomposition de la croissance développée par Solow (1957) à un cadre où les entreprises sont hétérogènes et où les ressources sont inefficacement allouées.

Une vaste littérature empirique a mis en évidence l'importance des flux de réallocation pour comprendre la dynamique de la productivité agrégée (Baily et al., 2001; Foster et al., 2001; Griliches and Regev, 1995; Bartelsman et al., 2009). Ces travaux se concentrent pour la plupart sur les changements de productivité à long terme et peu de faits ont donc été établis à l'horizon du cycle économique. Par ailleurs, ces articles utilisent une décomposition de la productivité agrégée qui s'appuie sur un indice de productivité agrégée défini comme étant la somme pondéré de la productivité globale des facteurs (PGF) des entreprises. Avec cette définition, la contribution des réallocations de facteurs est mesurée à partir de la corrélation entre les changements de parts de marché et le niveau des productivité des entreprises. Cependant, dès lors que l'on tient compte de la décroissance des productivités marginales des facteurs de production, la production totale peut diminuer si l'on réalloue les facteurs de production vers les entreprises dont la PGF est la plus élevée. La corrélation des parts de marché avec la PGF des entreprises n'est donc pas une mesure correcte des changements d'efficacité dans l'allocation des ressources. Ce chapitre en propose une mesure théoriquement fondée. Les facteurs de productions sont alloués efficacement lorsque la productivité marginale des facteurs est égalisée entre entreprises. La mesure d'efficacité allocative que nous proposons repose donc sur l'écart entre les productivités marginales des facteurs. Comme Restuccia and Rogerson (2008) et Chari et al. (2007), les frictions sont modélisées de façon générique, comme étant l'écart aux conditions d'optimalité gouvernant les décisions de capital et d'emploi d'une entreprise soumise à aucune friction. Les changements d'efficacité allocative peuvent alors être mesurés à partir des fluctuations de ces frictions. L'approche retenue pour modéliser les frictions inclut des frictions de natures diverses, telles que les coûts d'ajustement, les coûts de recherche d'emploi et les contraintes de crédit. La mesure de l'efficacité allocative capte ainsi les conséquences de toutes les imperfections de marché qui créent un écart entre les productivités marginales des entreprises. La productivité agrégée est dérivée à partir d'une fonction production agrégée, elle-même construite à partir de l'agrégation des fonctions de production microéconomiques. Le problème de l'agrégation est un problème classique en macroéconomie. Il a été établi qu'il n'est possible d'agréger les contraintes techniques d'entreprises hétérogènes qu'au prix de restrictions irréalistes sur les fonctions de production microéconomiques (Nataf, 1948). La fonction de production agrégée ne peut donc pas être une relation purement technique. Nous reprenons l'idée de Malinvaud (1993), qui envisage la fonction de production comme étant une relation entre la production agrégée et les facteurs de production agrégés, pour une allocation des ressources donnée. La fonction de production agrégée dépend alors de l'environnement économique et en particulier des éléments qui déterminent l'allocation des facteurs de production entre entreprises. Calculée à partir de cette fonction de production agrégée, la croissance de la productivité agrégée résulte des changements de productivité au sein de chaque entreprise ainsi que des changements dans l'allocation des ressources.

Cette décomposition est ensuite estimée sur données d'entreprises françaises sur la période 1991-2006. L'analyse empirique révèle que, contrairement à ce qui est suggéré par la théorie de l'effet de *cleansing* des récessions, l'entrée et la sortie des entreprises ont une contribution négligeable à la dynamique de la productivité agrégée. En revanche, les fluctuations dans l'efficacité allocative joue un rôle substantiel. En outre, il semblerait que ce qui tend à améliorer la productivité agrégée durant les récessions n'est pas la réallocation des ressources entre entreprises entrantes et sortantes, mais plutôt la réallocation entre entreprises existantes. Les fluctuations de l'efficacité allocative sont contracycliques, alors que la contribution de la marge extensive est procyclique. Ces résultats sont diamétralement opposés aux résultats obtenus en utilisant la décomposition de Foster et al. (2001). En effet, avec cette décomposition, la marge extensive joue un rôle significatif. Par ailleurs, l'efficacité allocative est positivement corrélée avec la croissance de la valeur ajoutée. Ce chapitre montre donc la sensibilité de l'analyse du rôle des réallocations dans les fluctuations de productivité agrégée à la méthode retenue pour construire la productivité agrégée. Ces résultats établissent des faits stylisés qui constituent une première étape dans la compréhension du rôle des frictions dans les fluctuations de la productivité agrégée. Ils appellent à de plus amples recherche sur les mécanismes théoriques à l'origine de la contracyclicité de l'efficacité allocative.

Chapitre 3 - Quand les fluctuations du taux de retour à l'emploi augmentent le niveau moyen du chômage

Les fluctuations du chômage résultent des variations sous-jacentes des taux de sortie et de retour dans l'emploi. Une récession se traduit par une hausse du taux de sortie et une baisse du taux de retour dans l'emploi, tandis qu'une période d'expansion est caractérisée à l'inverse par une baisse du taux de sortie et une hausse du taux de retour. Une littérature récente a cherché à identifier les contributions respectives des taux de sortie et de retour dans les fluctuations du chômage aux Etat-Unis. La récession crée-t-elle du chômage parce que les entreprises ralentissent le processus d'embauche ou bien parce qu'elles intensifient les destructions d'emploi ? Fujita and Ramey (2008) remettent ainsi partiellement en cause les résultats de Shimer (2007) qui établissent la prédominance du taux de retour dans l'emploi. Elsby et al. (2008) montrent que la contribution relative des sorties et des entrées est relativement équilibrée sur données européennes.

Dans ce chapitre, nous montrons qu'au-delà de leur contribution relative, les fluctuations des taux de sortie et de retour dans l'emploi ont des implications diamétralement opposées pour le niveau moyen de l'emploi. Parce que le stock de chômeurs covarie négativement avec le taux de retour dans l'emploi, l'augmentation du taux de retour à l'emploi produit un effet de plus en plus faible sur le niveau du chômage. A l'inverse, la baisse du taux de retour à l'emploi produit un effet qui s'accentue à mesure que le stock de chômage s'accroit. Ainsi, les chocs sur le taux de retour à l'emploi ont un effet asymétrique sur le chômage: cette non-linéarité explique que les fluctuations du taux de retour diminuent le niveau moyen de l'emploi. En revanche, parce que le stock d'emploi covarie négativement avec le taux de séparation, la hausse du taux de séparation à un effet plus faible qu'une baisse du taux de séparation : les fluctuations du taux de séparation augmentent le niveau moyen de l'emploi. En outre, bien que les fluctuations du taux de sortie de l'emploi contribuent à expliquer une fraction significative des fluctuations du chômage, il s'avère que leur impact sur le niveau moyen du chômage est négligeable au regard de l'effet des fluctuations du taux de retour à l'emploi : l'effet des fluctuations du taux de sortie de l'emploi est inférieur d'un ordre de grandeur à celui des fluctuations du taux de retour à l'emploi.

L'existence de ces non-linéarités dans l'interaction des flux et des stocks d'emploi implique que l'alternance de contractions et d'expansions peut augmenter le niveau moyen du chômage. Plus les fluctuations sont amples, plus ces non-linéarités sont susceptibles de créer des effets importants sur la moyenne de l'emploi. Au-delà de la volatilité du taux de retour à l'emploi, nous montrons que la persistance de leurs fluctuations peut également jouer un rôle important. En outre, le niveau du chômage structurel, défini comme le niveau du chômage indépendant des fluctuations économiques, rend la volatilité plus ou moins importante pour le niveau moyen du chômage: un niveau structurel plus élevé amplifie les effets de la volatilité.[6] Nous proposons

[6] On pense ici naturellement à l'économie française dont le taux de chômage stucturel est considéré comme plus élevé. Cependant reste à vérifier que les cycles du taux de retour en emploi ne sont pas moins amples en

une évaluation quantitative des pertes et des gains d'emploi pour les processus des taux de séparation et de retour estimés sur données américaines. Nous considérons simultanément les données fournies par Shimer (2007) et Hall (1991) qui diffèrent par leur définition du chômage. Nous montrons alors que les fluctuations du taux de retour dans l'emploi observées aux Etats-Unis ont pu causer jusqu'à près d'un point de chômage en plus en moyenne depuis la seconde guerre mondiale. Ceci est vrai pour une définition élargie du chômage à certains travailleurs hors de la population active comme le propose Hall. En revanche, quelle que soit la conception du chômage, les fluctuations du taux de séparation n'impliquent aucun effet significatif sur la moyenne du chômage. Ainsi, loin d'être une curiosité théorique, les non-linéarités dans l'interaction des flux et des stocks d'emploi font des fluctuations du taux de retour dans l'emploi un facteur explicatif du taux de chômage moyen.

Les implications de politique économique de ce travail sont potentiellement importantes. En augmentant le niveau de chômage moyen, les fluctuations conjoncturelles réduisent le bien-être des ménages, ce qui justifie la mise en oeuvre de politiques de stabilisation. Ce résultat est d'autant plus important que de nombreuses évaluations du coût en bien-être des fluctuations suggèrent qu'ils seraient négligeables. Lucas (1987) évalue l'incertitude que les fluctuations font peser sur la consommation des ménages à un coût équivalent à 0.01% de leur consommation annuelle. Les effets des fluctuations du taux de retour à l'emploi sur le niveau de chômage moyen s'élèvent à un coût équivalent à environ 1% de la consommation annuelle des ménages. Par ailleurs, nos résultats soulignent la forte interaction entre les politiques structurelles et le coût des fluctuations. Un taux de chômage structurel élevé peut également être dommageable parce qu'il rend plus couteux les fluctuations du taux de retour à l'emploi. Dans cette perspective, les politiques structurelles peuvent se substituer aux politiques de stabilisation pour diminuer les conséquences de la volatilité sur le niveau de l'emploi.

France, ce qui diminuerait l'effet sur le niveau moyen du chômage. Il n'existe toutefois pas encore de séries longues de taux de retour dans l'emploi pour l'économie française qui permettraient d'explorer cette question.

Chapter 1

All you need is loan. Credit market frictions and the exit of firms in recessions[1]

1.1 Introduction

The interplay between macroeconomic shocks and financial conditions has gained renewed interest after the global recession of 2008-2009. Several empirical studies document the role of credit market frictions on the severity and duration of recessions. Over a sample covering 21 OECD countries during the period 1960-2007, Claessens et al. (2009) show that recessions associated with credit crunches tend to be deeper and longer than other recessions. Moreover, a series of papers (Kroszner et al., 2007; Dell'Ariccia et al., 2008; Bricongne et al., 2009) emphasize that during banking crises, and in particular during the recent financial crisis, firms in financially vulnerable sectors were hit harder. As shown by Braun and Larrain (2005), industries highly dependent on external finance are more sensitive not only to banking crises, but also to recessions caused by non-financial factors. Braun and Larrain (2005) also show

[1] This chapter is joint with Francesco Pappadà.

that the number of establishments is more severely affected in those industries, suggesting that financial conditions affect not only the response of firms' output to business cycle shocks (the intensive margin) but also the response of the number of firms (the extensive margin). The extensive margin is an important channel for aggregate fluctuations. Data on the US private sector provided by the *Bureau of Labor Statistics* from 1994 to 2010 show that entering and exiting establishments account for 20% of the volatility of annual employment growth[2]. However, very little research has been devoted to understanding the firms' exit decision under financial frictions and its implication for business cycle fluctuations. Following Bernanke and Gertler (1989) and Kiyotaki and Moore (1997), the theoretical literature has focused on the implication of credit frictions on the intensive margin and leaves aside the consequences of credit frictions on the extensive margin.

The objective of this paper is to analyze how credit market frictions may alter business cycle dynamics by modifying the exit behavior of firms. Our main contribution is to show that the extensive margin yields a significant amplification mechanism. Credit market frictions amplify the fluctuations in the exit rate as they make the industry more vulnerable to negative aggregate shocks, thereby leading to a substantially larger output loss. In the presence of credit market frictions, the fall in aggregate productivity induces the exit of firms that would have survived in a frictionless economy. Unlike the standard financial accelerator, the mechanism emphasized in this paper does not depend on the sensitivity of firms' net worth to aggregate shocks. We show that when the balance sheet channel is shut down, credit market frictions amplify aggregate fluctuations by increasing the number of firms vulnerable to the adverse productivity shock.

We explore this mechanism in a model of firm dynamics with credit constraints. We analyti-

[2]We use the data from the Business Employment Dynamics program which are constructed from state unemployment insurance records.

cally characterize the exit decision of firms that face credit constraints and incur a fixed cost of production. This is a contribution to the existing literature on the exit decision of firms (McDonald and Siegel, 1985; Dixit, 1989; Hopenhayn, 1992).[3] In the model, as in Bernanke and Gertler (1989) and Carlstrom and Fuerst (1997), credit constraints endogenously arise from asymmetric information and costly state verification. The financial contract determines the amount the firm can borrow and the interest rate charged by the financial intermediary as a function of the firm's levels of productivity and net worth. When firms are hit by an adverse productivity shock, they may be unable to repay their debt, they default and are left with zero net worth. After default, most firms are excluded from the credit market and are therefore forced to exit. However, default is not the only motive for exit as firms also decide to leave the market when their expected profits are too low. This happens when firms are not sufficiently productive, as in the frictionless economy, but also when their balance sheets are too weak. Firms with a low net worth face tighter credit constraints and higher borrowing costs, which raise their probability to exit. As the firms' exit decision depends on their net worth, firms that exit are not necessarily the least productive ones. Credit market frictions therefore distort the selection of exiting firms: some high productivity firms are forced to exit in case of financial distress while some low productivity firms may survive. Despite this imperfect selection, we find that average idiosyncratic productivity rises more after a fall in aggregate productivity when firms face credit constraints. Credit market frictions magnify the increase in the exit rate as they increase the number of firms vulnerable to the fall in aggregate productivity. Though productivity is not the only determinant of survival, these exiting firms are, on average, less productive than surviving firms. Therefore, by amplifying the number of exiting firms, credit market frictions also amplify the increase in average idiosyncratic productivity.

We then investigate the consequences of the fluctuations in the exit rate on aggregate output. We focus on the effects of credit frictions on the vulnerability of the industry to a fall in ag-

[3] All these papers consider perfect financial markets, and to the best of our knowledge, the properties of the exit decision under credit constraints have not been derived analytically.

gregate productivity, and leave out the consequences of the endogenous change in net worth. In fact, the consequences of the fall in net worth are well-known at the intensive margin, and are likely to generate a similar mechanism at the extensive margin. In the economy with credit frictions, the decline in aggregate productivity leads to a significantly larger output loss. We disentangle the output loss induced by the intensive margin from that induced by the exit of firms. We show that, when the impact of the fall in aggregate productivity on the firms' net worth is shut down, credit frictions amplify the output loss at the exit margin while dampening aggregate fluctuations at the intensive margin. This amplification channel at the exit margin appears then more robust than the financial accelerator mechanism as it does not hinge on the sensitivity of firms' net worth to aggregate shocks. These results suggest that the exit margin is an important channel for understanding the aggregate implications of credit market frictions.

This paper contributes to the debate on the importance of the extensive margin for business cycle dynamics. While Samaniego (2008) finds that entry and exit respond very little to aggregate productivity shocks, Lee and Mukoyama (2008), Bilbiie et al. (2007), and Clementi and Palazzo (2010) show that the extensive margin plays a crucial role for aggregate fluctuations. These papers focus on the behavior of the entry rate, and highlight the effects of the procyclicality of the entry rate on the propagation of aggregate shocks. By contrast, we emphasize the role of the exit margin and assume a constant mass of potential entrants. Despite this assumption, the actual number of entrants is endogenous as firms enter the market only when their expected profits are sufficiently high. We show that the output loss due to the exit of incumbents and potential entrants accounts for a large fraction of the output drop, in particular when firms are subject to credit constraints. Our results therefore suggest that the exit of firms is an important dimension of aggregate fluctuations in less financially developed countries.

This paper is also related to the literature on firm dynamics in a credit constrained environment (Cooley and Quadrini, 2001; Miao, 2005; Arellano et al., 2009). The theoretical model is closely related to Cooley and Quadrini (2001). In their paper, they show how financial frictions can account for the size and age dependance of firm growth. Miao (2005) analyzes how the interaction between financing and production decisions affects firm turnover. Arellano et al. (2009) examine the link between financial development and firm growth.[4]

This paper is also related to the literature on the cleansing effect of recession (Caballero and Hammour, 1994; Barlevy, 2003; Ouyang, 2009). Barlevy (2003) studies how credit market frictions distort the selection of exiting firms. In his model, credit frictions may reverse the cleansing effect of recessions as they lead resources to flow from high to low productivity firms. This contrasts with our results as, in our setup, a higher productivity facilitates the firm's access to credit. We show that though some high productivity firms may exit during the recession, credit frictions mainly increase the exit rate of low productivity firms. Therefore, credit frictions lead average productivity to rise further during recessions, somewhat exacerbating the cleansing effect of recessions.

This paper is organized as follows. Section 2 describes the model of firm dynamics and credit constraints. In Section 3, we show analytically how the exit decision differs from the frictionless economy. In Section 4, we first analyze numerically the properties of the steady state economy. We then show how the distorted exit decision may amplify the consequences of a fall in aggregate productivity on average productivity and aggregate output. Section 5 concludes.

1.2 A model of firm dynamics and credit market frictions

In this section, we describe the model of firm dynamics with credit market imperfections. In what follows, we define the production technology and the timing of the firms' decisions, then

[4]Our model differs from these papers both in its focus and in its modeling choices, in particular with respect to exit. In these papers, exit is either exogenous or induced by default and financial imperfections. By contrast, in our model, firms endogenously exit when their expected profits are too low. As firms would exit even in the absence of financial frictions, our modeling assumption allows us to compare the exit behaviors in the credit constrained economy and in the frictionless economy.

present the frictionless economy and the economy with credit market frictions.

1.2.1 Technology and timing of decisions

The economy is constituted of risk neutral firms with a constant discount factor $0 < \beta < 1$. Firms are heterogenous with respect to their productivity and their net worth, and have access to a production technology with capital as the only input and decreasing returns to scale. Each period, firms incur a fixed operating cost to start production. After production, firms determine the amount of dividends to distribute and the amount of profits to reinvest. The firm can decide to stay in the market and reinvest its profit in production or invest in a risk-free asset. When the value from investing in the safe asset is higher than the value from producing, the firm chooses to exit and never enters again. Exiting firms lose the opportunity to receive future profits from production, but also avoid to pay the fixed cost. Firms therefore exit when their expected income from production is not sufficiently high to compensate the fixed cost.

After paying the fixed operating cost c, the firm produces output: $Z(\theta+\epsilon)k^{\alpha}$ with $0 < \alpha < 1$. The capital k used for production depreciates at rate $0 < \delta < 1$. Z denotes the value of aggregate productivity. Every period, firms are hit by a persistent idiosyncratic productivity shock $\theta \in [\theta_{\min}, \theta_{\max}]$, and a non-persistent idiosyncratic productivity shock $\epsilon \in [\epsilon_{\min}, \epsilon_{\max}]$. The non-persistent component ϵ is independently and identically distributed (i.i.d) across time and across firms, with distribution Φ, zero mean, and standard deviation σ. We impose $\epsilon_{\min} > -\theta_{\min}$, which ensures a non negative production whatever the value of the shock ϵ. The persistent component θ follows a Markov process independent across firms with conditional distribution $F(\theta'|\theta)$. The conditional distribution $F(\theta'|\theta)$ is assumed to be strictly decreasing in θ : the higher is the productivity shock at time t, the more likely are high shocks in period $t+1$. This assumption ensures that the value of the firm is an increasing function of the current productivity θ. In what follows, for any generic variable x, we adopt the notation x'

to define the next period value of the variable x.

The value of the persistent idiosyncratic shock is revealed one period in advance. Therefore, at the beginning of the period, firms choose their capital knowing their persistent idiosyncratic shock θ, the value of aggregate productivity Z, and their net worth e. At the beginning of the period, firms do not know their idiosyncratic shock ϵ. They observe the realization of ϵ after production, and must then reimburse their debt over the capital borrowed and the fixed operating cost $(c+k-e)$. They are left with the end-of-period net worth q. At the end of the period, a firm with net worth q observes its productivity shock θ', and decides its next period net worth e' (or equivalently the amount of dividends $(q-e')$ to distribute), and whether to exit or stay in the market. A firm decides to exit when its value from producing is lower than the value from investing in the safe asset, which is equal to $q_t + \Sigma_{s=0}^{+\infty}\beta^s[\beta(1+r)-1]e_{t+s+1}$. Note that if $\beta(1+r) \leq 1$, the value from investing in the safe asset simplifies to q_t. In that case, either the firm is indifferent about the timing of dividends ($\beta(1+r)=1$) or it prefers to distribute its end-of-period net worth as dividends ($\beta(1+r)<1$).

1.2.2 The frictionless economy

In the frictionless economy, firms borrow $(c+k-e)$ at the risk-free interest rate $r = 1/\beta - 1$. The value of a firm at the beginning of the period is:

$$V_{FL}(e,\theta,Z) = \max_k \mathbb{E}_\theta \int_{\epsilon_{\min}}^{\epsilon_{\max}} \max\left[q, \max_{e'}(q - e' + \beta V_{FL}(e',\theta',Z))\right] d\Phi(\epsilon),$$

where the end-of-period net worth is equal to

$$q = Z(\theta+\epsilon)k^\alpha + (1-\delta)k - (1+r)(c+k-e),$$

and $\mathbb{E}_\theta$ denotes expectations conditional on the current value of θ. The value of the firm depends on the expected outcome of its investment. Firm exit when the value from investing in the safe asset is higher than the value from investing in production. As $r = 1/\beta - 1$, the firm is indifferent about the timing of dividends and the value from investing in the safe asset is then equal to its end-of-period net worth q. Furthermore, the Modigliani-Miller theorem holds and the value of the firm is independent of its financing decision. In particular, the exit and capital decisions of the firm do not depend on its level of equity. It can be shown that, conditional on surviving, the program of the firm is equivalent to maximizing its expected profits:

$$\widehat{V}_{FL}(\theta, Z) = \max_{k} \mathbb{E}_{\theta} \int_{\epsilon_{\min}}^{\epsilon_{\max}} [Z(\theta+\epsilon)k^{\alpha} - (r+\delta)k - (1+r)c] d\Phi(\epsilon) + \beta \max \left[0, \widehat{V}_{FL}(\theta', Z)\right].$$

When credit markets are perfect, firms exit when they are not productive enough: they exit if $\theta' < \underline{\theta}_{FL}(Z)$, where $\underline{\theta}_{FL}(Z)$ is defined by $\widehat{V}_{FL}(\underline{\theta}_{FL}, Z) = 0$.

1.2.3 The economy with credit market frictions

As in Cooley and Quadrini (2001), we embed a one-period financial contract *à la* Bernanke and Gertler (1989) into a dynamic framework. Credit constraints arise from asymmetric information between the firm and the financial intermediary. After production, the non-persistent idiosyncratic shock ϵ is privately observed by the firm, whereas the financial intermediary can observe ϵ only at a cost μk^{α}. We consider a one-period debt contract in which the firm defaults when the shock is too low, and the financial intermediary monitors the firm's income only when the firm defaults. The terms of the financial contract depend on the value of the firm's net worth e, on its current productivity θ, and on the value of aggregate productivity Z, all observable by the financial intermediary and the firm at zero cost.

Assumption 1. The risk-free interest rate is such that: $\beta < \frac{1}{1+r}$.

As in Cooley and Quadrini (2001), this assumption implies that the risk-free rate is lower in the economy with credit frictions than in the frictionless economy, and guarantees that firms will not always choose to reinvest all their profits, thus giving an upper bound to their net worth. This condition can be interpreted as a general equilibrium property of economies with financial constraints. As it goes beyond the scope of this paper to analyze the impact of credit frictions on the risk-free rate, note that we choose to leave aside this general equilibrium effect when comparing the results in the credit constrained economy with the frictionless case. In the following, we therefore compare the credit constrained economy with the same economy without credit frictions but with the same risk-free rate r.

The capital chosen by the firm is financed by its equity e, and if $k + c > e$, the firm borrows $(k + c - e)$ at rate $\widetilde{r}$ from the financial intermediary. When a firm is not able to reimburse its debt, it defaults. In this case, the financial intermediary pays a cost to verify the firm's income and confiscates all the firm's income. The default threshold $\bar{\epsilon}$ is given by:[5]

$$Z(\theta + \bar{\epsilon})k^{\alpha} + (1 - \delta)k \geq (1 + \widetilde{r})(k + c - e). \tag{1.2.1}$$

Note that default leads to a zero net worth but does not necessarily lead to the exit of the firm, as observed empirically. Depending on its persistent productivity component θ, the firm could find profitable to stay in the market with zero net worth.

The financial intermediary lends $(k + c - e)$ to the firm only if its expected income from the loan is equal to the opportunity cost of the funds. The break even condition reads:

$$(1 + \widetilde{r})(k + c - e)[1 - \Phi(\bar{\epsilon})] + \int_{\epsilon_{\min}}^{\bar{\epsilon}} [Z(\theta + \epsilon)k^{\alpha} + (1 - \delta)k - \mu k^{\alpha}]\, d\Phi(\epsilon) \geq (1 + r)(k + c - e).$$

The expected income of the financial intermediary is equal to the repayment of the loan if the

[5] Note that the debt is never renegotiated after default. The financial intermediary could agree to reduce the debt to $(1+\widetilde{r})(k+c-e) - D$, with $0 \leq D \leq (1+\widetilde{r})(k+c-e) - (Z(\theta+\epsilon)k^{\alpha} + (1-\delta)k)$. This would leave the firm with end-of-period net worth $q = -D$. However, the renegotiation is never mutually profitable. Since there are no additional cost related to default, the firm always prefers to default and start the next period with zero net worth ($q = 0$) than to renegotiate the debt and have a negative net worth ($q = -D$).

firm does not default ($\epsilon \geq \bar{\epsilon}$) and to the firm's income net of monitoring costs when the firm defaults ($\epsilon < \bar{\epsilon}$). Using the default condition (Equation 1.2.1), we can rewrite the participation constraint of the financial intermediary as:

$$Z[\theta + G(\bar{\epsilon})]k^{\alpha} + (1-\delta)k - \mu k^{\alpha}\Phi(\bar{\epsilon}) \geq (1+r)(k+c-e),$$

with $G(\bar{\epsilon}) \equiv [1-\Phi(\bar{\epsilon})]\bar{\epsilon} + \int_{\epsilon_{\min}}^{\bar{\epsilon}} \epsilon d\Phi(\epsilon)$.

As it is more convenient to write the problem of the firm as a function of the default threshold $\bar{\epsilon}$, we characterize the financial contract by the couple $(k, \bar{\epsilon})$ and then derive the implied interest rate $\widetilde{r}$ charged by the financial intermediary from the default condition. Given Z, θ and e, the participation constraint indicates the amount the firm can borrow and the associated default threshold $\bar{\epsilon}$ required by the financial intermediary. A higher level of net worth relaxes the financial intermediary's participation constraint and allows the firm to borrow more capital. The credit constraint is tighter if the firm incurs a higher default rate when borrowing a larger amount. Assumption 2 gives the regularity condition on the distribution Φ that ensures a positive correlation between the amount the firm can borrow and the default threshold $\bar{\epsilon}$[6].

Assumption 2. The distribution function of the transitory shock is such that $\Phi'(\epsilon_{\min}) < \frac{Z}{\mu}$ and $\frac{\Phi'(\epsilon)}{1-\Phi(\epsilon)}$ is monotone in ϵ.

For some firms, the income of the financial intermediary is too low for its participation constraint to be satisfied. In fact, given θ and Z, there is a unique threshold $\underline{e}_b(\theta, Z)$ below which the financial intermediary refuses to lend any fund (see Appendix A). This threshold is defined as:

$$Z[\theta + G(\bar{\epsilon}_b)]k_b^{\alpha} + (1-\delta)k_b - \mu k_b^{\alpha}\Phi(\bar{\epsilon}_b) = (1+r)(k_b+c-\underline{e}_b), \tag{1.2.2}$$

[6] This condition will be necessary to prove the continuity of the value function. It implies that the income of the financial intermediary is either increasing in $\bar{\epsilon}$, or is an inverted U-shaped curve.

where $(\bar{\epsilon}_b, k_b)$ maximize the expected income of the financial intermediary. When the net worth of the firm is below $\underline{e}_b(\theta, Z)$, the financial intermediary would rather invest in the safe asset than lend to the firm. Proposition 1 shows that in that case the net worth of the firm is not sufficiently high to cover the fixed cost of production, and the firm is therefore forced to exit the market.

PROPOSITION 1. **Exit of firms due to credit rationing**
Firms that are not financed by the financial intermediary cannot cover their fixed cost of production:

$$\underline{e}_b(\theta, Z) \leq c, \quad \forall \theta \in [\theta_{\min}, \theta_{\max}], \quad \forall Z.$$

Proof: see Appendix A.

After production, the firm's end-of-period net worth is equal to:

$$q = \begin{cases} Z(\theta + \epsilon)k^{\alpha} + (1 - \delta)k - (1 + \tilde{r})(c + k - e) & \text{if } \epsilon > \bar{\epsilon} \\ 0 & \text{if } \epsilon \leq \bar{\epsilon}. \end{cases}$$

Using again the definition of the default threshold (Equation 1.2.1), the end-of-period net worth reads:

$$q = \max\{Zk^{\alpha}(\epsilon - \bar{\epsilon}); 0\}.$$

1.2.3.1 The firm's problem

Define V as the value of the continuing firm at the beginning of the period, before choosing its level of capital. The value of the firm depends on the outcome of its investment and on its exit decision. At the end of the period, the firm learns its next period productivity θ' and, depending on its end-of-period net worth, decides which fraction of its profit to distribute as dividends, and whether to stay or exit the market. When its end-of-period net worth is too low $q < \underline{e}_b(\theta', Z)$, the participation constraint of the financial intermediary is not satisfied.

As stated in Proposition 1, in that case the firm cannot finance the fixed cost of production and must therefore exit the market. On the other hand, when $q \geq \underline{e}_b(\theta', Z)$, the firm decides whether to stay in the market or exit by comparing the value from producing with the outside opportunity. As the discount rate is higher than the safe asset return r, the firm always prefers to distribute its end-of-period net worth as dividends rather than invest it in the safe asset. The firm therefore exits when its continuing value is lower than its end-of-period net worth q. We prove in Appendix A that the value function of the firm exists and is unique. The problem of the firm reads:

$$V(e, \theta, Z) = \max_{(k,\bar{\epsilon})} \mathbb{E}_\theta \left\{ \int_{\epsilon_{\min}}^{\epsilon_{\max}} I(q)q + (1 - I(q)) \max \left[q, \max_{e'} \left(q - e' + \beta V(e', \theta', Z) \right) \right] d\Phi(\epsilon) \right\}$$

with:

$$I(q) = \begin{cases} 0 & \text{if } q \geq \underline{e}_b(\theta', Z) \\ 1 & \text{if } q < \underline{e}_b(\theta', Z) \end{cases}$$

subject to:

$$Z[\theta + G(\bar{\epsilon})]k^\alpha + (1 - \delta)k - \mu k^\alpha \Phi(\bar{\epsilon}) \geq (1 + r)(k + c - e) \tag{1.2.3}$$

$$q = \max \left\{ Zk^\alpha(\epsilon - \bar{\epsilon}); 0 \right\} \tag{1.2.4}$$

$$\underline{e}_b(\theta', Z) \leq e' \leq q. \tag{1.2.5}$$

The firm maximizes its expected dividends subject to the participation constraint of the financial intermediary defined by Equation (1.2.3). Equation (1.2.4) describes the end-of-period net worth q, while equation (1.2.5) imposes that the firm cannot issue new shares[7] and can then increase its net worth only by reinvesting its profits. The firm faces a trade-off when deciding its level of capital. On the one hand, if the firm is solvent, a higher level of capital increases its next period level of production. On the other hand, it increases its probability to default as the default threshold required by the financial intermediary increases with the amount borrowed (Assumption 2).

We assume that the value function is differentiable. This allows us to derive analytical results on the exit decision of the firms. It also permits to characterize the dividend decision of the firm. Because the discount rate is higher than the risk-free rate (Assumption 1), the firm will not always choose to reinvest all its profits. It will distribute dividends if its end-of-period net worth $q \geq \bar{e}(\theta, Z)$ with $\bar{e}(\theta, Z)$ defined by $\beta \frac{\partial V(\bar{e},\theta,Z)}{\partial e} = 1$.

1.2.3.2 Exit conditions

By contrast with the frictionless economy, productivity is not the only determinant of firms' survival. Firms exit if they are not sufficiently productive, but they may also exit because their balance sheet position is too weak. In fact, in the presence of credit frictions, two additional motives for exit arise. At the end of the period, firms may exit because their level of net worth is not high enough for their participation constraint ($\beta V(q, \theta, Z) < q$) or for the participation constraint of the financial intermediary to be satisfied ($q < \underline{e}_b(\theta, Z)$). These exit conditions are described in the following proposition.

PROPOSITION 2. **Exit conditions**

For a given level of aggregate productivity Z, there exist three thresholds $\underline{\theta}(Z) < \theta^*(Z) < \theta^{**}(Z)$ that characterize the exit decision of the firm. These productivity thresholds[8] delimit

[7] Allowing $e' > q$ makes the financial constraints irrelevant as firms would finance all their investment with equity.

[8] The productivity thresholds $\underline{\theta}(Z)$, $\theta^*(Z)$ and $\theta^{**}(Z)$ are defined by the following equations:

$$\begin{aligned}
\bar{e}(\underline{\theta}, Z) &= \beta V(\bar{e}(\underline{\theta}, Z), \theta, Z) \\
\underline{e}_b(\theta^*, Z) &= \beta V(\underline{e}_b(\theta^*, Z), \theta^*, Z) \\
(1+r)(k_b + c) &= Z[\theta^{**} + G(\bar{\epsilon}_b)]k_b^{\alpha} - \mu k_b^{\alpha} \Phi(\bar{\epsilon}_b) + (1-\delta)k_b
\end{aligned}$$

four exit regions:

A. The firm exits when $\theta < \underline{\theta}(Z)$ whatever its level of net worth.

B. The firm exits when $\underline{\theta}(Z) \leq \theta < \theta^*(Z)$ if its end-of-period net worth is too low for its participation constraint to be satisfied: $q < \underline{e}_f(\theta, Z)$, where $\underline{e}_f(\theta, Z)$ is defined by $\underline{e}_f = \beta V(\underline{e}_f, \theta, Z)$.

C. The firm exits when $\theta^*(Z) \leq \theta < \theta^{**}(Z)$ if its end-of-period net worth is too low for the participation constraint of the financial intermediary to be satisfied: $q < \underline{e}_b(\theta, Z)$ where $\underline{e}_b(\theta, Z)$ is defined by equation (1.2.2).

D. The firm never exits when $\theta \geq \theta^{**}(Z)$ whatever its level of net worth.

PROOF: see Appendix A.

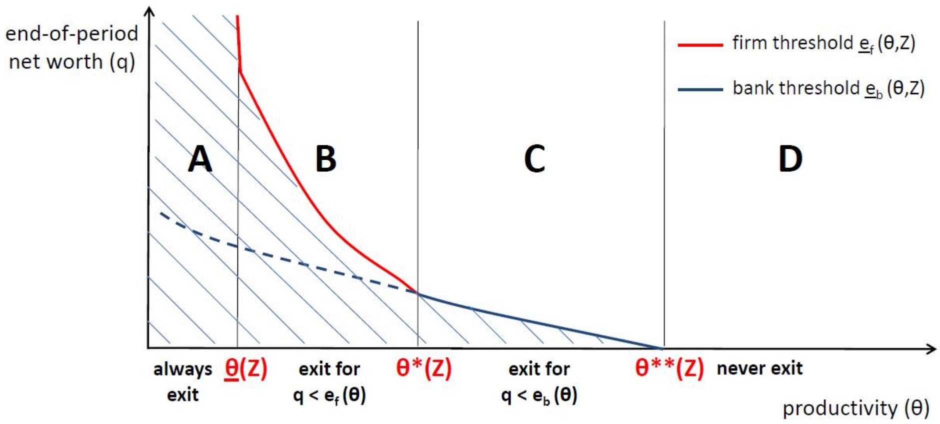


Figure 1.1: Exit frontier

Figure 1.1 represents the four exit regions defined in Proposition 2. All firms with a couple (θ, e) below the downward sloping frontier (solid line) exit the market, whereas all firms with

.

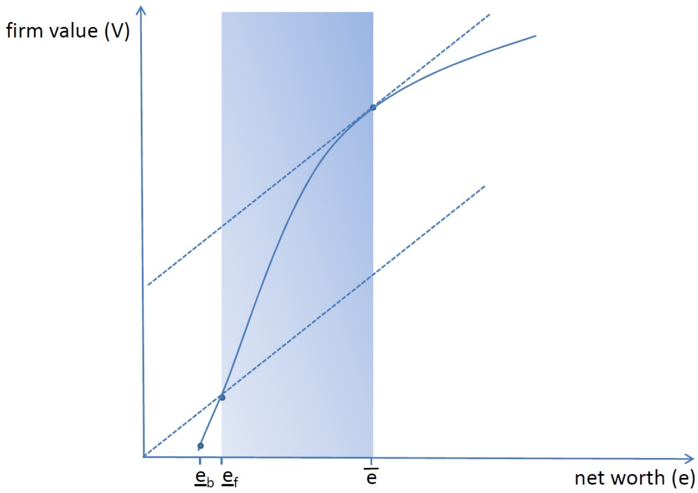


Figure 1.2: Firm value

a couple (θ, e) above the exit frontier produce and stay in the market. Firms with productivity $\underline{\theta}(Z) \leq \theta < \theta^*(Z)$ (region B) exit when their net worth is below $\underline{e}_f$. A low level of net worth raises the borrowing costs of the firm, which may then not be sufficiently profitable to stay in the market. Figure 1.2 shows how the firm's net worth threshold $\underline{e}_f$ is determined. This threshold is the value of net worth for which the firm is indifferent between investing this value into production or distributing it as dividends. Firms with a higher level of productivity always find it profitable to stay in the market. However, firms with productivity $\theta^*(Z) \leq \theta < \theta^{**}(Z)$ (region C) can be forced to exit the market when their net worth is too low for the participation constraint of the financial intermediary to be satisfied. In contrast with the frictionless economy, the firm's decision to exit is sensitive to the non-persistent idiosyncratic shock. An adverse shock ϵ may lead some firms to exit as it may lower their net worth below the thresholds $\underline{e}_b(\theta, Z)$ or $\underline{e}_f(\theta, Z)$. In particular, firms with $\theta < \theta^{**}(Z)$ necessarily exit after defaulting on their debt. However, this is not the case for high productivity firms $\theta \geq \theta^{**}(Z)$ (region D) as the financial intermediary accepts to lend to those firms even after they defaulted. Whatever their level of net worth, firms with $\theta \geq \theta^{**}(Z)$ never exit the market.

The exit thresholds $\underline{e}_b(\theta, Z)$ and $\underline{e}_f(\theta, Z)$ are both decreasing functions of the persistent component of productivity θ (see proof in Appendix A). This implies that high productivity firms have a lower probability to exit the market. This negative correlation between exit and productivity is likely to be reinforced once the endogenous distribution of net worth is accounted for. In fact, high productivity firms have a lower probability to exit also because they tend to accumulate more net worth.

Finally, for all firms with $\theta \geq \underline{\theta}(Z)$, the exit, default and dividend decisions restrict the net worth of continuing firms to $[\underline{e}(\theta, Z), \bar{e}(\theta, Z)]$ with $\underline{e}(\theta, Z) = \max\{0, \underline{e}_b(\theta, Z)\}$ if $\theta \geq \theta^*(Z)$ and $\underline{e}(\theta, Z) = \underline{e}_f(\theta, Z)$ if $\underline{\theta}(Z) \leq \theta < \theta^*(Z)$. The set $[\underline{e}(\theta, Z), \bar{e}(\theta, Z)]$ for firms with $\underline{\theta}(Z) \leq \theta < \theta^*(Z)$ is represented by the shaded area in Figure 1.2.

1.2.4 Entry, stationary distribution and aggregate output

This paper focuses on the consequences of credit market frictions on the exit rate and leaves aside the implications of credit frictions at the entry margin. We assume that the mass of potential entrants is constant. Despite this assumption, the actual number of entrants is endogenous as firms enter the market only when their expected profits are sufficiently high. The net worth e and productivity θ of potential entrants are characterized by the joint distribution ν. The distribution ν of potential entrants, the distributions Φ and F of the productivity shocks, together with the endogenous decision rules of the firms on capital, default, dividends and exit generate an endogenous joint distribution of productivity and net worth ξ. More specifically, these conditions give rise to a mapping Ω that indicates the next period joint distribution of net worth and productivity given the current distribution: $\xi' = \Omega(\xi)$. The stationary joint distribution is the fixed point of the mapping $\xi^* = \Omega(\xi^*)$[9]. We

[9] The distribution reads:

$$\begin{aligned}\xi'(\mathcal{E} \times \Lambda) &= \int_{\underline{\theta}(Z)}^{\theta_{\max}} \int_{\underline{e}(\theta,Z)}^{\bar{e}(\theta,Z)} \left(\int_{\theta' \in \Lambda \cap [\underline{\theta}(Z), \theta_{\max}]} Prob(e' \in \mathcal{E} \cap [\underline{e}(\theta', Z), \bar{e}(\theta', Z)] | \theta, e) dF(\theta'|\theta) \right) d\xi(\theta, e) \\ &\quad + \psi \int_{\Lambda \cap [\underline{\theta}(Z), \theta_{\max}]} \int_{\mathcal{E} \cap [\underline{e}(\theta,Z), \bar{e}(\theta,Z)]} d\nu(\theta, e),\end{aligned}$$

can now use the joint distribution of firms to write the aggregate production in the industry[10]:

$$Y(Z,\xi) = Z \int_{\underline{\theta}(Z)}^{\theta_{\max}} \int_{\underline{e}(\theta,Z)}^{\bar{e}(\theta,Z)} \theta \left[k(e,\theta,Z)\right]^{\alpha} d\xi(e,\theta). \tag{1.2.6}$$

1.3 Distortion of the exit decision

In this section, we show analytically how the exit decision of the firm is modified once credit market frictions are taken into account. We first show that credit market frictions distort the productivity distribution of exiting firms. In the frictionless economy, all exiting firms are less productive than surviving firms. In the credit constrained economy, since exit also depends on firms' net worth, some high productivity but financially fragile firms are forced to exit, while less productive firms may survive. In order to show this distortion, we measure the quality of the selection at the exit margin by the productivity gap between the most productive exiting firm θ^{**} and the less productive surviving firm $\underline{\theta}$. The larger the productivity gap, the lower is the quality of the selection at the exit margin.

PROPOSITION 3. **Imperfect selection**

The quality of the selection at the exit margin is negatively related to the degree of credit market frictions :

$$\frac{d(\theta^{**} - \underline{\theta})}{d\mu} > 0.$$

PROOF: See Appendix A.

---

where ψ is the mass of potential entrants. The associated stationary probability measure exists and is unique as the associated transition function is monotone, has the Feller property and satisfies a mixing condition. The proof is similar to Cooley and Quadrini (2001).

[10] We assume that the fraction of output required for the monitoring process is not destroyed.

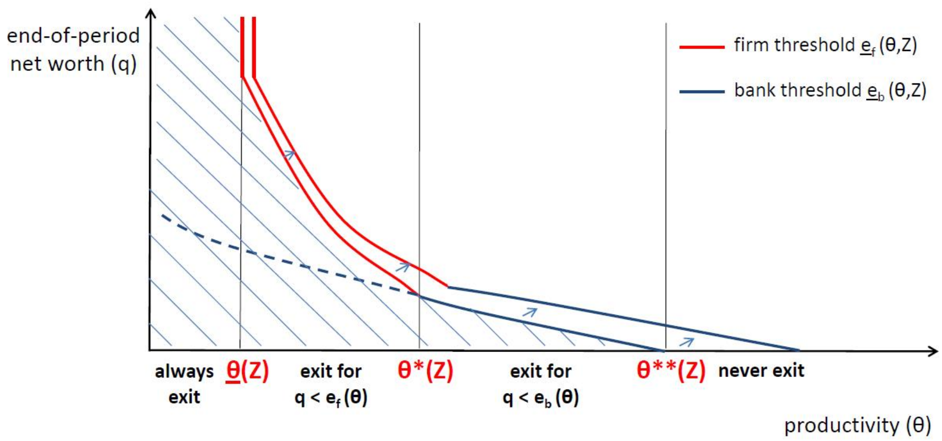


Figure 1.3: Distortion of the exit frontier due to an increase in credit frictions ($\Delta\mu > 0$)

Figure 1.3 provides an illustration of Proposition 3. An increase in the productivity gap between the most productive exiting firm θ^{**} and the least productive surviving firm $\underline{\theta}$ reveals larger distortions in the selection mechanism. This result implies that credit frictions increase the misallocation of resources at the exit margin. In fact, credit market frictions may generate an inefficient reallocation of resources as financial funds flow from high to low productivity firms. Besides modifying the distribution of exiting firms, credit market frictions also affect the mass of firms that exit the market. Proposition 4 describes the impact of credit frictions on the exit thresholds.

PROPOSITION 4. **Firm destruction**

The firms' probability to exit increases with the degree of credit market frictions:

$$\frac{d\underline{\theta}}{d\mu} > 0 \ ; \ \frac{d\underline{e}_f}{d\mu} > 0 \ ; \ \frac{d\underline{e}_b}{d\mu} > 0.$$

PROOF: See Appendix A.

Credit market frictions lead to the exit of firms with low productivity and low net worth that would have been economically viable absent any credit frictions. A higher degree of credit frictions raises firms' borrowing costs, which then increases their probability to exit.

So far, we have highlighted the distortions caused by credit market frictions in the absence of aggregate productivity shocks. Our main objective is to understand the consequences of these distortions on the propagation of aggregate shocks. The complexity of the model prevents us from solving analytically the response of output to a decline in aggregate productivity. However, the distortion in the firms' exit behavior suggests that the fall in output may be amplified by the fluctuations in the exit rate. In fact, higher credit frictions increase the set of firms vulnerable to shocks and may then magnify the impact of an aggregate productivity drop by leading more firms to exit the market. Recall that only firms with productivity $\theta < \theta^{**}$ may exit after an adverse idiosyncratic shock. Similarly, a fall in aggregate productivity $\Delta Z = Z' - Z < 0$ only raises the exit probability of firms with productivity $\theta < \theta^{**}(Z')$. More precisely, the aggregate productivity drop increases the mass of exiting firms by ΔN_X, which is an increasing function of θ^{**}:

$$\Delta N_X = \int_{\underline{\theta}(Z')}^{\theta^{**}(Z')} \int_{\underline{e}(\theta,Z)}^{\underline{e}(\theta,Z')} d\xi^*(e,\theta) + \int_{\underline{\theta}(Z)}^{\underline{\theta}(Z')} \int_{\underline{e}(\theta,Z)}^{\bar{e}(\theta,Z)} d\xi^*(e,\theta).$$

We therefore interpret θ^{**} to be a measure of the industry vulnerability to aggregate shocks.

PROPOSITION 5. **Industry vulnerability**

The industry vulnerability to aggregate shocks increases with the degree of credit market frictions:

$$\frac{d\theta^{**}}{d\mu} > 0.$$

PROOF: See Appendix A.

Proposition 5 suggests that aggregate productivity shocks are likely to generate a stronger response of the exit rate when credit frictions are taken into account. As the response of the economy to a fall in aggregate productivity cannot be derived analytically, we solve the model numerically and further investigate these questions in Section 4.

1.4 Aggregate implications of firm exit under credit market frictions

In this section, we analyze numerically how the exit of firms under credit market frictions modifies the impact of a fall in aggregate productivity. We solve the model using value function iteration. The method is described in Appendix B. We first present the benchmark calibration and describe the firm capital and exit decisions in steady state. Then, we illustrate how credit constraints distort the exit decision of the firm, and analyze its implication on average productivity and on output[11]. Finally, we show how the idiosyncratic volatility σ affects the results.

1.4.1 Benchmark calibration

The model period is one year. We calibrate the parameters on the frictionless economy and then analyze how the introduction of credit market frictions modifies the impact of a fall in aggregate productivity with respect to this benchmark. We normalize the value of aggregate productivity Z to one in the steady state and then consider the impact of a permanent decline in Z of one standard deviation. We use the standard deviation of the innovation to the Solow residual estimated by King and Rebelo (1999) to be 0.0072 for 1947-1996.

Consistently with the business cycle literature, we set the risk-free rate r to 4%, the discount rate β to 0.9606, and the depreciation rate δ to 7%. Following the estimates of Hennessy and Whited (2007), the returns to scale parameter of the production function α is set to 0.7.

We assume a Pareto steady state distribution of productivity θ for incumbents firms, normalize its mean to 0.35 and set the shape parameter to 4 which yields an interquartile ratio of

[11]Recall that we leave aside the implications of credit market frictions on the risk-free interest rate and therefore compare the economy with credit frictions with a frictionless economy characterized by the same risk-free rate.

1.32. This value is line with Del Gatto et al. (2008) who estimate the intra-industry shape parameter of total factor productivity to be between 3 and 4 in the Italian economy, as well as with the estimates of Foster et al. (2008) who reports an average interquartile ratio that ranges between 1.3 and 1.6 in the US manufacturing sector. We then discretize this steady state Pareto distribution into 5 levels of productivity $\{\theta_1, .., \theta_5\}$ and their associated distribution ${\xi^*_{FL}}$[12]. To capture the non linear effects of credit frictions, $\xi^*_{FL}(\theta_i)$ is set so that the discretized productivity levels θ_i are concentrated at the lower end of the distribution.

In the frictionless steady state economy, firms exit when their productivity is equal to θ_0, with θ_0 set to be 30 percent lower than the average productivity of incumbents firms. Together with the fixed cost c, the conditional distribution of productivity shocks $F(\theta'|\theta)$ is then chosen to match the stationary distribution of productivity ξ^*_{FL} and a steady state exit rate of 5%. This value is below usual estimates of the exit rate as we believe traditional measures are likely to overstate the number of exiting firms[13]. Dunne et al. (1989) report a 5-year exit rate of 36% in the US manufacturing sector, which induces a 7.2% annual exit rate, assuming that the number of firms remained constant during these 5 years. OECD firm level data (Scarpetta et al., 2002) exhibit a 8.4% exit rate for the US private sector. We also constrain $F(\theta'|\theta)$ to be decreasing in θ. Because of the discretization, we need to ensure that the aggregate shock yields a realistic increase in the exit rate. This was achieved by limiting the size of $\xi^*_{FL}(\theta_1)$[14] in order to generate a 6% exit rate after the adverse productivity shock. See Appendix C for more details on the calibration procedure.

[12] ξ^*_{FL} is the fixed point of the mapping:

$$\xi'_{FL}(\Lambda) = \int_{\theta\in[\theta_{\min},\theta_{\max}]} \int_{\theta'\in\Lambda\cap[\underline{\theta}_{FL},\theta_{\max}]} dF(\theta'|\theta) d\xi_{FL}(\theta) + \psi_{FL} \int_{\theta\in\Lambda\cap[\underline{\theta}_{FL},\theta_{\max}]} d\nu_\theta(\theta),$$

where ν_θ is the marginal productivity distribution of potential entrants and ψ_{FL} is the mass of potential entrants, normalized to have a unit mass of incumbents firms. θ_i is computed as the expectation of the truncated Pareto distribution in the interval $[m_{i-1}, m_i]$, with $Prob(m_{i-1} < \theta < m_i) = \xi^*_{FL}(\theta_i)$, $i = 1..5$ where m_0 is the scale of the Pareto distribution.

[13] The exit rate is usually measured as the number of firms disappearing from a given dataset. The disappearance could be due to reasons unrelated to exit such as mergers, acquisition, restructuring, etc.

[14] $\xi^*_{FL}(\theta_1) = exit(Z') - 4/5 exit(Z)$ where $exit(Z)$ is the exit rate for a Z level of aggregate productivity.

The remaining parameters ($\underline{e}_0$, $\overline{e}_0$, μ, σ) pertain to the credit constrained economy. Entrants are assumed to be uniformly distributed on $\{\theta_1, ..\theta_5\} \times [\underline{e}_0, \overline{e}_0]$. We choose $\underline{e}_0 = 0$ and $\overline{e}_0 = 5$ and verify that $\overline{e}_0$ is close to the highest level of net worth among incumbents firms. We report in Appendix D the sensitivity of the results to this assumption. We set the monitoring costs μ to match an average bankruptcy cost equal to 10% of capital. This value includes direct costs such as administrative and legal fees, but also indirect costs of bankruptcy linked to the efficiency of debt enforcement. Andrade and Kaplan (1998) estimate these costs to be between 10% and 20% of the firm's capital value. Concerning the dispersion of idiosyncratic shocks, we assume that these shocks are drawn from a normal distribution of mean zero and standard deviation σ truncated on $[-\theta_0, +\theta_0]$ to avoid a negative production for very low levels of the idiosyncratic shock. Since we could not pin down the dispersion of idiosyncratic shocks from standard targets[15], we choose $\sigma = 0.2$ as the benchmark and present in Section 4.5 the results for alternative values of σ. We report the set of values of the benchmark calibration in Table 1.

1.4.2 Steady state capital and exit behavior

Figure 1.4 displays how credit constraints link the firms' capital choice to their level of net worth. A higher level of net worth relaxes the financial intermediary's participation constraint and allows the firm to expand its production scale[16]. Firms with a high level of net worth are not subject to credit constraints and can invest as much as in the frictionless case. Moreover, a creditworthy firm has a lower probability to default on its debt as well as a lower probability to exit the market. An increase in the persistent productivity θ produces similar effects on the exit rate, but barely affects the default rate. As expected, productivity and creditworthiness

[15] In particular, the default rate and risk premium appear to be weakly sensitive to σ: the increase in idiosyncratic risk raises the default probability of firms, but it also increases the exit rate of firms with a high default probability, which tends to reduce the average default rate.

[16] For a low net worth, the capital function is not necessarily monotone. This is due to the non concavity of the firm's income (and hence of the value function) around the default threshold.

Table 1.1: Benchmark calibration

Parameter	Symbol	Value
Discount factor	β	0.9606
Risk-free rate	r	0.04
Depreciation rate	δ	0.07
Returns to scale	α	0.70
Aggregate productivity	Z	1
Aggregate shock	ΔZ	-0.72%
Persistent productivity	$\theta_1, .., \theta_5$	$0.2632, 0.2655, 0.2711, 0.2898, 0.4019$
Exit productivity	θ_0	0.2450
Productivity distribution	ξ^*_{FL}	$0.02, 0.05, 0.10, 0.30, 0.53$
Fixed cost	c	0.49
Idiosyncratic volatility	σ	0.20
Monitoring cost	μ	0.2127
Net worth of entrants	$[\underline{e}_0, \overline{e}_0]$	[0,5]

are highly correlated: an increase in the persistent level of productivity θ shifts the cumulative distribution of net worth to the right and raises the average net worth.

1.4.3 Imperfect selection and average productivity

We now consider a one standard deviation permanent fall in aggregate productivity ($\Delta Z = -0.72\%$) and simulate the impact response of the economy to this shock. In this section, we illustrate how credit constraints distort the selection at the exit margin and analyze its implication on average productivity.

As shown in Proposition 3, credit market frictions distort the productivity distribution of exiting firms. We capture the degree of the imperfect selection by computing the productivity gap between the most productive exiting firm and the less productive surviving firm. In the economy with credit market frictions, the productivity gap is equal to $\theta^{**}/\underline{\theta} - 1$. As this gap widens, the quality of the selection process deteriorates.

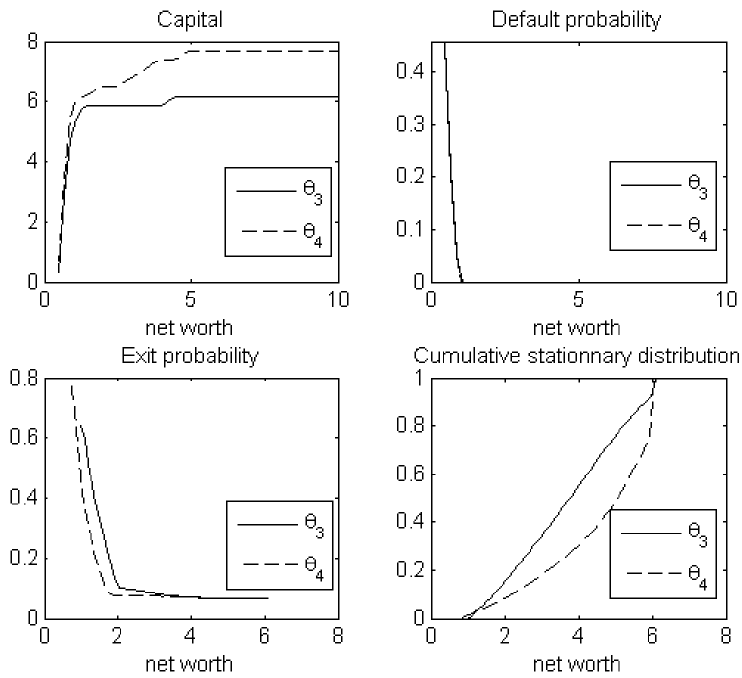


Figure 1.4: The firm's decision rules

Table 1.2: Credit frictions and imperfect selection

	Productivity gap	Average productivity
Frictionless	-0.9%	+0.49%
Credit frictions	51.3%	+0.95%

Note: the average productivity refers to the average of the productivity component θ among incumbents firms and is expressed in deviation from steady state.

Table 1.2 illustrates how the productivity gap is affected by credit market frictions. With credit frictions, some firms are forced to exit though their productivity is 51% higher than some surviving firms. In the frictionless economy, the productivity gap is slightly negative as exiting firms are strictly less productive than surviving firms. In a frictionless economy, the selection effect is usually associated with an increase in average productivity, as recessions reallocate resources from the least productive firms towards more productive ones. This is the intuition behind the cleansing effect of recessions, as emphasized in Caballero and Hammour (1994). Since credit frictions distort the selection mechanism, one could think that this distortion dampens the increase in average productivity. Nevertheless, the results reported in Table 1.2 indicate that this is not the case, as credit market frictions lead to a larger increase in productivity. Though credit market frictions lead to the exit of some high productivity firms, low productivity firms have a higher probability to exit. Since firms that exit are, on average, less productive than surviving firms, the fall in aggregate productivity tends to increase the average productivity of incumbent firms.

Table 1.3: Exit rate

	Z=1	Z =0.9928	Change
Frictionless	5.00%	6.00%	1.00 p.p.
Credit frictions	5.51%	8.00%	2.49 p.p.

The variation in the exit rate, shown in table 1.3, then explains why credit frictions induce a higher increase in idiosyncratic productivity than in the frictionless economy. The results reported in table 1.3 illustrate the increased destruction highlighted in Proposition 4 : the number of exiting firms in the credit constrained economy is higher than in the frictionless economy. More importantly, these results also indicate that credit market frictions amplify the fluctuations in the exit rate as suggested by Proposition 5. Credit market frictions increase the number of firms vulnerable to the aggregate productivity drop. In fact, the fall in aggregate productivity raises the exit rate by 2.49 percentage points in the economy with credit frictions versus a 1 percentage point increase in the frictionless economy. As credit market frictions amplify the increase in the number of exiting firms which are, on average less productive than surviving firms, they also amplify the increase in average idiosyncratic productivity.

1.4.4 Amplification at the exit margin

In this section, we investigate the consequences of the fluctuations in the exit rate on aggregate output. Our objective is to analyze how credit market frictions make the industry more vulnerable to a fall in aggregate productivity. In order to focus on this new mechanism, we leave out the effects of the endogenous change in net worth. In fact, the consequences of the fall in net worth are well known at the intensive margin and are likely to generate a similar mechanism at the extensive margin. We therefore consider the impact response of aggregate output, when the fall in aggregate productivity has not produced yet any effect on the firms' net worth. We decompose the impact of the decline in aggregate productivity (Z drops to Z') on aggregate output as follows:

$$\Delta Y = \underbrace{\Delta Z \int_{\underline{\theta}'}^{\theta_{\max}} \int_{\underline{e}'}^{\bar{e}'} \theta k'^{\alpha} d\xi^*}_{\text{Direct effect}} + \underbrace{Z \left[\int_{\underline{\theta}'}^{\theta_{\max}} \int_{\underline{e}'}^{\bar{e}'} \theta k'^{\alpha} d\xi^* - \int_{\underline{\theta}'}^{\theta_{\max}} \int_{\underline{e}'}^{\bar{e}} \theta k^{\alpha} d\xi^* \right]}_{\text{Intensive margin}}$$
$$\underbrace{- Z \left[\int_{\underline{\theta}'}^{\theta^{**}} \int_{\underline{e}}^{\underline{e}'} \theta k^{\alpha} d\xi^* + \int_{\underline{\theta}}^{\underline{\theta}'} \int_{\underline{e}}^{\bar{e}} \theta k^{\alpha} d\xi^* \right]}_{\text{Exit margin}}$$

The fall in aggregate productivity lowers the productivity of incumbents firms (*direct effect*), leading to a reduction in their investment (*intensive margin*), and inducing some firms that have become unprofitable to exit (*exit margin*)[17].

Table 1.4: A decomposition of the output loss upon impact

	Aggregate production	Direct effect	Intensive margin	Exit margin
Frictionless	3.09%	0.70%	1.66%	0.70%
Credit frictions	3.63%	0.70%	1.54%	1.39%

[17] Note that the exit margin also includes the exit of potential entrants. The number of actual entrants also declines as fewer potential entrants find it profitable to stay in the market and immediately exit.

Table 1.4 reports the result of this output decomposition for the economy with credit frictions and the frictionless economy. The aggregate productivity drop causes a decline in output about 20% larger in the economy with credit frictions. The decomposition shows that the overall impact on output masks a larger effect at the exit margin. In the economy with credit frictions, the output loss induced by business shutdowns is twice as big as in the frictionless economy. In the latter, the exit margin generates a 0.7% loss in output while inducing a 1.39% loss when credit frictions are accounted for. By contrast, the output decomposition points to a dampening effect along the intensive margin. The intensive margin reduces output by 1.54% in the economy with credit frictions and by 1.66% in the frictionless economy. Firms reduce less their capital when they face tighter credit constraints. This is the outcome of two counteracting effects. On the one hand, the fall in aggregate productivity raises the financing costs of credit constrained firms, which tends to exacerbate the decline in investment. On the other hand, the marginal cost of financing also decreases as firms reduce the amount borrowed, which tends to dampen the impact of the aggregate shock. As reported in Table 1.4, the second effect dominates, and therefore leads to an overall dampening effect[18]. Following the fall in aggregate productivity, creditworthy firms decrease their investment further than firms with low net worth. This result extends to the comparison of the frictionless economy and the economy with credit frictions. For a given distribution of net worth, a fall in aggregate productivity reduces capital further in the frictionless economy. As highlighted by the standard financial accelerator, the intensive margin may however amplify aggregate fluctuations once the fall in net worth is accounted for. In that respect, the amplification at the exit margin is more robust than the financial accelerator as it does not hinge on the sensitivity of firms' net worth to aggregate shocks. We show that, when the balance sheet effect is shut down, credit frictions lead to a substantially larger output loss at the exit margin. Moreover, table 1.4 indicates that the exit margin accounts for a significant part of output fluctuations, especially in presence of credit frictions. In that case, the exit margin contributes as much

[18]A similar dampening effect is at work in Carlstrom and Fuerst (1997). On impact, output respond less to an aggregate productivity fall in the economy with credit frictions. However, after the initial period, the decline in net worth raises the cost of borrowing and may reduce output further.

as the intensive margin to the decline in aggregate output. All in all, these results suggest that the exit margin is an important channel for understanding the aggregate implications of credit market frictions.

1.4.5 The role of idiosyncratic productivity

In this section we show how the volatility of the non-persistent component of productivity ϵ affects the amplification mechanism at the exit margin. We use the benchmark parameters presented in section 4 and study the response of the economy to a fall in aggregate productivity when the idiosyncratic volatility σ varies from 0.10 to 0.40. Table 1.5 shows that the fluctuations in aggregate output and average productivity are amplified when idiosyncratic volatility is high. An increase in idiosyncratic volatility exacerbates the sensitivity of the exit rate to aggregate productivity. Following the aggregate productivity drop, the negative output response is 10% larger than the frictionless economy when $\sigma = 0.10$, and 28% larger when $\sigma = 0.40$. For high levels of idiosyncratic volatility, the contribution of the exit margin to the output loss is more important than the contribution of the intensive margin. These results show that the interaction between credit market frictions and the volatility of idiosyncratic shocks is crucial for the amplification of aggregate shocks.

These results also illustrate the importance of idiosyncratic volatility for aggregate dynamics in the economy with credit market frictions. In the frictionless economy, the volatility of the non-persistent idiosyncratic productivity ϵ does not matter for aggregate fluctuations. The firm exit and capital decisions are driven by expected profits which are not affected by the non-persistent component of productivity. By contrast, the volatility of ϵ plays a key role when firms face credit constraints. A high idiosyncratic volatility reduces the expected income of the financial intermediary and raises the firm's financing costs and exit probability. The amplification at the exit margin therefore results from the interplay of credit market frictions and idiosyncratic volatility.

Table 1.5: The role of idiosyncratic volatility

	Aggregate production	Direct effect	Intensive margin	Exit margin	Average productivity
$\sigma = 0.10$	3.39%	0.70 %	1.61%	1.08%	0.74 %
$\sigma = 0.15$	3.48%	0.70 %	1.54%	1.24%	0.85 %
$\sigma = 0.20$	3.63%	0.70 %	1.54%	1.39%	0.95 %
$\sigma = 0.25$	3.78%	0.70 %	1.53%	1.55%	1.06 %
$\sigma = 0.30$	3.81%	0.70%	1.42%	1.69%	1.16 %
$\sigma = 0.40$	3.92%	0.70%	1.55%	1.67%	1.14 %

1.5 Conclusion

In this paper, we analyze the exit decision of firms subject to credit constraints, and then investigate its implication for the propagation of business cycle fluctuations. We show that, in the presence of credit frictions, the extensive margin is an important channel for the amplification of aggregate fluctuations: credit market frictions exacerbate the fluctuations in the exit rate by increasing the number of firms vulnerable to aggregate shocks. Unlike the standard financial accelerator, this new amplification mechanism does not depend on the sensitivity of firms' net worth to aggregate shocks. Note that the industry vulnerability and the financial accelerator mechanisms are complementary and an accurate estimate of the business cycle implications of credit frictions should take into account both mechanisms. This work is left for further research.

This paper also emphasizes how credit frictions distort the selection of exiting firms: high productivity but financial fragile firms may exit during recessions while low productivity and high net worth firms may survive. This distortion suggests that credit frictions lead to a inefficient reallocation of resources during recessions, as resources may flow from high to low productivity firms. Interestingly, this imperfect selection does not weaken the cleansing effect of recessions. On the contrary, the presence of credit frictions contributes to a higher increase in average idiosyncratic productivity during recessions. Finally, our results support the recent line of research (Bilbiie et al., 2007; Clementi and Palazzo, 2010) which takes into account firm dynamics in the study of business cycles. Our results suggest that the exit behavior of firms plays an important role in explaining the response of average productivity and aggregate output to business cycle shocks.

Appendix A

Financial intermediary net worth threshold $\underline{e}_b$

Let us define the net income of the financial intermediary as $B(e,k,\bar{\epsilon})$ where

$$B(e,k,\bar{\epsilon}) = Z[\theta + G(\bar{\epsilon})]k^{\alpha} + (1-\delta)k - \mu k^{\alpha}\Phi(\bar{\epsilon}) - (1+r)(k+c-e).$$

The participation constraint of the financial intermediary is not satisfied when the firm's net worth is below $\underline{e}_b(\theta, Z)$, where this threshold is defined by:

$$B(\underline{e}_b, k_b, \bar{\epsilon}_b) = 0,$$

with $(k_b, \bar{\epsilon}_b)$ being the values of capital and default threshold that maximize the income of the financial intermediary. Note that $\bar{\epsilon}_b$ can be a corner solution depending on the shape of Φ:

$$\bar{\epsilon}_b = \begin{cases} \epsilon_{\max} & \text{if } \frac{\Phi'}{\Phi}(\epsilon) \text{ is decreasing in } \epsilon \\ \text{such that } Z(1-\Phi(\bar{\epsilon})) = \mu\Phi'(\bar{\epsilon}) & \text{otherwise.} \end{cases}$$

As we assume $\epsilon_{\min} > -\theta_{\min}$, $Z(G(\bar{\epsilon}_b) + \theta) - \mu\Phi(\bar{\epsilon}_b) > 0$ and the financial intermediary's income is a concave function of capital. Therefore k_b is an interior solution:

$$k_b = \left(\frac{\alpha\left[Z(G(\bar{\epsilon}_b) + \theta) - \mu\Phi(\bar{\epsilon}_b)\right]}{\delta + r} \right)^{\frac{1}{1-\alpha}}.$$

Since the income of the financial intermediary $B(e,k,\bar{\epsilon})$ is strictly increasing in the net worth e, there is a unique net worth threshold $\underline{e}_b$ such that $B(\underline{e}_b, k_b, \bar{\epsilon}_b) = 0$.

PROOF of Proposition 1. Exit when the financial intermediary refuses the loan

The firm exits the market when the participation constraint of the financial intermediary is not satisfied, that is when the end-of-period net worth is too low $q < \underline{e}_b(\theta, Z)$. Indeed, we show that $e_b(\theta, Z) \leq c$, $\forall \theta \in [\theta_{\min}, \theta_{\max}]$ and therefore the firm which is rationed from the credit market cannot self finance its fixed operating cost.

Recall that $\underline{e}_b(\theta, Z)$ is defined by the following equation:

$$\max_{(k,\bar{\epsilon})} \left(Z[\theta + G(\bar{\epsilon})]k^{\alpha} - (\delta + r)k - \mu k^{\alpha}\Phi(\bar{\epsilon})\right) = (1 + r)(c - \underline{e}_b)$$

Notice that $\max_{(k,\bar{\epsilon})} \left(Z[\theta + G(\bar{\epsilon})]k^{\alpha} - (\delta + r)k - \mu k^{\alpha}\Phi(\bar{\epsilon})\right) \geq 0$ since the financial intermediary can always choose $k = 0$ and have 0. It follows that $c - \underline{e}_b(\theta, Z) \geq 0$.

PROOF of existence and uniqueness of the value function

Consider the problem of the firm defined as:

$$T(V)(e, \theta, Z) = \max_{(k,\bar{\epsilon})\in\Gamma(e)} \mathbb{E}_{\theta} \left\{ \int I(q)q + (1 - I(q)) \max \left[q, \max_{e'\in\Upsilon(q)} \left(q - e' + \beta V(e', \theta, Z) \right) \right] d\Phi(\epsilon) \right\}$$

with:

$$I(q) = \begin{cases} 0 & \text{if } q \geq \underline{e}_b \\ 1 & \text{if } q < \underline{e}_b. \end{cases}$$

$$\Gamma(e) = \{(k, \bar{\epsilon}) \in \mathbb{R}^+ \times [\epsilon_{\min}, \epsilon_{\max}] : Z[\theta + G(\bar{\epsilon})]k^{\alpha} + (1 - \delta)k - \mu k^{\alpha}\Phi(\bar{\epsilon}) \geq (1 + r)(k + c - e)\}$$

$$\Upsilon(q) = \{e' \in \mathbb{R} : \underline{e}_b \leq e' \leq q\}$$

$$q = \max \{Zk^{\alpha}(\epsilon - \bar{\epsilon}); 0\}.$$

In the following, we prove that there exist a unique function V that satisfies the functional equation $V = T(V)$ assuming the idiosyncratic productivity level θ is constant. The proof extends to non-permanent level of $\theta \in [\theta_{\min}, \theta_{\max}]$.

First note that the participation constraint of the financial intermediary and Assumption 1 ($\beta(1 + r) < 1$) limits the space for net worth of continuing firms to $X = [\underline{e}_b(\theta, Z), \bar{e}(\theta, Z)]$. Because $\beta(1 + r) < 1$, entrepreneurs will not always reinvest their net worth in the firm, and will start distributing dividend when their net worth is sufficiently high. In particular, there exists a threshold $\bar{e}(\theta, Z)$ above which the firm will stop accumulating net worth.

We then show that the value of the continuing firm $V : X \to \mathbb{R}^+$, is necessarily bounded. The value of the firm is the discounted sum of the income from production and/or investing in the safe asset. As the decreasing returns to scale technology put an upper bound on the profits of the firm and Assumption 1 limits net worth accumulation, the value of the firm is bounded. This also means that the function resulting from the mapping TV is bounded and T maps the space of bounded functions $B(X)$ into itself. Then, we observe that the operator T is a contraction since it satisfies the Blackwell conditions of monotonicity and discounting. The condition for discounting is verified as $\forall V \in B(X)$,

$$\begin{aligned} T(V + c) &= \max_{(k,\bar{\epsilon}) \in \Gamma(e)} \left\{ \int I(q)q + (1 - I(q)) \max \left[q, \max_{e' \in \Upsilon(q)} (q - e' + \beta V(e', \theta, Z) + \beta c) \right] d\Phi(\epsilon) \right\} \\ &\leq \max_{(k,\bar{\epsilon}) \in \Gamma(e)} \left\{ \int I(q)q + (1 - I(q)) \max \left[q + \beta c, \max_{e' \in \Upsilon(q)} (q - e' + \beta V(e', \theta, Z) + \beta c) \right] d\Phi(\epsilon) \right\} \\ &\leq TV + \beta c, \end{aligned}$$

where $c > 0$, and $0 < \beta < 1$ by definition.

Since $B(X)$ is a complete metric space (see for example, Godement (2001)), the Contraction Mapping Theorem applies and the operator T has a unique fixed point V, which is bounded.

Let us show that V is a continuous function. The participation constraint of the bank generates a discontinuity in the firm's end-of-period outcome. For $q < \underline{e}_b$, the value of the firm is simply q and if $q = \underline{e}_b$, the firm can invest in production and obtains the value $\beta V(\underline{e}_b)$ which has no reason to coincide with q. However, we can show that, though the end-of period value of the firm is discontinuous, the expectation of this value is a continuous function of the threshold $\frac{e_b}{Zk^\alpha} + \bar{\epsilon}$ below which the firm cannot borrow from the financial intermediary. This appears clearly when rewriting the value function as follows:

$$V(e,\theta,Z) = \max_{(k,\bar{\epsilon})\in\Gamma(e)} \left\{ \int_{\epsilon_{\min}}^{\frac{e_b}{Zk^\alpha}+\bar{\epsilon}} q d\Phi(\epsilon) + \int_{\frac{e_b}{Zk^\alpha}+\bar{\epsilon}}^{\epsilon_{\max}} \max\left[q, \max_{e'\in\Upsilon(q)} \left(q - e' + \beta V(e',\theta,Z)\right)\right] d\Phi(\epsilon) \right\}.$$

Despite the discontinuity at $\underline{e}_b$ in the end-of-period value of the firm, the discontinuity disappears in the objective function of the continuing firm. Furthermore, note that if V is a continuous function, then the objective function of the firm deciding its next period net worth e' is also continuous. Because the correspondence $\Upsilon(q)$ that describes the feasibility constraint for e' is non-empty, continuous and compact-valued, the theorem of the maximum ensures that the maximum exists and the function resulting from this dividend choice is continuous. As the threshold $\frac{e_b}{Zk^\alpha} + \bar{\epsilon}$ below which the firm cannot borrow from the financial intermediary is a continuous function of e, the objective function of the firm deciding its capital level is the sum of two continuous functions and is therefore a continuous function of e. Using again the theorem of the maximum, we can finally show that the function resulting from the mapping $T(V)(e)$ is continuous since the correspondence Γ that describe the feasibility constraint for k and $\bar{\epsilon}$ is non-empty, continuous[19]. and compact-valued. This means that T maps the space of continuous and bounded functions into itself, $T : C(X) \to C(X)$. As $C(X)$ is a closed subset of the complete metric space of bounded functions $B(X)$, the fixed point V is a continuous function by the corollary of the contraction mapping theorem[20].

[19]The continuity of Γ derives from the continuity of the participation constraint of the bank $B(e,k,\bar{\epsilon})$ and from the uniqueness of $\bar{\epsilon}_b$ guaranteed by Assumption 2.

[20]See Corollary 1 of Theorem 3.2 in Stokey et al. (1989).

Let us now characterize more precisely the value function V. Notice that Υ and Γ are increasing correspondences: $q_1 \leq q_2$ implies $\Upsilon(q_1) \subseteq \Upsilon(q_2)$ and $e_1 \leq e_2$ implies $\Gamma(e_1) \subseteq \Gamma(e_2)$. A higher net worth e relaxes the credit constraint of the firm and allows the firm to reach a higher end-of-period net worth q. This means that the period return function q is strictly increasing in e. This implies that T maps the space of bounded continuous and strictly increasing functions into itself. As this is a closed subset of the space of bounded functions $B(X)$, V is a strictly increasing function of e. By the same reasoning, we can show that V is also strictly increasing in θ[21].

To further characterize the value function, let us write the Lagrangian of the firm's problem:

$$\mathcal{L}(k, \bar{\epsilon}, \lambda) = \int_{\epsilon_{\min}}^{\frac{e_b}{Zk^\alpha}+\bar{\epsilon}} \left[Zk^\alpha(\epsilon - \bar{\epsilon})\right] d\Phi(\epsilon)$$

$$+ \int_{\frac{e_b}{Zk^\alpha}+\bar{\epsilon}}^{\epsilon_{\max}} \max\left\{Zk^\alpha(\epsilon - \bar{\epsilon}), \max_{e' \in \Upsilon(q)} Zk^\alpha(\epsilon - \bar{\epsilon}) - e' + \beta V(e', \theta, Z)\right\} d\Phi(\epsilon)$$

$$+\lambda g(k, \epsilon, e)$$

with $g(k, \bar{\epsilon}, e) = \big(Z[\theta + G(\bar{\epsilon})]k^\alpha + (1-\delta)k - \mu k^\alpha \Phi(\bar{\epsilon}) - (1+r)(k + c - e)\big)$.

In the following, we assume that V is differentiable. We can then compute its second order derivative to show that V is a concave function. Using the envelop theorem, we can write:

$$\frac{\partial^2 V}{\partial e^2} = \frac{\partial \lambda}{\partial e}(1+r)$$

We then take the total differential of the financial intermediary's participation constraint with

[21] In the case of stochastic productivity θ, we further assume that the transition function $F(\theta'|\theta)$ is strictly decreasing in θ.

respect to λ and e:

$$\frac{d\lambda}{de} = -\frac{1+r}{\frac{\partial}{\partial\lambda}g(k(\lambda),\epsilon(\lambda),e)}$$

Applying the implicit function theorem to the first order optimality conditions, we can write:

$$\frac{\partial}{\partial\lambda}g(k(\lambda),\bar{\epsilon}(\lambda),e) = -\begin{pmatrix} \frac{\partial g}{\partial\bar{\epsilon}} \\ \\ -\frac{\partial g}{\partial k} \end{pmatrix}' H \begin{pmatrix} \frac{\partial g}{\partial\bar{\epsilon}} \\ \\ -\frac{\partial g}{\partial k} \end{pmatrix},$$

where H is the hessian matrix of the Lagrangian function[22]. As long as the second order optimality conditions of the firm problem are satisfied, H is negative definite. Then $\frac{\partial}{\partial\lambda}g(k(\lambda),\bar{\epsilon}(\lambda),e) > 0$ and V is a concave function of e.

The concavity of the firm value function also allows us to characterize the dividend decision of the firm. There exists a unique threshold $\bar{e}(\theta, Z)$ above which the firm decides to distribute some dividends. This optimal threshold is given by:

$$\beta\frac{\partial V}{\partial e}(\bar{e},\theta,Z) = 1. \tag{A.1}$$

Therefore, the optimal next period net worth is $e' = \min[q, \bar{e}(\theta, Z)]$.

PROOF of Proposition 2. Exit conditions

Under Assumptions 1 and 2, let us show that for a given Z, the thresholds $\underline{\theta}(Z)$, $\theta^*(Z)$ and $\theta^{**}(Z)$ exist and are unique.

- No exit threshold θ^{**}:

  We have already shown that firms exit the market when the participation constraint of the financial intermediary is not satisfied. However, for some high productivity firms the

[22] $H = \begin{pmatrix} \frac{\partial\mathcal{L}^2}{\partial\bar{\epsilon}^2} & \frac{\partial\mathcal{L}^2}{\partial\bar{\epsilon}\partial k} \\ \frac{\partial\mathcal{L}^2}{\partial\bar{\epsilon}\partial k} & \frac{\partial\mathcal{L}^2}{\partial k^2} \end{pmatrix}$

participation constraint of the financial intermediary is always satisfied. The financial intermediary accepts to lend to firms with $\theta \geq \theta^{**}$ whatever their level of net worth. Therefore, these firms never exit because they are not sufficiently creditworthy. The no exit threshold θ^{**} is characterized by:

$$Z\left[\theta^{**} + G(\bar{\epsilon}_b)\right] k_b^{\alpha} - \mu k_b^{\alpha} \Phi(\bar{\epsilon}_b) - (\delta + r)k_b - (1+r)c = 0. \tag{A.2}$$

Let us show that this threshold is unique. Denote $\hat{\theta} = \frac{\mu}{Z}\Phi(\bar{\epsilon}_b) - G(\bar{\epsilon}_b)$, the level of productivity below which the net income of the financial intermediary is decreasing in k. The left hand side of the no exit threshold condition is strictly increasing in θ for all $\theta > \hat{\theta}$. Furthermore, as this expression is negative for $\theta = \hat{\theta}$, this implies that the threshold θ^{**} is unique.

- Credit market exit threshold θ^* :

Firms also exit if their participation constraint is not satisfied. Using the optimal dividend decision (Equation A.1), firms exit when:

$$e' > \beta V(e', \theta, Z), \qquad \text{with } e' = \min[q, \bar{e}(\theta, Z)].$$

Note that if the firm is sufficiently productive, their participation constraint is always satisfied. Recall $V(e, \theta, Z)$ is defined on $[\underline{e}_b(\theta, Z), \bar{e}(\theta, Z)]$. Then, if $\underline{e}_b(\theta, Z) < \beta V(\underline{e}_b, \theta, Z)$, the firm finds it profitable to stay in the market whatever its level of net worth. The productivity threshold θ^* above which firms always satisfy their participation constraint is characterized by:

$$\underline{e}_b(\theta^*, Z) = \beta V(\underline{e}_b(\theta^*, Z), \theta^*, Z). \tag{A.3}$$

Above this threshold, firms exit only when the participation constraint of the financial intermediary is not satisfied. To show that this threshold is unique, we start with the fact that a firm with productivity $\theta_0 \leq \underline{\theta}_{FL}$ is not profitable and therefore exits the market in the frictionless economy but also in the credit constrained economy (as the financing cost faced by the firm is higher in this case). Therefore: $\underline{e}_b(\theta_0, Z) > \beta V(\underline{e}_b(\theta_0, Z), \theta_0, Z)$. We now need to show that $\underline{e}_b(\theta, Z)$ is a decreasing function and $V(\underline{e}_b(\theta, Z), \theta, Z)$ a strictly increasing function of θ.

Recall that the threshold $\underline{e}_b(\theta, Z)$ below which the financial intermediary refuses to loan any funds is defined by:

$$Z[\theta + G(\bar{\epsilon}_b)]k_b^\alpha + (1-\delta)k_b - \mu k_b^\alpha \Phi(\bar{\epsilon}_b) = (1+r)(k_b + c - \underline{e}_b).$$

Taking the total differential of this equation indicates that firms with a high productivity are less frequently rationed from the market: $d\underline{e}_b/d\theta \leq 0$.

We now need to show that $V(\underline{e}_b(\theta), \theta, Z)$ is increasing in θ.

Using the envelop theorem, it comes:

$$\begin{aligned}
\frac{\partial V}{\partial e} = \frac{\partial \mathcal{L}}{\partial e} &= \lambda(1+r) \\
\frac{\partial V}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial \theta} &= \lambda Z k^\alpha + \beta \int_{\frac{\underline{e}_b}{Zk^\alpha}+\bar{\epsilon}}^{\epsilon_{\max}} \frac{\partial V}{\partial \theta}(e', \theta, Z) \mathbb{1}_{e' < \beta V(e', \theta, Z)} d\Phi(\epsilon) \\
&+ \frac{1}{Zk^\alpha}\frac{d\underline{e}_b}{d\theta}\left(\underline{e}_b - \beta V(\underline{e}_b, \theta, Z)\right)\Phi'\left(\frac{\underline{e}_b}{Zk^\alpha} + \bar{\epsilon}\right)\mathbb{1}_{\underline{e}_b < \beta V(\underline{e}_b, \theta, Z)}.
\end{aligned}$$

Then, it follows that $V(\underline{e}_b(\theta), \theta, Z)$ is strictly increasing in θ,

$$\begin{aligned}
\frac{dV}{d\theta}(\underline{e}_b(\theta, Z), \theta, Z) &= \frac{\partial V}{\partial e}(\underline{e}_b, \theta, Z)\frac{d\underline{e}_b}{d\theta} + \frac{\partial V}{\partial \theta}(\underline{e}_b, \theta, Z) \\
&= -\lambda(1+r)\frac{Zk_b^\alpha}{1+r} + \lambda Z k_b^\alpha + \beta \int_{\frac{\underline{e}_b}{Zk^\alpha}+\bar{\epsilon}}^{\epsilon_{\max}} \frac{\partial V}{\partial \theta}(e', \theta, Z) \mathbb{1}_{e' < \beta V(e', \theta, Z)} d\Phi(\epsilon) \\
&+ \frac{1}{Zk_b^\alpha}\frac{Zk_b^\alpha}{1+r}\left(\beta V(\underline{e}_b, \theta, Z) - \underline{e}_b\right)\Phi'\left(\frac{\underline{e}_b}{Zk_b^\alpha} + \bar{\epsilon}\right)\mathbb{1}_{\underline{e}_b < \beta V(\underline{e}_b, \theta, Z)} \\
&> 0,
\end{aligned}$$

where the last line follows from the fact that V is strictly increasing in θ.

- Full exit threshold $\underline{\theta}$:

  Low productivity firms always exit whatever their level of net worth. We can find a productivity threshold $\underline{\theta}$ below which the participation constraint of the firm is never satisfied. As $e' \leq \bar{e}(\underline{\theta}, Z)$ if $\bar{e}(\underline{\theta}, Z) = \beta V(\bar{e}(\underline{\theta}, Z), \underline{\theta}, Z)$ the firm never finds it profitable to stay in the market. The threshold $\underline{\theta}$ is therefore defined as:

$$\bar{e}(\underline{\theta}, Z) = \beta V(\bar{e}(\underline{\theta}, Z), \underline{\theta}, Z). \tag{A.4}$$

  Firms with productivity $\theta_0 \leq \underline{\theta}_{FL}$ are not profitable and therefore exit the market : $\bar{e}(\theta_0, Z) > \beta V(\bar{e}(\theta_0, Z), \theta_0, Z)$. We complete the proof by showing that the dividend threshold $\bar{e}$ increases with θ less than the value function. Using the dividend decision condition (Equation A.1), we can show:

$$\begin{aligned} \beta \frac{dV}{d\theta}(\bar{e}(\theta, Z), \theta, Z) &= \beta \frac{\partial V}{\partial e}(\bar{e}, \theta) \frac{d\bar{e}}{d\theta} + \beta \frac{\partial V}{\partial \theta}(\bar{e}, \theta) \\ &= \frac{d\bar{e}}{d\theta} + \beta \frac{\partial V}{\partial \theta}(\bar{e}, \theta) \\ &> \frac{d\bar{e}}{d\theta}. \end{aligned}$$

Firms' net worth threshold $\underline{e}_f$

For $\underline{\theta} \leq \theta < \theta^*$, firms exit if their participation constraint is not satisfied: they exit if $e < \underline{e}_f(\theta, Z)$ with $\underline{e}_f(\theta, Z)$ defined by:

$$\underline{e}_f = \beta V(\underline{e}_f, \theta, Z). \tag{A.5}$$

Given θ, we show that this threshold is unique by observing that $\beta \frac{dV(e,\theta,Z)}{de} > 1$ as long as $e < \bar{e}(\theta, Z)$ and $\underline{e}_b(\theta, Z) > \beta V(\underline{e}_b(\theta), \theta, Z)$ for any $\underline{\theta} \leq \theta < \theta^*$.

Furthermore we can show that the exit threshold $\underline{e}_f$ is decreasing in θ as $\frac{d\underline{e}_f}{d\theta} = \frac{\beta \frac{\partial V}{\partial \theta}}{1 - \beta \frac{\partial V}{\partial e}}$.

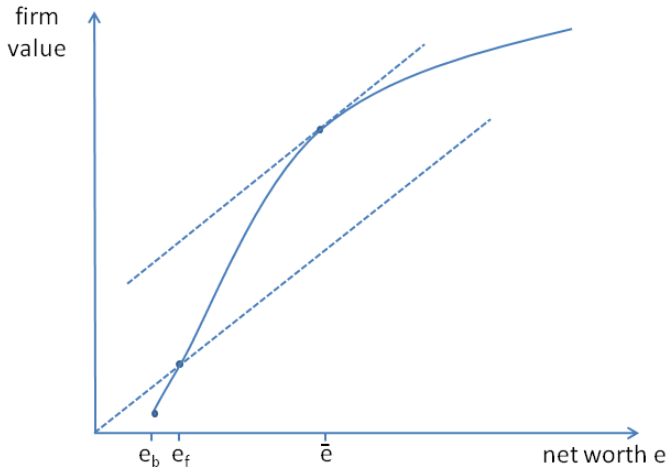


Figure A.1: Firms' net worth threshold $\underline{e}_f$

Figures A.1, A.2 and A.3 illustrate how the thresholds $\underline{e}_f$, $\underline{\theta}$ and θ^* are determined.

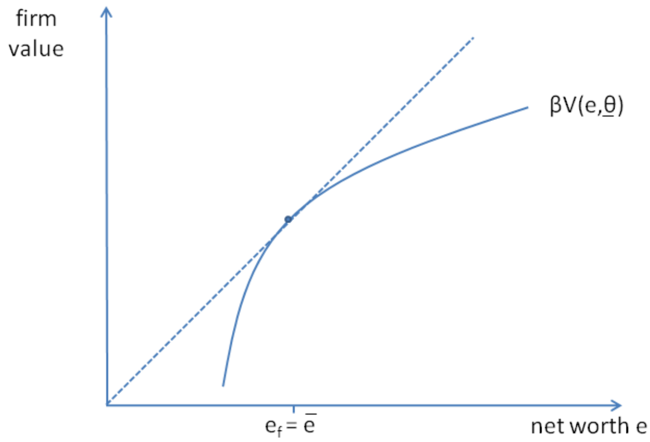


Figure A.2: Productivity threshold $\underline{\theta}$

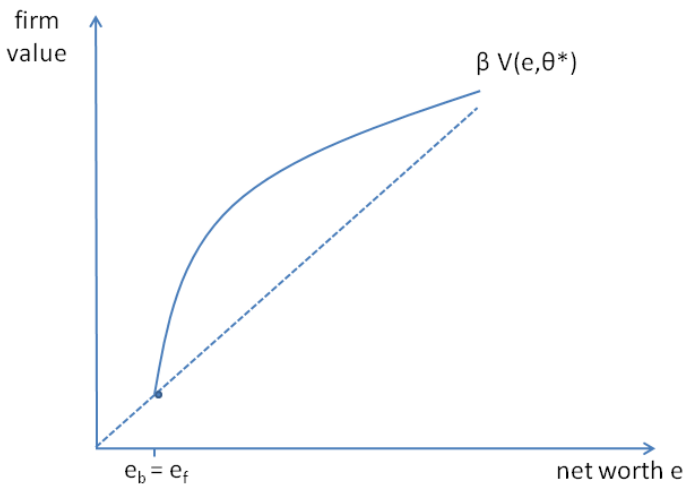


Figure A.3: Productivity threshold θ^*

PROOF of Propositions 3, 4 and 5:

The no exit threshold $\theta^{**}(Z)$, the credit market exit threshold $\theta^{*}(Z)$ and the full exit threshold $\underline{\theta}(Z)$ are defined respectively by Equations (A.2), (A.3) and (A.4). Taking the total differential of Equations (A.2), (A.3), (A.4) and (A.5) yields the following results:

$$\frac{d\underline{\theta}}{d\mu} = -\frac{\frac{\partial V}{\partial \mu}(\bar{e}, \underline{\theta}, Z)}{\frac{\partial V}{\partial \theta}(\bar{e}, \underline{\theta}, Z)} > 0;$$

$$\frac{d\underline{e}_b}{d\mu} = \frac{\Phi(\bar{\epsilon}_b) k_b^{\alpha}}{1+r} > 0;$$

$$\frac{d\theta^{**}}{d\mu} = \Phi(\bar{\epsilon}_b) > 0;$$

$$\frac{d\underline{e}_f}{d\mu} = \beta \frac{\frac{\partial V}{\partial \mu}(\underline{e}_f, \theta, Z)}{1 - \beta \frac{\partial V}{\partial e}(\underline{e}_f, \theta, Z)} > 0.$$

Appendix B: Numerical procedure

The model is solved using value function iteration on the discretized state space, using splines to approximate between grid points.

1. Discretize the shocks ϵ.

2. Start by choosing a grid for the net worth of firms with productivity θ_1. The lower point is initialized at $\max\{\underline{e}_b(Z, \theta_1), 0\}$ and the upper bound at a guess $\bar{e}$.

3. Choose a grid for k. The lower point is initialized at 0 and the upper bound at the frictionless level.

4. Compute for each e, the maximum level of capital that the firm can borrow $k_{\max}(e)$, with $k_{\max}(e)$ defined by:

$$Z[\theta + G(\bar{\epsilon}_b)]k_{\max}^{\alpha} + (1-\delta)k_{\max} - \mu k_{\max}^{\alpha}\Phi(\bar{\epsilon}_b) = (1+r)(k_{\max} + c - e)$$

   where $\bar{\epsilon}_b$ maximizes the income of the financial intermediary.

5. Compute the default threshold $\bar{\epsilon}(k, e)$ for $k < k_{\max}(e)$

6. For each ϵ, compute the end-of-period net worth $q = (\epsilon - \bar{\epsilon}(k, e))Zk^{\alpha}$

7. Compute the solution of the static problem to be used as the initial point for the value function $V^0(e, \theta, Z)$.

8. Repeat steps 2 to 7, for the other levels of productivity θ.

9. For each ϵ, each θ' and each k, solve for the optimal exit and dividend decision d:

$$\max_d d + \beta \left\{\max\,(1+r)(q-d); V^0(q-d, Z, \theta')\right\}$$

10. Repeat the last step for the other levels of productivity θ.

11. Solve for the optimal capital and compute the corresponding value function :

$$V^1(e,\theta,Z) = \max_k \mathbb{E}_\theta \left\{ \int_{\epsilon_{\min}}^{\epsilon_{\max}} \max_d \left[d + \beta \max \left\{ (1+r)(q-d), V^0(q-d,\theta',Z) \right\} \right] d\Phi(\epsilon) \right\}$$

12. Update the guess with this new found value function. Iterate step 9 to 11 until convergence.

Appendix C: Calibration of the transition matrix

The transition matrix Q is chosen to match the distribution of steady state productivity ξ^*_{FL} and a steady state exit rate of 5%. We also constrain $F(\theta'|\theta)$ to be decreasing in θ. Note that these conditions do not pin down a unique value for Q.

The transition matrix is obtained as follows:

1. Find the values that solve $\xi^*_{FL} = Q\xi^*_{FL} + \psi_{FL}e$ where e is a column vector with unit elements and $\psi_{FL} = 0.05/5$ consistent with the stationarity (exit rate=entry rate=0.05) and the uniform entry assumptions. The solution to this system is not unique, and the set of solutions can be characterized by adding the linear combination of the basis vector of the homogenous equation to a particular solution of the full equation.

2. The set of solution found in step 1 is then reduced to keep only the solutions for which:

- $F(\theta'|\theta) \geq 0$
- $\sum_{\theta'} F(\theta'|\theta) = 1$
- $F(\theta'|\theta)$ is decreasing in θ

Appendix D: Alternative calibration

In this appendix, we check the sensitivity of the results to an alternative assumption about the distribution of the potential entrants. In particular, we report the case in which potential entrants have a higher average net worth, and assume that they are uniformly distributed over $[\theta_1, \theta_5] \times [0, 10]$.

Table A.1: Net worth of entrants uniformly distributed over$[\theta_1, \theta_5] \times [0, 10]$

	Aggregate production	Direct effect	Intensive margin	Exit margin	Productivity increase
$\sigma = 0.05$	3.21%	0.70 %	1.65%	0.85%	0.44 %
$\sigma = 0.10$	3.25%	0.70 %	1.63%	0.91%	0.47 %
$\sigma = 0.15$	3.27%	0.70 %	1.58%	0.99%	0.66 %
$\sigma = 0.20$	3.38%	0.70 %	1.59%	1.09%	0.73 %
$\sigma = 0.25$	3.54%	0.70 %	1.59%	1.24%	0.83 %
$\sigma = 0.30$	4.38%	0.69%	1.45%	2.25%	1.53 %
$\sigma = 0.35$	4.50%	0.69%	1.56%	2.24%	1.52 %
$\sigma = 0.40$	4.53%	0.69%	1.60%	2.24%	1.52 %

Chapter 2

Aggregate productivity growth and the allocation of resources over the business cycle

2.1 Introduction

Recessions are often viewed as time where the economy is "cleansed": the least productive firms are forced to exit the market, allowing resources to be reallocated towards more productive uses. This Schumpeterian view of recessions, which has been emphasized by Caballero and Hammour (1994), suggests that the efficiency in the allocation of resources improves during recessions. While the theoretical literature has focused on the contribution of entry and exit, the dynamics of aggregate productivity could also be driven by changes in the efficiency of resource allocation across existing firms. Is allocative efficiency an important determinant of aggregate productivity changes over the business cycle? To answer this question, I propose a novel approach to separate out the variations in aggregate productivity which are due to within-firm productivity changes from those due to changes in the allocation of resources between incumbents, entering and exiting firms. To this end, I derive the link between micro and aggregate productivity dynamics in a framework where firms are heterogeneous and where market frictions distort the allocation of resources across firms. This approach extends the Solow (1957) growth accounting exercise to a framework with firm heterogeneity and allocative inefficiency.

The importance of resource reallocation for aggregate productivity growth has been documented in many empirical papers (Baily et al., 2001; Foster et al., 2001; Griliches and Regev, 1995; Bartelsman et al., 2009). However, all these papers focus on long run productivity changes and therefore provide little evidence on the contribution of resource reallocation at business cycle frequency. Furthermore, their decomposition relies on an aggregate productivity index computed as the weighted average of firm-level total factor productivity (TFP). The contribution of input reallocations is then measured by the correlation between changes in input shares and firm-level productivity.[1] In this paper, I argue that this correlation is not a well-defined measure of allocative efficiency. Once decreasing marginal productivity of inputs is accounted for, shifting resources towards high TFP firms may reduce allocative efficiency and lower aggregate production. In the general case, *aggregate* productivity is not equal to *average* productivity and must be derived from an aggregate production function. When derived from the aggregation of firm-level production functions, the contribution of input reallocation is exactly equal to the change in the efficiency of resource allocation and depends on the dispersion in the marginal products of capital and labor, not on the correlation between input shares and firm-level TFP. In fact, resources are efficiently allocated when the value of marginal productivities are equalized across firms. The level of allocative inefficiency in then measured with respect to this first-best benchmark as the dispersion in labor and capital marginal productivities. In this paper, I propose a measure of changes in allocative efficiency based on this criterion. Contrary to the existing decompositions, the reallocation component captures only the shifts in input shares that lead to a change in the level

[1] The decompositions found in this literature slightly differ from one another by the weights used (previous or/and current period, labor or market shares) and whether or not firm-level productivity is normalized (relative to the aggregate productivity index)

of allocative inefficiency.

The decomposition of aggregate productivity growth is derived in a model where frictions in output and input markets generate production inefficiencies. For a given level of aggregate inputs, a higher level of output could be reached if resources were more efficiently allocated. Following Restuccia and Rogerson (2008) and Chari et al. (2007), I do not specify the frictions that induce this resource misallocation. Rather, the frictions are modelled as wedges between the firms' marginal products. The model therefore encompasses various sources of distortions such as adjustment costs, search frictions, financial constraints or distortionary regulation. Within this framework, I show how to aggregate the heterogenous production functions into an aggregate production function. The aggregation of heterogeneous production functions is a classical problem in macroeconomics. It is well known that if one allows the individual firm inputs to take any values, the aggregation of firm-level technological constraints is impossible unless very restrictive conditions are imposed. The usual solution consists in defining the aggregate production function as the efficient frontier of the production possibilities set (e.g. Fisher (1969), Houthakker (1955)). As the focus is on misallocation and production inefficiencies, this is not the approach considered here. Following Malinvaud (1993), I define the aggregate production function as the relation between aggregate output and input for a given allocation of resources. Computed from this aggregate production function, aggregate productivity growth captures the variations in output that are due to changes in resource allocation, as well as those due to within-firm productivity change[2]. Then, using a method similar to the index number approach, I decompose aggregate productivity growth between productivity changes at the firm-level, changes in allocative efficiency, and changes in the pattern of entry and exit.

The decomposition of aggregate productivity growth is estimated from 1991 to 2006 on French

[2] I do not investigate the source of within-firm productivity change. Just as the aggregate productivity also depends on allocative efficiency, within-firm productivity could also depend on the allocation of resources within the firm. This dimension is left aside in the paper.

firm-level data from the manufacturing sector.[3] I use a dataset collected annually by the tax administration and combined with survey data in the INSEE unified system of business statistics (SUSE). The empirical analysis leads to the following findings: 1) entry and exit contribute very little to the dynamics of aggregate productivity growth. 2) movements in allocative efficiency are somewhat countercyclical, with a correlation to real value added growth of -0.25. 3) within-firm productivity changes are procyclical, with a correlation coefficient equal to 0.64. These results differ substantially from those obtained when implementing the decomposition proposed by Foster et al. (2001). Using their decomposition, the extensive margin accounts for a larger share of aggregate productivity growth and the reallocation component appears procyclical.

The finding that entry and exit flows have a negligible role for aggregate productivity growth contrasts sharply with the literature on the cleansing effect of recession (Caballero and Hammour, 1994; Barlevy, 2003; Ouyang, 2009). While the cleansing effect would imply a countercyclical extensive margin component, I find that, not only is the contribution of entry and exit small, it is also positively correlated to real value added growth. In fact, it is the reallocation of resources between incumbents, and not that between entering and exiting firms that tends to raise aggregate productivity during recessions. Despite the heterogeneity between sectors, the countercyclicality of allocative efficiency also holds for most sectors. This finding suggests new directions for future theoretical work as very little is known on the mechanisms behind the cyclical patterns of allocative efficiency.

This paper is not the first to advocate the use of a well-defined measure of allocative efficiency. Several recent papers (Petrin and Levinsohn, 2005; Basu et al., 2009; Petrin et al., 2011) have emphasized that the reallocation component should capture changes in the allocation of inputs between firms with different marginal values. None of these papers investigate the role of the extensive margin. More importantly, the methods used, as well as the results obtained

[3] Note that the link between micro and macro productivity is likely to be different if plant-level data were used instead. However, the availability of plan-level data is limited in France, and in particular do not allow to compute productivity.

are substantially different from these papers. Their decompositions are based on the Solow residual measure of productivity at the firm-level. They all build on Basu and Fernald (2002)'s insight that, under some conditions, the Solow residual approximates the welfare change of a representative consumer even when the allocation of resources is distorted by imperfect competition. Their measure of aggregate productivity growth therefore includes the effects of changes in aggregate input reflecting the fact that, under imperfect competition, welfare increases with aggregate input use. Moreover, their measure of allocative efficiency only captures the consequences of reallocation across firms with different markups and therefore does not account for all changes in the dispersion of marginal products. Basu et al. (2009) compute this decomposition for several European countries over the period 1995-2005 and find that allocative efficiency is not an important component of aggregate productivity. In particular, for France, they find that within-firm productivity changes explain all the changes in the Solow residual. This contrasts with my results in which allocative efficiency reduces by 51% the volatility of sectoral productivity growth. Contrary to Basu and Fernald (2002), I explicitly address the aggregation issue to investigate the link between the Solow residual and firm-level productivity change. By taking explicit account of firm heterogeneity and microeconomic frictions, I provide a complementary analysis to Hall (1991) who highlights the consequence of market imperfections for the measure of aggregate productivity growth in a representative firm framework. I show that the Solow residual gives a biased measure of productivity change when the heterogeneity in factor elasticities is accounted for, even when resources are efficiently allocated.

This paper is also closely related to Hsieh and Klenow (2009) who study the role of resource misallocation in explaining the TFP differential between China, India and the United States. As in their paper, resource misallocation is captured by the dispersion in the marginal products of capital and labor. However, both the objective and the decomposition differ from their paper. They analyze TFP differentials across countries, and quantify misallocation by measuring the distance between observed and first-best TFP levels. By contrast, I focus on

TFP variation across time, and propose a decomposition of observed TFP using an approach similar to the index number theory. Furthermore, contrary to Hsieh and Klenow (2009), I use a unified approach at both the sectoral and aggregate levels. This allows me to provide an estimation of allocative efficiency not only within sectors but also between sectors.

The paper is organized as follows. Section 2 lays out the framework and shows how to aggregate heterogeneous production units to derive the aggregate productivity index. Section 3 presents the decomposition of aggregate productivity both within and between sectors. Section 4 presents the estimation method and the results obtained on French firm-level data. Section 5 concludes.

2.2 Aggregation of heterogenous production units

Aggregate productivity is a concept which is intrinsically related to the production function. Aggregate productivity is defined as the change in real output not accounted for by the change in real input. To analyze changes in aggregate productivity, it is therefore necessary to derive the link between aggregate inputs and aggregate output. This section lays out the setup and shows how to derive the aggregate production function in a framework where producers are heterogeneous and face allocation frictions.

2.2.1 Framework

Consider an economy with S sectors. In each sector $s = 1...S$, there are N_s potential firms indexed by $i = 1, ..., N_s$. Firms produce a differentiated good Y_{it} using a Cobb-Douglas technology[4]:

$$Y_{it} = z_{it} K_{it}^{\alpha_s} L_{it}^{\beta_s},$$

[4]To simplify notations, sector subscripts s are omitted for the firms's output and inputs.

where K_{it} denotes capital, L_{it} labor and z_{it} the firm-level total factor productivity. Depending on the value of $\gamma_s \equiv \alpha_s + \beta_s$, firms face either decreasing or constant returns to scale $\gamma_s \leq 1$, for all $s = 1, ..., S$. The factor elasticities are assumed to be identical within sectors, but may vary from one sector to another. Firms face a downward sloping demand curve. For simplicity let us assume that the demand functions are iso-elastic. The price demand for good i reads:

$$p_{it} = b_{it} Y_{it}^{\theta_s - 1},$$

where $0 < \theta_s \leq 1$. The parameter θ_s governs the price elasticity of demand in sector s and b_{it} is a firm-specific demand shock. It is important to note that the size of the firm is indeterminate if firms use a constant returns to scale technology and behave competitively ($\gamma_s = 1$ and $\theta_s = 1$) In this case, the allocation rule is not unique and the aggregate production function is then not defined. I therefore restrict the analysis to the case where $\theta_s \leq 1$ and $\gamma_s \leq 1$, with at least one strict inequality.

In this economy, the allocation of input across firms is distorted by market frictions. Following Restuccia and Rogerson (2008) and Chari et al. (2007), let us remain agnostic about the nature of the frictions. Consider a generic model in which distortions are captured by the presence of wedges $\tau_{it} = (\tau_{it}^K, \tau_{it}^L)$ that disrupt the firms' input decisions with respect to the frictionless first-order conditions:

$$\alpha_s \theta_s \frac{p_{it} Y_{it}}{K_{it}} = r_t (1 + \tau_{it}^K) \quad (1)$$

$$\beta_s \theta_s \frac{p_{it} Y_{it}}{L_{it}} = w_t (1 + \tau_{it}^L), \quad (2)$$

The distortions τ_{it} reduce the efficiency of aggregate production as they prevent the value of marginal products to be equalized across firms. The cross-sectional inefficiency may differ across inputs: τ_{it}^K denotes distortions that affect specifically the marginal product of capital, and τ_{it}^{L}, the marginal product of labor. As already mentioned, this model encompasses different types of frictions. Adjustment costs, search frictions, financial constraints and distortionary regulation all generate gaps between the firms' marginal products. Note that frictions that generate a wedge between the marginal product and the marginal cost of an input are not a source of allocative inefficiency if they affect all the firms in the same way. Hence, in this framework, imperfect competition does not distort the allocation of resources across firms within sectors as the markup $(1/\theta_s)$ is assumed to be identical within sectors. However, imperfect competition may reduce the efficiency of resource allocation across sectors. The allocation of resources across firms from different sectors is distorted both by firm-specific distortions and by sector-specific markups.

The firm does not always find it profitable to produce; depending on its level of productivity and distortions, the firm could decide to exit. The firm exits if $I_{it}(z_{it}, b_{it}, \tau_{it}) < 0$ and produces if $I_{it}(z_{it}, b_{it}, \tau_{it}) \geq 0$. The function I_{it} characterizes the participation decision of the firm and may vary across firms and across time as the decision to produce depends on factor prices, firm-specific fixed costs, and whether the firm consider to enter or exit. Combining the participation decision with equations (1) and (2), and with the demand curve, we can derive the input levels as a function of the firm's distortions, productivity and demand shock:

$$K_{it} = \begin{cases} \theta_s z_{it}^{\frac{\theta_s}{1-\gamma\theta_s}} b_{it}^{\frac{1}{1-\gamma\theta_s}} \left(\frac{r}{\alpha_s}(1+\tau_{it}^{K})\right)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}} \left(\frac{w}{\beta_s}(1+\tau_{it}^{L})\right)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}} & \text{if } I_{it}(z_{it}, b_{it}, \tau_{it}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

$$L_{it} = \begin{cases} \theta_s z_{it}^{\frac{\theta_s}{1-\gamma\theta_s}} b_{it}^{\frac{1}{1-\gamma\theta_s}} \left(\frac{r}{\alpha_s}(1+\tau_{it}^{K})\right)^{\frac{-\alpha_s\theta_s}{1-\gamma\theta_s}} \left(\frac{w}{\beta_s}(1+\tau_{it}^{L})\right)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}} & \text{if } I_{it}(z_{it}, b_{it}, \tau_{it}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

2.2.2 Aggregation

The aggregation of heterogeneous firm-level production functions is a classical problem in macroeconomics. The link between aggregate input and output is not straightforward to derive. In fact, changes in aggregate input may affect aggregate output differently, depending on the way the additional input is allocated across firms. It is well known that, if one allows the firms' inputs to take any values, the aggregation of firm-level technological constraints is possible only under very restrictive conditions.[5] In this paper, as the objective is to disentangle the effects of aggregate input from that of allocative efficiency and within-firm productivity changes, the aggregate production function need not be a pure technical relationship : we are looking for the relationship between aggregate output and input for a given allocation of resources. This is Malinvaud (1981)'s approach to the aggregate production function. With this definition, the aggregate production function becomes specific to the economic environment and the conditions for aggregation become less restrictive.

The aggregate production function is derived in three steps. First, we must take into account the heterogeneity in the goods produced. In particular, the impact of a reallocation of resources between two different goods on aggregate productivity should depend on the relative value of the goods produced. Then, I show how to derive the aggregate production function at the sectoral level, where factor and demand elasticities are homogeneous. Finally, I take into account the heterogeneity across sectors and aggregate the sectoral production functions. Deriving first the sectoral production functions allows us to disentangle the changes in the allocative efficiency within and between sectors.

2.2.2.1 Accounting for goods heterogeneity

Since goods are heterogeneous, the relative value of the goods should be accounted for and the individual production functions described in section 2.1 should therefore not be aggregated directly. This is all the more important when studying the impact of reallocations on aggregate productivity. Suppose that all the firms use a constant returns to scale technology. If the impact of reallocation on the relative value of the goods is not accounted for, aggregate production is maximized when resources are reallocated towards the firm with the highest TFP level, leading the economy to produce only one type of good. However, this does not

[5] The individual production functions must be linear and have the same slopes (Nataf, 1948)

correspond to the optimal allocation of resources when goods are heterogeneous. The optimal allocation is the one that equalizes the *value* of marginal productivity across firms. Reallocating all the resources toward the production of one good would decrease the value of this good, and raise that of non-produced goods. To obtain a measure of aggregate productivity consistent with this optimal allocation rule, we must take into account the impact of input changes not only on production but also on relative prices. This leads to a modified production function, where the output of the firm depends on the value of the goods produced. The output of the firm is then $\frac{p_{it}}{P_{st}}Y_{it} = f_i(K_{it}, L_{it})$, where the production function is given by:

$$f_i(K_{it}, L_{it}) = A_{it}K_{it}^{\alpha_s\theta_s}L_{it}^{\beta_s\theta_s}, \tag{5}$$

with $A_{it} \equiv z_{it}^{\theta_s}b_{it}^{1-\theta_s}/P_{st}$ and P_{st} is the sectoral price index. When accounting for changes in the relative value of goods, the actual productivity of the firm combines both technical efficiency and demand factors. Note that A_{it} is also the usual measure of productivity at the firm level. Because firm-level prices are rarely available, the firm's output is usually measured by its nominal output divided by a sectoral deflator. Though it hinders the measure of technical efficiency at the firm-level[6], the absence of firm-level prices is not problematic here. What matters for aggregate productivity is the measured firm-level productivity A_{it}. In fact, the whole model can be rewritten as a function of measurable variables only. In particular, the input decision described in equations (3) and (4) can be rewritten as $K_i(P_{st}A_{it}, \tau_{it}, r_t, w_t)$ and $L_i(P_{st}A_{it}, \tau_{it}, r_t, w_t)$:

$$K_{it} = \begin{cases} \theta_s(P_{st}A_{it})^{\frac{1}{1-\gamma\theta_s}}\left(\frac{r}{\alpha_s}(1+\tau_{it}^K)\right)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}}\left(\frac{w}{\beta_s}(1+\tau_{it}^L)\right)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}} & \text{if } I_{it}(z_{it}, b_{it}, \tau_{it}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$L_{it} = \begin{cases} \theta_s(P_{st}A_{it})^{\frac{\theta_s}{1-\gamma\theta_s}}\left(\frac{r}{\alpha_s}(1+\tau_{it}^K)\right)^{\frac{-\alpha_s\theta_s}{1-\gamma\theta_s}}\left(\frac{w}{\beta_s}(1+\tau_{it}^L)\right)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}} & \text{if } I_{it}(z_{it}, b_{it}, \tau_{it}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

[6]The consequences of the absence of firm-level prices for the measure of firm-level TFP have been recently emphasized by Foster et al. (2008).

2.2.2.2 The sectoral production functions

Let us now aggregate the modified production functions described above. To find the relation between aggregate output ($Y = \sum_i \frac{p_i}{P} Y_i$) and aggregate inputs ($K = \sum_i K_i$ and $L = \sum_i L_i$), I use Malinvaud (1993)'s approach. The idea is to find the allocation rules, i.e how aggregate inputs are allocated between firms, and then aggregate output over firms using the allocation rules. More formally, once the allocation rules $K_i = k_i(K, L)$ and $L_i = l_i(K, L)$ are known, the aggregate production function is simply: $Y = \sum_i f_i(k_i(K, L), l_i(K, L))$.

The allocation rules can be derived from the equilibrium condition on the inputs markets. Specifically, they are obtained after inverting the aggregate input equations. At the sectoral level, total inputs are given by:

$$\begin{aligned} K_{st} &= \sum_{i=1}^{N_s} K_i(P_{st} A_{it}, \tau_{it}, r_t, w_t) \\ L_{st} &= \sum_{i=1}^{N_s} L_i(P_{st} A_{it}, \tau_{it}, r_t, w_t). \end{aligned}$$

Inverting this system of equations allows us to write the factor prices as a function of aggregate capital, labor input and the vector of firm-level productivity $\widetilde{A}_{st} = \{A_{1t}, ..., A_{N_s t}\}$ and distortions $\widetilde{\tau}_t = \{\tau_{1t}, ..., \tau_{N_s t}\}$.[7] The input levels can then be written as a function of aggregate inputs, firm-level productivities and distortions. The sectoral production function follows:

$$F_s(K_{st}, L_{st}, \widetilde{A}_{st}, \widetilde{\tau}_{st}) = \sum_{i=1}^{N_s} f_i(k_i(P_{st}\widetilde{A}_{st}, \widetilde{\tau}_{st}, K_{st}, L_{st}), l_i(P_{st}\widetilde{A}_{st}, \widetilde{\tau}_{st}, K_{st}, L_{st})).$$

In general, the aggregate production function does not share the same functional form as the individual production functions. However, when the factor and demand elasticities are identical across firms (as assumed within sectors), I show in Appendix A that the aggregate production function is also Cobb-Douglas. The sectoral production function is:

$$F_s(K_{st}, L_{st}, \widetilde{A}_{st}, \widetilde{\tau}_{st}) = TFP_{st} K_{st}^{\alpha_s \theta_s} L_{st}^{\beta_s \theta_s},$$

[7] This system is loacally invertible if the Jacobian of the application $\{K_s(r, w), L_s(r, w)\}$ is non-zero.

where TFP_{st} is a function of both firm-level productivity $\widetilde{A}_{st}$ and distortions $\widetilde{\tau}_{st}$.

And the change in sectoral productivity can be computed as:

$$\frac{dTFP_{st}}{TFP_{st}} = \frac{dY_{st}}{Y_{st}} - \alpha_s\theta_s\frac{dK_{st}}{K_{st}} - \beta_s\theta_s\frac{dL_{st}}{L_{st}} \tag{8}$$

2.2.2.3 From the sectoral to the aggregate production function

In order to aggregate the sectoral production functions, let us characterize each sector's representative firm.The sectoral input demand function are identical to the firm-level functions (equations (6) and (7) with the participation equation always satisfied). The sectoral input levels $K_s(P_{st}TFP_{st}, \omega_{st}, r_t, w_t)$ and $L_s(P_{st}TFP_{st}, \omega_{st}, r_t, w_t)$ are functions of sectoral level of distortions $\omega_{st} = (\omega_{st}^K, \omega_{st}^L)$ which are given by:

$$1 + \omega_{st}^K = \sum_i \frac{K_{it}}{K_{st}}(1 + \tau_{it}^K) \tag{9}$$

$$1 + \omega_{st}^L = \sum_i \frac{L_{it}}{L_{st}}(1 + \tau_{it}^L). \tag{10}$$

Like at the sectoral level, the aggregate production function is obtained after inverting the aggregate input equations.[8] Since the factor and demand elasticities are allowed to be heterogenous across sectors, the aggregate production function is not Cobb-Douglas. Real aggregate output Y is given by $\sum_s Y_s = F(K_t, L_t, \widetilde{TFP}_t, \widetilde{\omega}_t, P_{st})$, where:

$$F(K_t, L_t, \widetilde{TFP}_t, \widetilde{\omega}_t, \widetilde{P}_t) = \sum_{s=1}^{S} TFP_{st}K_{st}^{\alpha_s\theta_s}L_{st}^{\beta_s\theta_s}$$

[8] The aggregate input equations are:
$K = \sum_{s=1}^{S} K_s(P_{st}TFP_{st}, \omega_{st}, r_t, w_t)$
$L = \sum_{s=1}^{S} L_s(P_{st}TFP_{st}, \omega_{st}, r_t, w_t)$

$$\text{with } K_{st} = k_s(\widetilde{TFP_t}, \widetilde{\omega}_t, \widetilde{P}_t, K_t, L_t)$$
$$L_{st} = l_s(\widetilde{TFP_t}, \widetilde{\omega}_t, \widetilde{P}_t, K_t, L_t).$$

$\widetilde{TFP}_t = \{TFP_{1t}, ..., TFP_{St}\}$ is the vector of sectoral level productivity, $\widetilde{\omega}_t = \{\omega_{1t}, ..., \omega_{St}\}$ the vector of sectoral level distortions and $\widetilde{P}_t = \{P_{1t}, ..., P_{St}\}$ is the vector of sectoral price indexes.

Aggregate TFP growth is then:

$$\frac{dTFP_t}{TFP_t} = \frac{dY}{Y} - \varepsilon_K(\widetilde{\alpha_s}, \widetilde{\beta})\frac{dK}{K} - \varepsilon_L(\widetilde{\alpha_s}, \widetilde{\beta})\frac{dL}{L}, \tag{11}$$

with $\widetilde{\alpha} = \{\alpha_1\theta_1, ..., \alpha_S\theta_S\}$, $\widetilde{\beta} = \{\beta_1\theta_1, ..., \beta_S\theta_S\}$.

Note that the aggregate elasticity to capital ε_K is not a weighted average of sector-level elasticities to capital, and similarly, the aggregate elasticity to labor ε_L is not a weighted average of sector-level elasticities to labor. The aggregate elasticity of capital and labor both depend on the sector-level elasticities of labor and capital and on sectoral price elasticities. To see more explicitly this point, let us assume that the decreasing returns to scale and the elasticity of demand parameters are identical across sectors $\gamma_s = \gamma$ and $\theta_s = \theta$, for all s. The change in aggregate TFP is then:

$$\frac{dTFP_t}{TFP_t} \equiv \frac{dY}{Y} - \left(\frac{\alpha^Y(1-\alpha^L) - \beta^Y\alpha^L}{1-\alpha^L-\beta^K}\right)\frac{dK}{K} - \left(\frac{\beta^Y(1-\beta^K) - \alpha^Y\beta^K}{1-\alpha^L-\beta^K}\right)\frac{dL}{L},$$

with $\alpha^X = \sum_s \frac{X_s}{X}\alpha_s\theta$ and $\beta^X = \sum_s \frac{X_s}{X}\beta_s\theta$ for $X = Y, K, L$.

As will be shown in section 4, the heterogeneity in factor elasticities may bias the estimate of aggregate factor elasticities and of aggregate productivity growth.

2.3 Accounting for changes in aggregate productivity

This section presents the decomposition of aggregate productivity growth into changes in firm-level productivity, changes in the efficiency of resource allocation within and between sectors, and changes in entry and exit patterns. I first decompose sectoral productivity growth. Then, I use the aggregate production function to aggregate the within-sector components and to derive the contribution of allocative efficiency between sectors.

2.3.1 Decomposition of sectoral productivity

As shown in section 2.2.2 and Appendix A, the sectoral production function is given by: $Y_{st} = TFP_{st} K_{st}^{\alpha_s \theta_s} L_t^{\beta_s \theta_s}$, where sectoral productivity is a function of firm-level productivity A_{it} and distortions $\tau_{it} = (\tau_{it}^K, \tau_{it}^L)$:[9]

$$TFP_{st} = \frac{\sum_{i=1}^{N_s} g_s^Y(A_{it}, \tau_{it})}{\left(\sum_{i=1}^{N_s} g_s^K(A_{it}, \tau_{it})\right)^{\alpha_s \theta_s} \left(\sum_{i=1}^{N_s} g_s^L(A_{it}, \tau_{it})\right)^{\beta_s \theta_s}}. \tag{1}$$

To show more specifically how aggregate productivity depends on the cross-sectional distribution of firm-level productivity and distortions, let us assume that A_i, $(1+\tau_i^K)$ and $(1+\tau_i^L)$ are jointly lognormally distributed. Sectoral TFP can then be written as a function of the first two moments of the joint distribution:

$$\begin{aligned} \ln TFP_{st} &= (1-\gamma_s\theta_s)\ln N_{st} + E_t \ln A_i + \frac{1}{2}\frac{1}{1-\gamma_s\theta_s} V_t \ln A_i \\ &\quad -\frac{1}{2}\frac{\alpha_s\beta_s\theta_s^2}{1-\gamma_s\theta_s}\left(\frac{1-\beta_s\theta_s}{\beta_s\theta_s} V_t \ln(1+\tau_i^K) + \frac{1-\alpha_s\theta_s}{\alpha_s\theta_s} V_t \ln(1+\tau_i^L) + Cov_t(\ln(1+\tau_i^K), \ln(1+\tau_i^L))\right), \end{aligned}$$

where E_t, V_t and Cov_t denote the period t cross-sectional expectation, variance and covariance. Sectoral TFP depends on the average and the dispersion of firm-level productivity, but also on

[9] The g functions are given by:
$g_s^Y(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}$
$g_s^K(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}$
$g_s^L(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}}$

the dispersion in the distortions, on the covariance between labor and capital distortions, and on the number of firms in the sector. Idiosyncratic (uncorrelated) shocks to either productivity or distortions have an impact on aggregate productivity only to the extent that these shocks also modify the cross-sectional variance. This expression also shows that average distortions, as well as the covariance between firm level productivity and distortion plays no role for aggregate productivity growth. Though the result on the covariance will probably not hold in a more general framework, the insensitivity of aggregate productivity to a common distortion factor holds beyond the case of the log-normal distribution.

Let us now decompose sectoral TFP in the general case, without specifying the joint distribution of A_i, $(1+\tau_i^K)$ and $(1+\tau_i^L)$. In appendix B, I show that aggregate productivity growth (ΔTFP_{st}) can be decomposed into changes in within-firm productivity $(\Delta \mathrm{TE}_s)$, changes in allocative efficiency $(\Delta \mathrm{AE}_s)$ and changes at the extensive margin $(\Delta \mathrm{EX}_s)$[10]:

$$\Delta TFP_{st} \approx \Delta \mathrm{TE}_s + \Delta \mathrm{AE}_s + \Delta \mathrm{EX}_s,$$

An approximation for each component is given below. The decomposition between changes in within-firm productivity and changes in allocative efficiency is similar to the decompositions we can find in the index number literature. To avoid any asymmetry induced by the functional form of the indexes, I measure in the data each component with a Fisher-like index, i.e. the geometric mean of the Laspeyres-like and Paasche-like indexes. The exact decomposition with the Fisher-like index is described in Appendix B. For simplicity, I present here an approximation in which the effects of within-firm productivity changes are measured with a Laspeyres-like index, while those of allocative efficiency are measured with a Paasche-like index.

The change in firm-level productivity can be approximated as a combination of weighted averages of the firm level productivity changes:[11]

[10] For convenience, we denote the aggregate productivity growth by $\Delta TFP_{st} \equiv TFP_{st}/TFP_{st-1} - 1$

[11] Note that the case of perfect competition and constant returns to scale is not determinate in this framework

$$\Delta TE_s \simeq \frac{1}{1-\gamma_s\theta_s}\sum_{i\in C_s}\frac{\Delta A_{it}}{A_{it-1}}\left[\frac{p_{it-1}Y_{it-1}}{\sum\limits_{i\in C_s}p_{it-1}Y_{it-1}} - \alpha_s\theta_s\frac{K_{it-1}}{\sum\limits_{i\in C_s}K_{it-1}} - \beta_s\theta_s\frac{L_{it-1}}{\sum\limits_{i\in C_s}L_{it-1}}\right], \qquad (2)$$

with C_s is the set of continuing firms ($I_{it-1} \geq 0$ and $I_{it} \geq 0$). Note that the impact of within-firm productivity change is measured for a given level of allocative efficiency. Therefore, this measure includes the effect of within-firm productivity changes for a given distribution of inputs, as well as the consequences of within-firm productivity changes on input shares for a given level of distortions. The level of distortions is measured here by the relative marginal productivities of input across firms.[12] Therefore, the firm-level efficiency component captures the impact of within-firm productivity changes, had relative marginal productivities remained constant.

The change in allocative efficiency is a combination of weighted averages of firm-level changes in distortion:

$$\begin{aligned}\Delta AE_s \simeq\ & \frac{\alpha_s\theta_s}{1-\gamma_s\theta_s}\sum_{i\in C_s}\frac{\Delta(1+\tau^K_{it})}{1+\tau^K_{it-1}}\left[(1-\beta_s\theta_s)\frac{K_{it}}{\sum\limits_{i\in C_s}K_{it}} + \beta_s\theta_s\frac{L_{it}}{\sum\limits_{i\in C_s}L_{it}} - \frac{p_{it}Y_{it}}{\sum\limits_{i\in C_s}p_{it}Y_{it}}\right] \\ & + \frac{\beta_s\theta_s}{1-\gamma_s\theta_s}\sum_{i\in C_s}\frac{\Delta(1+\tau^L_{it})}{1+\tau^L_{it-1}}\left[\alpha_s\theta_s\frac{K_{it}}{\sum\limits_{i\in C_s}K_{it}} + (1-\alpha_s\theta_s)\frac{L_{it}}{\sum\limits_{i\in C_s}L_{it}} - \frac{p_{it}Y_{it}}{\sum\limits_{i\in C_s}p_{it}Y_{it}}\right] \qquad (3)\end{aligned}$$

The contribution of the extensive margin to aggregate productivity growth depends of the relative weights of entering and exiting firms:

[12] An alternative measure of distortions is the *difference* between marginal productivities (instead of the *ratio*). Note that in this case, the decomposition is identical to the case where elasticities are heterogeneous (equations (7) to (10)) but with a different values for aggregate elasticities ε_K and ε_L. In fact, the way distortions are measured affects not only the measure of changes in allocative efficiency, but also the impact of changes in aggregate inputs and within-firm productivity on aggregate output as these are computed holding fixed the level of distortions.

$$\Delta EX_s \simeq \frac{\sum\limits_{i\in E_s}\frac{p_{it}Y_{it}}{P_{st}Y_t}-\sum\limits_{i\in X_s}\frac{P_{it-1}Y_{it-1}}{P_{st-1}Y_{t-1}}}{1-\sum\limits_{i\in E_s}\frac{p_{it}Y_{it}}{P_{st}Y_t}} - \alpha_s\theta_s\frac{\sum\limits_{i\in E_s}\frac{K_{it}}{K_t}-\sum\limits_{i\in X_s}\frac{K_{it-1}}{K_{t-1}}}{1-\sum\limits_{i\in E_s}\frac{K_{it}}{K_t}} - \beta_s\theta_s\frac{\sum\limits_{i\in E_s}\frac{L_{it}}{L_t}-\sum\limits_{i\in X_s}\frac{L_{it-1}}{L_{t-1}}}{1-\sum\limits_{i\in E_s}\frac{L_{it}}{L_t}} \tag{4}$$

with E_s the set of new entrants ($I_{it-1} < 0$ and $I_{it} \geq 0$) and X_s the set of exiting firms ($I_{it-1} \geq 0$ and $I_{it} < 0$) in sector s. The extensive margin contributes positively to aggregate productivity growth if entering firms are more productive and face less distortions than exiting firms. As long as there are decreasing returns to scale $\gamma_s\theta_s<1$ (in production or in demand), the positive effect of higher output shares is not offset by larger input shares. All else equal, a higher returns to scale parameter reduces the role of entry and exit for aggregate productivity growth. Note that with decreasing returns to scale, the efficiency of production increases with the number of firms.

When resources are perfectly allocated across firms ($1+\tau_i = 1+\tau$, for all $i = 1, ..., N_s$), the allocation of labor, capital and output depends only on the firms' relative productivity. Moreover, in this case, input and output shares are equal ($\frac{K_i}{K} = \frac{L_i}{L} = \frac{p_{it}Y_{it}}{P_{st}Y_t}$), and the components of aggregate productivity simplify to:

$$\Delta TE_s^{FL} \simeq \sum_{i\in C_s}\frac{\Delta A_{it}}{A_{it-1}}\frac{p_{it-1}Y_{it-1}}{\sum\limits_{i\in C_s}p_{it-1}Y_{it-1}}$$

$$\Delta AE_s^{FL} = 0$$

$$\Delta EX_s^{FL} \simeq (1-\gamma_s\theta_s)\frac{\sum\limits_{i\in E_s}\frac{p_{it}Y_{it}}{P_{st}Y_t}-\sum\limits_{i\in X_s}\frac{p_{it-1}Y_{it-1}}{P_{st-1}Y_{t-1}}}{1-\sum\limits_{i\in E_s}\frac{p_{it}Y_{it}}{P_{st}Y_t}}$$

In the absence of allocation distortions, the impact of within-firm productivity changes can be measured by a simple weighted average. The allocative efficiency component is null as resources are always efficiently allocated, and the extensive margin component simplifies to the difference between the output shares of entering and exiting firms.

2.3.2 Decomposition across sectors

The aggregate production function, derived in section 2.3.3, is a function of the vector of sectoral productivity $\widetilde{TFP}_t = \{TFP_{1t}, ..., TFP_{St}\}$, and sectoral distortions $\widetilde{\omega}_t = \{\omega_{1t}, ..., \omega_{St}\}$:

$$\sum_s Y_s = F(K_t, L_t, \widetilde{TFP}_t, \widetilde{\omega}_t, P_{st}), \tag{5}$$

As defined in equation (11), aggregate productivity growth is the change in output that is not due to the change in aggregate inputs. Combining this equation with the total differential of equation (5), aggregate productivity can also be written as the change in aggregate output explained by changes in sectoral productivity and distortions. This leads to the following decomposition:[13]

$$\frac{dTFP_t}{TFP_t} = \underbrace{\sum_{s=1}^{S} \frac{\partial F}{\partial TFP_s} \frac{TFP_s}{Y} \frac{dTFP_s}{TFP_s}}_{\text{Within sectors}} + \underbrace{\sum_{s=1}^{S} \frac{\partial F}{\partial \omega_s^K} \frac{1+\omega_s^K}{Y} \frac{d\omega_s^K}{1+\omega_s^K} + \frac{\partial F}{\partial \omega_s^L} \frac{1+\omega_s^L}{Y} \frac{d\omega_s^L}{1+\omega_s^L}}_{\text{Between sectors}}.$$

The first term measures changes in within-sector productivity and the second term measures changes in the efficiency of resource allocation across sectors. Using the implicit function theorem to compute the derivative of F with respect to TFP_s and replacing the within-sector productivity change by the decomposition derived at the sectoral level yields the decomposition of aggregate productivity growth (ΔTFP_t) into changes in firm-level productivity $\Delta\overline{\text{TE}}$, changes in the allocation of resources between $\Delta\overline{\text{AE}}_{\text{between}}$ and within sectors $\Delta\overline{\text{AE}}_{\text{within}}$, as well as changes in entry and exit patterns $\Delta\overline{\text{EX}}$:

$$\Delta TFP_t = \Delta\overline{\text{TE}} + \Delta\overline{\text{AE}}_{\text{within}} + \Delta\overline{\text{EX}} + \Delta\overline{\text{AE}}_{\text{between}}. \tag{6}$$

[13] Note that the impact of changes in sectoral prices are neglected. In the empirical results we verify that the last term of the decomposition $\sum_{s=1}^{S} \frac{\partial F}{\partial P_s} \frac{P_s}{Y} \frac{dP_s}{P_s}$ is indeed negligible.

The details of the derivation are given in Appendix B.[14] The within-sector productivity decomposition is aggregated at the macro level using input and output sectoral shares and the elasticities of the aggregate production function with respect to aggregate inputs, ε_K and ε_L:

$$\Delta\overline{\text{TE}} = \sum_{s=1}^{S} \frac{1}{1-\gamma_s\theta_s}\left(\frac{Y_s}{Y} - \varepsilon_K\frac{K_s}{K} - \varepsilon_L\frac{L_s}{L}\right)\Delta\text{TE}_s \tag{7}$$

Similarly, the change in within-sector allocative efficiency and the aggregate impact of the extensive margin can be computed as:

$$\Delta\overline{\text{AE}}_{\text{within}} = \sum_{s=1}^{S} \frac{1}{1-\gamma_s\theta_s}\left(\frac{Y_s}{Y} - \varepsilon_K\frac{K_s}{K} - \varepsilon_L\frac{L_s}{L}\right)\Delta\text{AE}_s \tag{8}$$

$$\Delta\overline{\text{EX}} = \sum_{s=1}^{S} \frac{1}{1-\gamma_s\theta_s}\left(\frac{Y_s}{Y} - \varepsilon_K\frac{K_s}{K} - \varepsilon_L\frac{L_s}{L}\right)\Delta\text{EX}_s \tag{9}$$

Finally, the change in between-sector allocative efficiency is computed as follows:

$$\begin{aligned}\Delta\overline{\text{AE}}_{\text{between}} \quad = \quad & \sum_{s=1}^{S} \frac{1}{1-\gamma_s\theta_s}\left(\varepsilon_K(1-\beta_s\theta_s)\frac{K_s}{K} + \varepsilon_L\alpha_s\theta_s\frac{L_s}{L} - \alpha_s\theta_s\frac{Y_s}{Y}\right)\frac{d\omega_s^K}{1+\omega_s^K} \\ & + \sum_{s=1}^{S} \frac{1}{1-\gamma_s\theta_s}\left(\varepsilon_K\beta_s\theta_s\frac{K_s}{K} + \varepsilon_L(1-\alpha_s\theta_s)\frac{L_s}{L} - \beta_s\theta_s\frac{Y_s}{Y}\right)\frac{d\omega_s^L}{1+\omega_s^L} \quad (10)\end{aligned}$$

If factor elasticities were equal across sectors, all these expressions would be exactly equal to the decomposition derived at the sectoral level.[15]

2.4 Estimation

Aggregate and sectoral productivity growth can be computed directly from aggregate data. However, estimating the decomposition given by equation (6) requires the use of firm-level

[14]Note that, contrary to the within-sector decomposition, the decomposition across sectors requires no specific assumptions about the individual production functions ; it requires though the existence of a unique allocation rule, necessary for the existence of the aggregate production function.

[15]We cannot, however, use the same method at the sectoral level. Because of the discontinuity introduced by the entry and exit decisions, we cannot take the total derivative of the sectoral production function.

data. First, the firm-level efficiency, allocative efficiency and extensive margin components are computed at the sectoral level using equations (2) to (4). Then the within-sector components can be aggregated at the macro level as described by equation (7) to (9). The change in between-sector allocative efficiency is computed with equation (10). In this section, I present the data used, the estimation strategy and the results of this decomposition estimated on French firm-level data.

2.4.1 Data description

Following the bulk of the literature, I investigate the dynamics of aggregate productivity in the manufacturing industry. The data used in this analysis are collected every year by the French tax administration, and combined with survey data in the INSEE unified system for business statistics (SUSE). I use the database of private businesses that declare their profits under the "normal" regime (*Bénéfice réel normal*) from 1989 to 2007. In 2003, the "normal" regime accounted for 24.4 % of firms and 94.3% of total sales. Firms are required to provide balance sheet data, which includes measures of the firms' value added, expenditures on capital and number of employees. I use sectoral value added price indexes to deflate firm-level value added and reconstruct the real capital stock using the perpetual inventory method.

Potential entries and exits are detected by following firms with at least one employee that appear in and disappear from the database.[16] Since firms are followed through their identification number (SIREN), they appear and disappear from the database not only when they actually open or shut down their businesses, but also in case of restructuring or takeover as this induces a change in their identification number. Entry and exit rates are therefore likely to be overestimated. I try to limit this bias by excluding firms that disappear or appear with a number of employees higher than the 99.8 percentile among exiting and entering

[16]Note that entering flows are likely to be overestimated as firms appear in the database when they cross the sales threshold of 763 000 euros above which the *Bénéfice réel normal* regime is mandatory. However, these spurious entry flows are likely to be limited as the "normal" tax regime is widely chosen by firms which are below the threshold: in 2003, 46% of the sample was below the sales threshold. Furthermore the dataset covers a lot of small firms, in 2003 71% of the sample (restricted to firms with at least one employee) had less than 20 employees

firms[17]. Furthermore, to avoid spurious entry or exit flows induced by missing values, firms that temporarily move below the threshold of one employee, or temporarily disappear from the database are excluded. The procedure used for excluding temporary exits is likely to overestimate the entry rate in the first years of the sample, and overestimate the exit rate in the last years of the sample. Since about 75% of the firms that temporarily disappear from the database are absent only one year, dropping the first and the last year of the sample considerably reduces the amount of spurious entry and exit flows. I therefore remove the first two years (entry cannot be identified in the first year) and the last year of the sample, and undertake the analysis over the period 1991-2006. I also remove outliers as the decomposition is quite sensitive to extreme values[18]. After this trimming procedure, the dataset is constituted by about 126 000 observations each year. Figure A.1 in appendix C, shows that the aggregate value added is not affected much by this trimming procedure and that it compares quite well with the national accounts data. Appendix C also provides a more detailed description of the data and the sources.

2.4.2 Estimation method

To implement the decomposition, I need estimates of firm-level and aggregate factor elasticities, firm-level and sectoral distortions, and firm-level productivity. In the following, I present the issues that arise and the assumptions required to estimate these parameters.

[17] This correspond to firms appearing in or disappearing from the database with about 1000 employees. Despite this correction, the entry and exit rates are likely to be overestimated since the threshold is quite high in light of the median size of target firms. Using French data over the period 2000-2004, Bunel et al. (2009) find a median size between 30 and 87 employees, depending on the exact nature of the restructuring operation. Note however that the average entry and exit rates displayed in Table A.1 are in line with what has been found for the US manufacturing sector (Dunne et al., 1989).

[18] Firms whose TFP is multiplied or divided by 2.5 are excluded from the analysis.

2.4.2.1 Estimation of production functions

The firm-level productivity estimates are likely to suffer from the traditional bias linked to unobserved factor utilization[19]. I abstract from this well-known measurement issue and focus here on the specific issues that arise in the presence of heterogeneity in factor elasticities and unobserved allocation frictions. First, I show that the standard growth accounting approach gives a biased measure of aggregate factor elasticities when the underlying factor elasticities are heterogeneous. Second, I investigate the biases induced by unobserved micro frictions.

Heterogeneity bias

To isolate the effects of the heterogeneity bias from the biases generated by market imperfections, let us assume that resources are perfectly allocated across firms ($\tau_{it}^K = \tau_{it}^L = 0$ for all i), and that firms behave competitively ($\theta_s = 1$, for all $s = 1, ...S$). In that case, as shown in section 2.2.3, the aggregate capital and labor elasticity are both a combination of capital and labor elasticities. In particular, when the decreasing to scale and the demand parameters are identical across sectors ($\gamma_s = \gamma$ for all $s = 1, ...S$), the aggregate elasticities are given by:

$$\varepsilon_K(\widetilde{\alpha}, \widetilde{\beta}) = \frac{\alpha^Y(1-\alpha^L) - \beta^Y \alpha^L}{1-\alpha^L-\beta^K} \text{ and } \varepsilon_L(\widetilde{\alpha}, \widetilde{\beta}) = \frac{\beta^Y(1-\beta^K) - \alpha^Y \beta^K}{1-\alpha^L-\beta^K}, \tag{1}$$

with $\alpha^X = \sum_s \frac{X_s}{X}\alpha_s$ and $\beta^X = \sum_s \frac{X_s}{X}\beta_s$ for $X = PY, K, L$.

Standard growth accounting method yield $\widehat{\varepsilon_K} = rK/PY$ and $\widehat{\varepsilon_L} = wL/PY$. Using equation (1) and (2), we find:

$$\widehat{\varepsilon_K} = \alpha^Y \text{ and } \widehat{\varepsilon_L} = \beta^Y,$$

which is different from the actual elasticity. Though there are no market imperfections, the Solow residual does not properly measure aggregate productivity changes when factor

[19] As emphasized in section 2.2.1, another bias comes from the absence of firm-level data which leads firm-level productivity estimates to combine both demand and technical efficiency components.

elasticities are heterogeneous. Aggregate elasticities are therefore computed using equations (1).

Unobserved micro distortions

There is no heterogeneity bias when using the Solow residual at the firm-level. However, because of the presence of allocation distortions, the growth accounting approach cannot be used at the firm-level for it is not possible to identify separately the factor elasticities from the firm-level distortions (equations 1 and 2).[20] I use the assumption that factor elasticities are identical within sectors, and estimate the firms' production functions using growth accounting at the sectoral level. To derive an estimate for $\beta_s\theta_s$, I use equation (2) aggregated over firms and over time :

$$\beta_s\theta_s\frac{1}{T}\sum_t\frac{1}{(1+\omega^L_{st})}=\frac{1}{T}\sum_t\frac{w_tL_{st}}{P_{st}Y_{st}}$$

In order to identify the factor elasticity, I assume that the average labor sectoral distortions are null over the period considered.[21] This assumption may bias the estimates of the factor elasticities but has no impact on the contribution of firm-level distortions to aggregate productivity. As shown in the next subsection, what matters for aggregate productivity growth is the change in the relative level of distortions and not in their absolute level. The elasticity of demand is pinned down using estimates of the elasticity of substitution between goods. For simplicity, I assume that the elasticity is identical across sectors. I set $\theta = 0.8$, which corresponds to an elasticity of substitution between goods of 5, in line with Broda and Weinstein (2006) estimates.[22] Finally, capital elasticities are derived by assuming constant return to scale in each sector $\alpha_s = 1 - \beta_s$. Once the factor elasticities are known, firm-level

[20]Note that the presence of unobserved allocation distortion also rules out the use of standard semi-parametric methods such as Olley Pakes or Levinshon Petrin. These methods, which consists in using a proxy for productivity, can accommodate only a unique unobservable state variable. In the present framework, firms decisions depend on productivity, but also on capital and labor distortions, which raise to three the number of unobservable state variables.

[21]More specifically, it is the average of the inverse of sector specific distortions that is assumed to be zero. Actually assuming an average of zero leads to estimate $\beta_s\theta_s$ as $1/(\frac{1}{T}\sum_t\frac{P_{st}Y_{st}}{w_tL_{st}})$. The results are not affected to this choice

[22]For the 3-digit aggregation level, they report a mean of 6.8 over the period 1972-1988 and of 4 over the period 1990-2001.

productivities can be estimated from output and input levels using equation (5).

2.4.2.2 Estimation of firm-level distortions

Firm-level distortions (τ_i^K, τ_i^L) are computed as the wedge between the marginal productivity of labor and capital and factor prices (equations (1) and (2)). Firm-level distortions are therefore sensitive to factor prices, which are known to be difficult to measure. Luckily, both changes in sectoral and aggregate TFP do not depend on the absolute level of distortions, but only on changes in relative distortions. Hence, the value chosen for factor prices has no impact on aggregate TFP growth and on its decomposition. To illustrate this point, suppose that we use biased measures of factor prices ($\hat{r}_t \neq r_t$ and $\hat{w}_t \neq w_t$) to measure firm-level distortions. This leads to the following biased measures:

$$\begin{aligned} 1+\widehat{\tau_{it}^K} &= (1+\tau_{it}^K)\frac{r_t}{\widehat{r_t}} \\ 1+\widehat{\tau_{it}^L} &= (1+\tau_{it}^L)\frac{w_t}{\widehat{w_t}} \end{aligned}$$

Using equation (1), it can be shown that the obtained measure of sectoral TFP is unbiased despite the measurement errors on firm-level distortions.

$$\begin{aligned} \widehat{TFP_{st}} &= \frac{(r_t/\widehat{r_t})^{\frac{-\alpha_s\theta_s}{1-\gamma\theta_s}}(w_t/\widehat{w_t})^{\frac{-\beta_s\theta_s}{1-\gamma\theta_s}}}{\left((r_t/\widehat{r_t})^{\frac{-(1-\beta_s\theta_s)}{1-\gamma\theta_s}}(w_t/\widehat{w_t})^{\frac{-\beta_s\theta_s}{1-\gamma\theta_s}}\right)^{\alpha_s\theta_s}\left((r/\widehat{r_t})^{\frac{-\alpha_s\theta_s}{1-\gamma\theta_s}}(w_t/\widehat{w_t})^{\frac{-(1-\alpha_s\theta_s)}{1-\gamma\theta_s}}\right)^{\beta_s\theta_s}} TFP_{st} \\ &= TFP_{st} \end{aligned}$$

At the aggregate level, we can show that this measurement error in factor prices has no impact on changes in aggregate productivity if the returns to scale and demand parameters are homogenous across sectors ($\gamma_s = \gamma$ and $\theta_s = \theta$).The measured sectoral distortions read:

$$\frac{d\widehat{\omega_t^K}}{1+\widehat{\omega_t^K}} = \frac{d\omega_t^K}{1+\omega_t^K} + \frac{d(r_t/\widehat{r_t})}{r_t/\widehat{r_t}}$$
$$\frac{d\widehat{\omega_t^L}}{1+\widehat{\omega_t^L}} = \frac{d\omega_t^L}{1+\omega_t^L} + \frac{d(w_t/\widehat{w_t})}{w_t/\widehat{w_t}}$$

Using equation (8) and replacing with the value of ε_K and ε_L, it comes:

$$\begin{aligned}\Delta\widehat{\text{AE}}_{\text{between}} &= \Delta\overline{\text{AE}}_{\text{between}} + \frac{1}{1-\gamma\theta}\left(\varepsilon_K(1-\beta^K) + \varepsilon_L\alpha^L - \alpha^Y\right)\frac{d(r_t/\widehat{r_t})}{r_t/\widehat{r_t}} \\ &\quad + \frac{1}{1-\gamma\theta}\left(\varepsilon_K\beta^K + \varepsilon_K(1-\alpha^L) - \beta^Y\right)\frac{d(w_t/\widehat{w_t})}{w_t/\widehat{w_t}} \\ &= \Delta\overline{\text{AE}}_{\text{between}}\end{aligned}$$

Therefore, the values chosen for factor prices have no impact on the measure of changes in aggregate and sectoral productivity.

2.4.3 Empirical results

In line with the rest of the literature, I implement the decomposition on the manufacturing industry.[23] First, I investigate the role of allocative efficiency, firm-level efficiency and the extensive margin in explaining the dynamics of productivity growth within sectors in the manufacturing industry. Then, I present the contribution of each component to the industry-wide productivity growth, and show the role of between-sector allocative efficiency. Finally, I compare these aggregate results to the standard decomposition used in the literature.

2.4.3.1 Decomposition at the sectoral and aggregate levels

Table 1 gives the average productivity growth for each sector as well as its decomposition in a firm-level efficiency, allocative efficiency and an extensive margin component. This de-

[23] I dropped the electronic components sector as the estimation led to $\alpha<0$.

composition is computed using the exact decomposition described in Appendix B. Over the period 1991-2006, average productivity grew by 2.3% in those sectors. Firm-level efficiency and allocative efficiency contribute both significantly to average TFP growth. While the contribution of firm-level efficiency is positive in virtually all sectors (+3.9 percentage points on average), changes in allocative efficiency tend to decrease sectoral productivity (-1.3 p.p.). By contrast, the contribution of the extensive margin appears to be negligible for most sectors. Entry and exit flows have increased sectoral productivity by an average of 0.1 percentage points. As shown in Appendix D, net entry rates are positive and quite large (average of 2 p.p.) for most years, which tends to enhance aggregate productivity as the latter increase with the number of firms in the industry. This effect is however partially offset by the relatively small size of entrants, suggesting that entrants face either more distortion or have lower T FP than exiting firms. Figure A.2, A.3 A.4 in Appendix D depicts the TFP and distortion distributions of exiting and entering firms. These figures indicate that the TFP of entrants is larger than that of exiting firms, but that entering firms face higher distortions especially in the labor market. This result is line with empirical studies that point to the existence of important financial constraints that hampers the development of new firms. These frictions limits the impact of entering firms on aggregate productivity. [24]

Let us analyze the role of each of these components for the volatility of sectoral productivity growth. The variance of sectoral productivity growth can be computed as follows:

$$V(\Delta TFP) = Cov(\Delta TFP, \Delta TE) + Cov(\Delta TFP, \Delta AE) + Cov(\Delta TFP, \Delta EX)$$

The contribution of firm-level efficiency is therefore measured as $Cov(\Delta TFP, \Delta TE)/V(\Delta\Delta TFP)$[25]. It gives the variation in ΔTFP which is due to variations in within-firm productivity ΔTE, both directly and through its correlations with ΔEX and ΔAE. Correspondingly, the con-

[24]Note also that we measure here only the impact of the extensive margin on the short run productivity. On the long run, the impact of entering firms could be higher if entering firms are characterized by a higher productivity growth rate.

[25]Note that this is the parameter obtained when regressing ΔTE on ΔTFP. As noted by Fujita and Ramey (IER, 2009), this measure is conceptually equivalent to the beta used in finance.

Table 1: Sectoral productivity decomposition, 1991-2006 average (in %)

Sector	ΔTFP_s	ΔTE_s	ΔAE_s	ΔEX_s
Food products*	0.54	0.09	1.01	-0.32
Wearing apparel and leather products	3.53	6.88	-3.97	1.10
Printing and reproduction	2.03	3.14	-0.84	0.15
Pharmaceutical and perfumes products	3.52	2.91	2.26	-0.11
Furniture	1.95	4.75	-2.74	0.14
Motor vehicle	0.29	2.59	-1.14	0.09
Other transportation equipment	2.97	8.33	-4.65	-0.06
Mechanical equipments*	2.94	4.08	-1.46	0.46
Mineral products	1.74	3.70	-1.75	0.10
Textiles	1.80	5.18	-3.64	0.58
Wood and paper products	1.79	2.77	-1.04	0.14
Rubber and plastics products*	4.95	5.60	-0.76	0.39
Fabricated metal products*	0.26	1.61	-1.35	0.15
Electronic products	6.26	8.54	-2.20	0.20
Weighted average	2.29	3.90	-1.33	0.13

Note: This table presents average sectoral productivity growth over the period 1991-2006, as well as the average level of its components. The first column gives the average sectoral productivity growth ΔTFP_s, the second column gives the average change in firm-level efficiency ΔTE_s, the third column gives the average change in allocative efficiency ΔAE_s, and the last column gives the average change in sectoral productivity due to the extensive margin ΔEX_s. Sectoral productivity has been computed using equation (1) and the components ΔTE_s , ΔAE_s and ΔEX_s have been computed using the formulas given in Appendix B. Because of approximation errors, the sum of the components is not exactly equal to sectoral productivity growth. Each sector denoted with * represents more than 10% of the total manufacturing industry value added.

tribution of within-sector allocative efficiency and the extensive margin are measured as $Cov(\Delta TFP, \Delta AE)/V(\Delta TFP)$ and $Cov(\Delta TFP, \Delta EX)/V(\Delta TFP)$. Table 2 reports the contribution of each component to the volatility of sectoral TFP.

Firm-level efficiency appears to be the main driver of sectoral productivity dynamics. In fact, changes in firm-level efficiency tend to exacerbate the volatility of sectoral productivity growth in many sectors. On the contrary, the contribution of changes in allocative efficiency to the volatility of productivity is negative, and allocative efficiency component thus dampens the movement in sectoral productivity growth. In absolute terms, the contribution of allocative efficiency is smaller than firm-level efficiency but still large. The extensive margin component plays a negligible role in the volatility of sectoral productivity growth in virtually all sectors.

Table 2: Contribution to sectoral volatility, 1991-2006

Sector	Δ TFP stand. dev.	contribution to ΔTFP volatility		
		ΔTE_s	ΔAE_s	ΔEX_s
Food products*	2.36 %	1.26	-0.30	0.04
Wearing apparel and leather products	4.34 %	1.21	-0.33	0.12
Printing and reproduction	2.69 %	1.98	-1.05	0.07
Pharmaceutical and perfumes products	2.85 %	2.26	-1.20	-0.06
Furniture	2.66 %	0.91	0.01	0.08
Motor vehicle	9.35 %	1.54	-0.54	0.00
Other transportation equipment	5.98 %	1.47	-0.48	0.00
Mechanical equipments*	4.16 %	1.34	-0.39	0.05
Mineral products	3.74 %	1.36	-0.37	0.02
Textiles	3.96 %	1.00	-0.03	0.03
Wood and paper products	5.04 %	1.10	-0.09	-0.01
Rubber and plastics products*	3.72 %	1.38	-0.30	-0.07
Fabricated metal products*	6.39 %	1.15	-0.19	0.04
Electronic products	5.22 %	1.02	-0.09	0.07

Note: This table presents the contribution of each components to the volatility of sectoral productivity growth. The first column gives the standard deviation of sectoral productivity growth. The second column gives the contribution of firm-level efficiency (ΔTE_s) to the volatility of sectoral productivity growth. The third column gives the contribution of allocative efficiency (ΔAE_s), and the last column gives the contribution of the extensive margin (ΔEX_s). The contributions are computed as described in the text. The sum of the contribution of ΔTE_s, ΔAE_s and ΔEX_s equals 1. To avoid the discrepancies caused by approximation errors, the contributions to volatility have been computed with respect to the volatility of ($\Delta\text{TE}_s + \Delta\text{AE}_s + \Delta\text{EX}_s$). Each sector denoted with * represents more than 10% of the total manufacturing industry value added.

Table 3 reports the correlation of aggregate productivity growth and each of its components with the sector's real value added growth. Sectoral TFP growth is highly correlated with changes in the sector's activity. The firm-level efficiency component also exhibits a positive, though smaller, correlation. This result is not surprising, since in this framework firm-level productivity depends both on technical and demand factors, a higher sectoral demand then leads to a higher within-firm productivity component.[26] In most sectors, the extensive margin component also is procyclical. This procyclicality contrasts with theories on the cleansing effect of recessions, according to which the higher exit rate of low productivity firms in recessions would enhance aggregate productivity. The results suggest that it is the reallocation of

[26]Recall that firm-level productivity reads $A_{it} = z_{it}^{\theta_s} b_{it}^{1-\theta}/P_{st}$, where z_{it} is technical efficiency and b_{it} is the firm-specific demand shock.

Table 3: Correlation with sectoral value added growth, 1991-2006

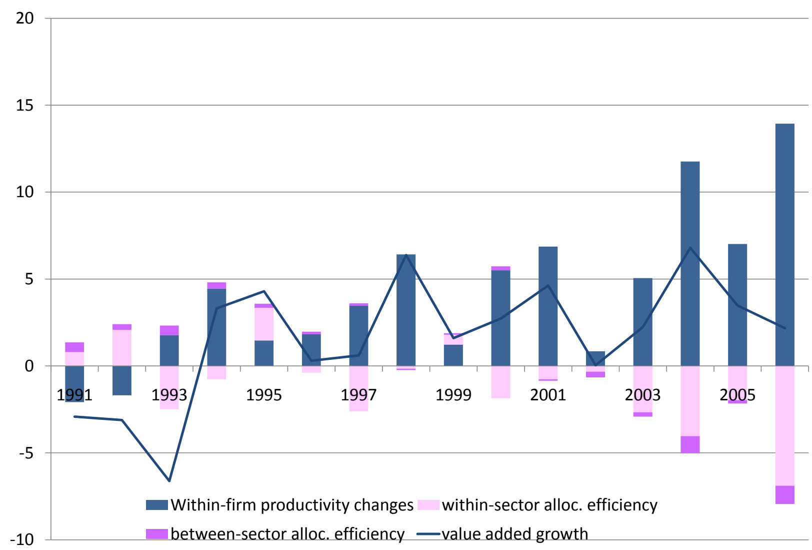
Sector	ΔTFP_s	ΔTE_s	ΔAE_s	ΔEX_s
Food products*	0.84	0.30	0.04	0.40
Wearing apparel and leather products	0.93	0.58	-0.10	0.81
Printing and reproduction	0.70	0.38	-0.18	0.46
Pharmaceutical and perfumes products	0.92	0.64	-0.46	-0.03
Furniture	0.84	0.47	0.11	0.31
Motor vehicle	0.84	0.54	-0.21	-0.17
Other transportation equipment	0.94	0.89	-0.66	0.09
Mechanical equipments*	0.91	0.80	-0.41	0.48
Mineral products	0.97	0.73	-0.31	0.28
Textiles	0.86	0.49	0.23	-0.21
Wood and paper products	0.96	0.86	-0.11	-0.01
Rubber and plastics products*	0.90	0.58	-0.08	-0.26
Fabricated metal products*	0.97	0.90	-0.34	0.35
Electronic products	0.89	0.82	-0.18	0.36

Note: This table presents the correlation of sectoral productivity growth and its components with real value added growth. The first column gives the correlation of sectoral productivity growth (ΔTFP_s) to real value added growth. The second column gives the correlation of firm-level efficiency (ΔTE_s). The third column gives the contribution of allocative efficiency (ΔAE_s), and the last column that of the extensive margin (ΔEX_s). Each sector denoted with * represents more than 10% of the total manufacturing industry value added.

resources between continuing firms, and not between entering and exiting firms, that tends to raise aggregate productivity during recessions. For most sectors, allocative efficiency is indeed countercyclical.

Figure 1 and 2 present the results at the aggregate level computed using equations (7) to (10) (the numbers are provided in Appendix C). These figures, which display the value added growth of the entire manufacturing industry, complete the picture given by sectoral data. They give the aggregation of the sectoral firm-level efficiency, allocative efficiency and extensive margin components, as well as the contribution of inter-sectoral allocative efficiency. Like at the sectoral level, changes in within-firm productivity growth and in allocative efficiency are the main determinants of aggregate productivity growth. Figure 1 and 2 show that the variation of aggregate productivity due to the extensive margin is small compared to the other components. While the firm-level efficiency component is procyclical, the overall allocative efficiency component is countercyclical. Their correlation with the manufacturing industry's value added is respectively 0.64 and -0.25. Note that the between-sector allocative efficiency is more procyclical (-0.55) than within-sector allocative efficiency (-0.17). Figure A.6 and A.7 in Appendix D show that the overall picture is not modified when a higher elasticity of substitution between goods is chosen (equal to 6).

Figure 1: Firm-level productivity and allocative efficiency (manufacturing industry)

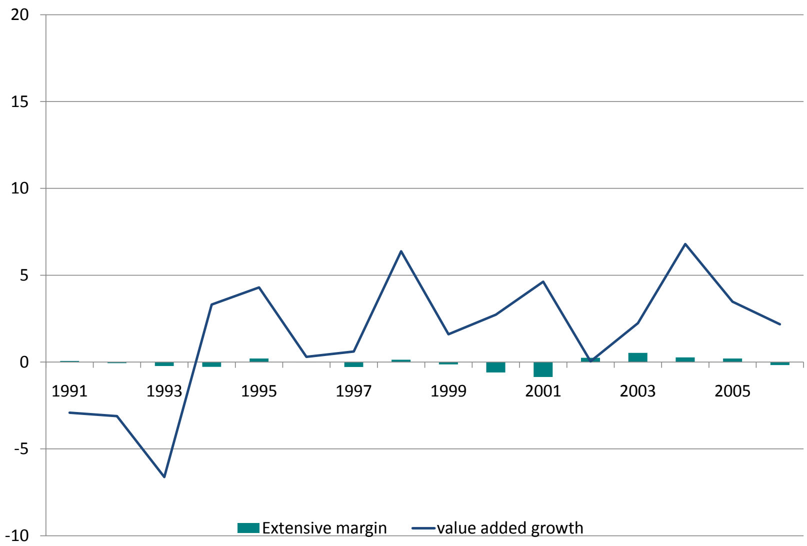


2.4.3.2 Comparison with the existing approach

I now show how these results differ from those obtained with the approach used in the applied literature. With this approach, aggregate productivity is not derived from an aggregate production function but is computed as the weighted average of firm-level TFP.

$$\ln TFP_t = \sum_i s_{it} \ln A_{it},$$

Figure 2: The extensive margin (manufacturing industry)



where $s_{it} = Y_{it}/Y_t$ is the output share of firm i.

Let us consider the decomposition proposed by Foster, Haltiwanger and Krizan (hereafter FHK, 2001):

$$\begin{aligned}\Delta \ln TFP_t \quad = \quad & \underbrace{\sum_{\text{stay}} s_{it-1}\Delta \ln A_{it}}_{\text{within}} + \underbrace{\sum_{\text{stay}} \Delta s_{it}(\ln A_{it-1} - \ln TFP_{t-1})}_{\text{reallocation}} \\ & + \underbrace{\sum_{\text{entry}} s_{it}(\ln A_{it} - \ln TFP_{t-1}) - \sum_{\text{exit}} s_{it-1}(\ln A_{it-1} - \ln TFP_{t-1})}_{\text{extensive margin}} \\ & + \underbrace{\sum_{\text{stay}} \Delta \ln A_{it}\Delta s_{it}}_{\text{cross term}}\end{aligned} \tag{2}$$

The first component captures changes in within-firm productivity holding fixed the market shares. The second component measures the impact of changes in market shares on aggregate productivity. The third component measures the contribution of entering and exiting firms, and the last component is a cross term that measures the correlation between changes in within-firm productivity and changes in market shares. This decomposition is completely different from the one derived from the aggregation of firm-level production functions (equations (2) to (4), and (7) to (10)). The main differences lie in the measure of the contribution of the extensive margin and that of cross-sectional efficiency. In the FHK decomposition, the contribution of entry and exit depends on the relative TFP of entering an exiting firms. Compared to equation (4), the FHK component does not take into account the impact of the distortions face by entering and exiting firms on aggregate productivity growth. Furthermore, this measure does capture the impact of entry and exit and the total number of firms in the industry. When firms use a decreasing returns to scale technology or when consumers have a love for variety, entry and exit flows affect aggregate productivity not only through a composition effect, but also trough their impact on the number of firms. The FHK measure of cross-sectional allocation is also completely different from my approach. Changes in cross-sectional allocation are measured by the correlation between changes in market shares and firm-level productivity. This measure captures changes in allocative efficiency when goods are homogeneous and produced using a technology with constant returns to scale and only one input. However, in the more general case where goods are heterogeneous or the marginal productivity of one input is decreasing, reallocating resources towards high TFP firms could decrease aggregate output. In this case, allocative efficiency depends on the dispersion of the value of marginal productivity and not on the correlation between market shares and firm-level productivity.

To illustrate these differences, I estimate the decomposition proposed by FHK and compare them with my results. Figure A.5 in Appendix C shows that the dynamics of *average* productivity given by the FHK measure is very close to that of *aggregate* productivity. The discrepancies are, however, substantial for the decomposition which are depicted in Figure 2.4.3.2 and 2.4.3.2 (numbers are provided in appendix D). The results are qualitatively similar to those obtained by FHK on the US manufacturing sector. The cross term is the main determinant of the mean and the volatility of aggregate productivity growth. On average, the net entry component contributes positively and reallocation negatively to aggregate productivity growth. When compared with Figure 1 and 2, the year-to-year changes in the contribution of the extensive margin, within-firm productivity and reallocations are considerably different. In particular, entry and exit play a relatively larger role for aggregate productivity growth in the FHK decomposition. Computed in absolute values, the contribution of entry exit is about 2 times lower than that of input reallocations in the FHK decomposition, and 7 times lower using my decomposition. Furthermore, the cyclical properties of the reallocation component are at odds with my findings. While my results indicate a countercyclical allocation efficiency, the reallocation component is here positively correlated with value added growth (correlation of 0.19). All in all, these results illustrate that the micro determinants of average and aggregate productivity are substantially different.

Figure 3: The FHK decomposition

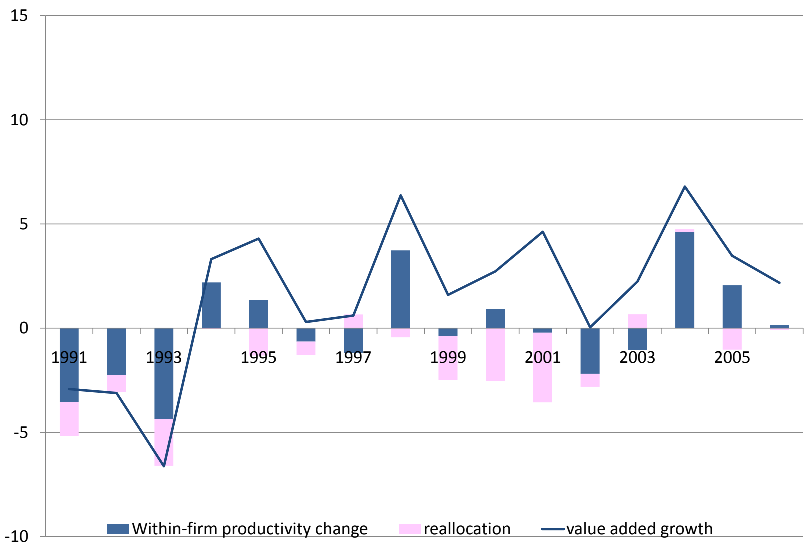
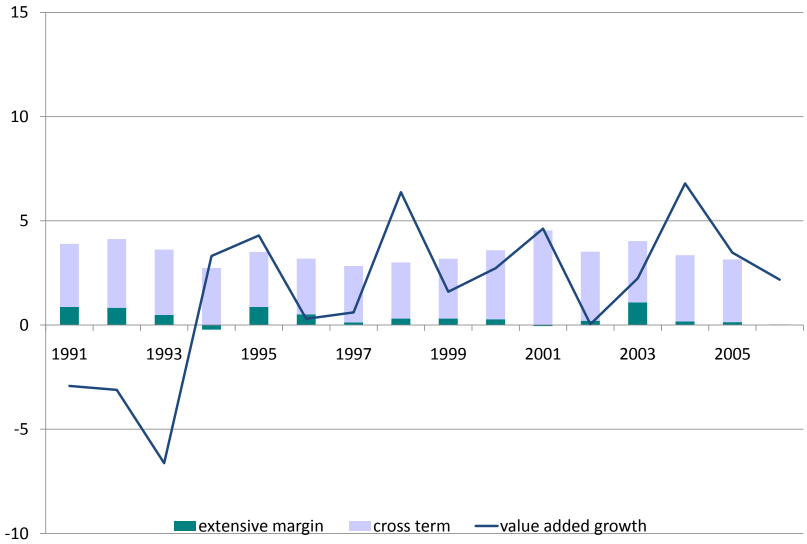


Figure 4: The FHK decomposition



2.5 Conclusion

This paper proposes a novel approach for analyzing the micro determinants of aggregate productivity dynamics. After deriving aggregate productivity from the aggregation of firm-level production functions, I show how to decompose aggregate productivity into changes in firm-level efficiency, changes in allocative efficiency and changes in entry and exit flows. Contrary to both existing empirical and theoretical works that emphasize the role of entry and exit for aggregate productivity growth, this paper shows that the extensive margin plays a negligible role in the dynamics of aggregate productivity growth and highlights the contribution of changes in the efficiency of resource allocation across firms. I show that changes in allocative efficiency are countercyclical and tend to stabilize the movements in aggregate productivity growth. Furthermore, I find no evidence in support of the cleansing effect of recession as the extensive margin component is not only small but also procyclical. However, it must be emphasized that these results give only the contemporaneous impact of entry and exit on aggregate productivity growth. The effects of entry and exit flows on long run productivity growth are not captured in this paper. New firms have a small contribution to aggregate productivity growth in the year they enter, but may have a large contribution on the long run if they have a higher productivity growth than exiting firms. Finally, the results calls for further research on the behavior of allocative efficiency over the business cycle. In fact, a deeper understanding of the dynamics of allocative efficiency would require a more precise modelling of frictions. Here, any distortion that generate a wedge between the firms' value of marginal productivity is captured in our measure of allocative efficiency, and we cannot therefore identify among the various sources of frictions those that are the more relevant.

Appendix A: Deriving the sectoral aggregate production function

In this Appendix, we show how to derive the sectoral production functions. When factor and demand elasticities are identical across firms the allocation rules can be derive explicitly. Using equations (6) and (7), we can write:

$$K_{it} = \left(\frac{A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}}{\sum A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}} \right) K_{st}$$

$$L_{it} = \left(\frac{A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}}}{\sum A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}}} \right) L_{st}$$

and the sectoral production function is then:

$$F_s(K_{st}, L_{st}, \widetilde{A}_{st}, \widetilde{\tau}_{st}) = \sum_i A_{it} \left(\frac{g_s^K(A_{it}, \tau_{it})}{\sum_i g_s^K(A_{it}, \tau_{it}} K_{st} \right)^{\alpha_s\theta_s} \left(\frac{g_s^K(A_{it}, \tau_{it})}{\sum_i g_s^K(A_{it}, \tau_{it}} L_{st} \right)^{\beta_s\theta_s}$$

which can be written

$$F_s(K_{st}, L_{st}, \widetilde{A}_{st}, \widetilde{\tau}_{st}) = TFP_{st} K_{st}^{\alpha_s\theta_s} L_{st}^{\beta_s\theta_s}$$

Hence, sectoral TFP is a function of both firm-level productivity $\widetilde{A}_{st}$ and distortions $\widetilde{\tau}_{st}$:

$$TFP_{st} = \frac{\sum_{i=1}^{N_s} g_s^Y(A_{it}, \tau_{it})}{\left(\sum_{i=1}^{N_s} g_s^K(A_{it}, \tau_{it})\right)^{\alpha_s\theta_s} \left(\sum_{i=1}^{N_s} g_s^L(A_{it}, \tau_{it})\right)^{\beta_s\theta_s}}.$$

where the g functions are given by:

$$g_s^Y(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}$$

$$g_s^K(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{1-\beta_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{\beta_s\theta_s}{1-\gamma\theta_s}}$$

$$g_s^L(A_{it}, \tau_{it}) = A_i^{\frac{1}{1-\gamma\theta_s}} (1+\tau_i^K)^{-\frac{\alpha_s\theta_s}{1-\gamma\theta_s}} (1+\tau_i^L)^{-\frac{1-\alpha_s\theta_s}{1-\gamma\theta_s}}$$

Appendix B: Decomposition of sectoral productivity: Exact and approximate formulas

This appendix presents the decomposition of sectoral TFP. First, I describe the method and show how changes in sectoral TFP can be approximated. Then, I give the formulas for the exact decomposition.

From equation (1), the change in sectoral aggregate productivity can be computed as:

$$\frac{\Delta TFP_{st}}{TFP_{st}} \simeq \ln\left(\frac{\sum g_s^Y(A_{it},\tau_{it})}{\sum g_s^Y(A_{it-1},\tau_{it-1})}\right) - \alpha_s\theta_s \ln\left(\frac{\sum g_s^K(A_{it},\tau_{it})}{\sum g_s^K(A_{it-1},\tau_{it-1})}\right) - \beta_s\theta_s \ln\left(\frac{\sum g_s^L(A_{it},\tau_{it})}{\sum g_s^L(A_{it-1},\tau_{it-1})}\right)$$

Sectoral TFP growth can then be decompose into changes in within-firm productivity (ΔTE_s), changes in allocative efficiency (ΔAE_s) and changes in the extensive margin (ΔEX_s), by decomposing each component as follows:

$$\ln\frac{\sum g_s^Y(A_{it},\tau_{it})}{\sum g_s^Y(A_{it-1},\tau_{it-1})} = \underbrace{\ln\frac{\sum_{i\in C_s} g_s^Y(A_{it},\tau_{it-1})}{\sum_{i\in C_s} g_s^Y(A_{it-1},\tau_{it-1})}}_{\Delta \text{TE}_Y} + \underbrace{\ln\frac{\sum_{i\in C_s} g_s^Y(A_{it},\tau_{it})}{\sum_{i\in C_s} g_s^Y(A_{it},\tau_{it-1})}}_{\Delta \text{AE}_Y} + \underbrace{\ln\frac{1-\frac{\sum_{i\in X_s} g_s^Y(A_{it-1},\tau_{it-1})}{\sum_{i=1}^{N_{t-1}} g_s^Y(A_{it-1},\tau_{it-1})}}{1-\frac{\sum_{i\in E_s} g_s^Y(A_{it},\tau_{it})}{\sum_{i=1}^{N_t} g_s^Y(A_{it},\tau_{it})}}}_{\Delta \text{EX}_Y}$$

with C_s is the set of continuing firms ($I_{it-1} \geq 0$ and $I_{it} \geq 0$), E_s the set of new entrants ($I_{it-1} < 0$ and $I_{it} \geq 0$) and X_s the set of exiting firms ($I_{it-1} \geq 0$ and $I_{it} < 0$) in sector s. Aggregate productivity growth can then be expressed as:

$$\frac{\Delta TFP_{st}}{TFP_{st}} \simeq \Delta \text{TE}_s + \Delta \text{AE}_s + \Delta \text{EX}_s,$$

where the within-firm productivity component is defined as $\Delta \text{TE}_s = \Delta \text{TE}_Y - \alpha_s\theta_s \Delta \text{TE}_K - \beta_s\theta_s \Delta \text{TE}_L$, and the allocative efficiency ΔAE_s and the extensive margin ΔEX_s components are symmetrically defined.

The decomposition between changes in within-firm productivity and changes in allocative efficiency is similar to the decompositions we can find in the index number literature. In the expression given above the effects of within-firm productivity changes are measured with a Laspeyres-like index, while those of allocative efficiency (ΔAE_Y) are measured with a Paasche-like index. Using the expressions for g_s^Y, g_s^K and g_s^L, together with the approximation $ln(1+x) \simeq x$ yield the expression given in equations (2) to (4).

For simplicity, the main text only provides approximations. The decomposition is computed from firm-level data using the exact formulas. The change in TFP can be broken down into three components.

$$\frac{TFP_{st}}{TFP_{st-1}} = ITE\ IAE\ IEX$$

The extensive margin component is measured as follows:

$$IEX = \frac{\frac{1-\sum_{X_s}\frac{P_{it-1}Y_{it-1}}{P_{st-1}Y_{t-1}}}{1-\sum_{E_s}\frac{P_{it}Y_{it}}{P_{st}Y_t}}}{\left(\frac{1-\sum_{X_s}\frac{K_{it-1}}{K_{t-1}}}{1-\sum_{E_s}\frac{K_{it}}{K_t}}\right)^{\alpha_s\theta_s}\left(\frac{1-\sum_{X_s}\frac{L_{it-1}}{L_{t-1}}}{1-\sum_{E_s}\frac{L_{it}}{L_t}}\right)^{\beta_s\theta_s}}$$

To avoid any asymmetry induced by the functional form of the indexes, I measure the firm-level and allocative efficiency component with a Fisher-like index, i.e. the geometric mean of the Laspeyres-like and Paasche-like indexes. The Fisher-like index for within-productivity changes is computed as follows:

$$ITE = ITE0^{0.5} ITE1^{0.5}$$

where

$$ITE0 = \frac{\sum \frac{A_{it}}{A_{it-1}}^{\frac{1}{1-\gamma\theta_s}} \frac{p_{it}Y_{it-1}}{\sum p_{it}Y_{it-1}}}{\left(\sum \frac{A_{it}}{A_{it-1}}^{\frac{1}{1-\gamma\theta_s}} \frac{K_{it-1}}{\sum K_{it-1}}\right)^{\alpha_s\theta_s}\left(\sum \frac{A_{it}}{A_{it-1}}^{\frac{1}{1-\gamma\theta_s}} \frac{L_{it-1}}{\sum L_{it-1}}\right)^{\beta_s\theta_s}}$$

$$ITE1 = \frac{\left(\sum \frac{A_{it-1}}{A_{it}}^{\frac{1}{1-\gamma\theta_s}} \frac{K_{it}}{\sum K_{it}}\right)^{\alpha_s\theta_s}\left(\sum \frac{A_{it-1}}{A_{it}}^{\frac{1}{1-\gamma\theta_s}} \frac{L_{it}}{\sum L_{it}}\right)^{\beta_s\theta_s}}{\sum \frac{A_{it-1}}{A_{it}}^{\frac{1}{1-\gamma\theta_s}} \frac{p_{it}Y_{it}}{\sum p_{it}Y_{it}}}$$

The change in allocative efficiency is similarly computed. Then, each component of aggregate productivity growth is computed as $\Delta TE = ITE - 1$, $\Delta AE = IAE - 1$ and $\Delta EX = IEX - 1$.

I now provide some indications for the decomposition of aggregate productivity growth across sectors. The heterogeneity in elasticities makes the calculation more tedious. For simplicity, I will describe the method in the case where there is only one input ($\beta = 0$). In this case $Y = \sum_s TFP_s K_s^{\alpha_s}$, with $K_s = K_s(P_s TFP_s, \omega_s, r)$

Taking the derivative with respect to TFP_s gives:

$$\frac{\partial F}{\partial TFP_s}\frac{TFP_s}{Y} = \sum_s \frac{Y_s}{Y}\frac{dTFP_s}{TFP_s} + \frac{\alpha_s}{1-\alpha_s}\frac{Y_s}{Y}\frac{dTFP_s}{TFP_s} + \alpha_s \frac{Y_s}{Y}\frac{\partial K_s}{\partial r}\frac{r}{K_s}\frac{dr}{r}$$

where dr/r is the change in interest rate induced by changes in firm-level efficiency holding fixed aggregate inputs and distortions. Using the implicit function theorem, this change can be computed as:

$$\begin{aligned} 0 &= -\sum_s \frac{1}{1-\alpha_s} K_s \frac{dr}{r} + \sum_s \frac{1}{1-\alpha_s} K_s \frac{dTFP_s}{TFP_s} \\ \frac{dr}{r} &= \frac{1}{\sum_s \frac{1}{1-\alpha_s} K_s} \sum_s \frac{1}{1-\alpha_s} K_s \frac{dTFP_s}{TFP_s} \end{aligned}$$

The derivative with respect to TFP_s can then be written:

$$\begin{aligned} \frac{\partial F}{\partial TFP_s}\frac{TFP_s}{Y} &= \sum_s \frac{1}{1-\alpha_s}\frac{Y_s}{Y} - \left(\frac{\sum_s \frac{\alpha_s}{1-\alpha_s}\frac{Y_s}{Y}}{\sum_s \frac{1}{1-\alpha_s}\frac{K_s}{K}} \right) \frac{1}{1-\alpha_s}\frac{K_s}{K}\frac{dTFP_s}{TFP_s} \\ &= \sum_s \frac{1}{1-\alpha_s}\left(\frac{Y_s}{Y} - \varepsilon_K \frac{K_s}{K} \right) \frac{dTFP_s}{TFP_s}. \end{aligned}$$

which gives the weights use in equations (7) to (9) with $\beta = \varepsilon_L = 0$. The changes in between-sector allocative efficiency are computed similarly.

Appendix C: Data

Data used are collected by the French tax administration and controlled for inconsistencies and combine with the responses to the annual business surveys in the INSEE unified system for business statistics (SUSE). I consider firms that declare under the “normal" regime (SUSE-BRN) from 1989 to 2007. This tax regime is mandatory for firms with sales above 763 000 euros (230 000 in the service sector), but is also widely chosen by firms which are below the threshold: in 2003 46% of the sample restricted to the manufacturing sector had sales below 763 000 euros. As described in section 4.2, this dataset is corrected for spurious entry and exit flows, as well as for outliers in terms of productivity growth.

Description of variables

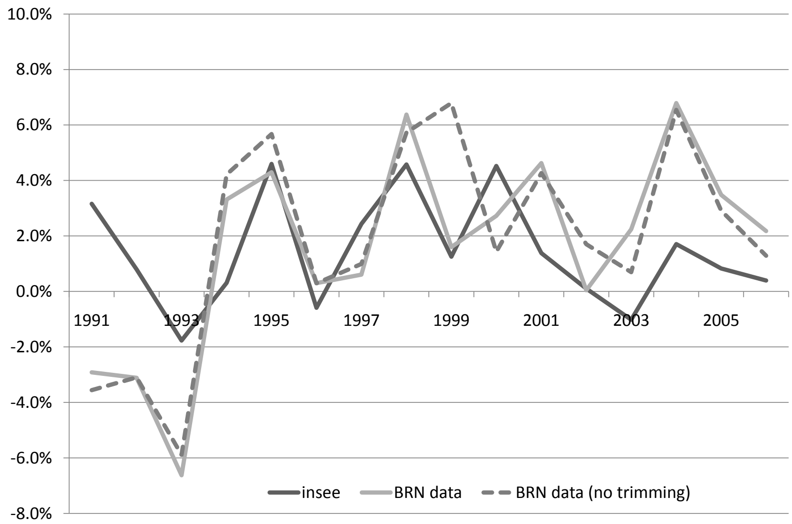
Output: gross value added less operating subsidies plus taxes, deflated by the NES36 sectoral price index published by the national accounts (INSEE, base 2000).

Labour: average number of employees over the year.

Capital: constructed using the perpetual inventory method. Real capital stock are computed from the previous period undepreciated stock and the period's real investment in tangible and intangible assets. I use a linear sector-specific depreciation rate, based on Sylvain (2003) estimates of equipments life span. The the NES36 sectoral investment deflator is derived from the national accounts (INSEE, base 1995 and 2000). As Dunne et al. (1989), the initial year capital stock is estimated by assuming that the firm's relative real capital stock is equal to its relative book value of assets.

Sectoral labor share: computed at the N36 level from BRN data as 1- ratio of the sectoral gross operating surplus over value added.

Figure A.1: Growth of real value added in the manufacturing industry: national accounts vs SUSE-BRN



Appendix D: Additional tables and figures

Table A.1: Entry and Exit rates in the manufacturing industry

	Entry rate	Exit rate	Entry rate (empl.weighted)	Exit rate (empl.weighted)
1991	0.125	0.158	0.044	0.057
1992	0.134	0.170	0.037	0.062
1993	0.212	0.267	0.052	0.085
1994	0.161	0.120	0.048	0.054
1995	0.106	0.114	0.041	0.046
1996	0.095	0.099	0.037	0.048
1997	0.115	0.088	0.041	0.040
1998	0.082	0.073	0.032	0.037
1999	0.080	0.076	0.031	0.039
2000	0.096	0.081	0.035	0.041
2001	0.099	0.099	0.042	0.043
2002	0.081	0.057	0.031	0.038
2003	0.077	0.069	0.030	0.036
2004	0.077	0.067	0.031	0.034
2005	0.080	0.063	0.027	0.034
2006	0.084	0.065	0.029	0.034
average	0.107	0.104	0.037	0.045

Note: This table presents entry and exit rate and their employment weighted counterparts. Note that the data are corrected for temporary exits and outliers as described in section 4.1

Table A.2: Decomposition of the manufacturing industry TFP growth

	ΔTFP	$\Delta\overline{\text{TE}}$	$\Delta\overline{\text{AE}}_{\text{within}}$	$\Delta\overline{\text{EX}}$	$\Delta\overline{\text{AE}}_{\text{between}}$
1991	-0.92	-2.08	0.80	0.06	0.56
1992	0.32	-1.69	2.08	-0.06	0.33
1993	-1.92	1.77	-2.49	-0.23	0.56
1994	4.74	4.44	-0.77	-0.28	0.36
1995	4.01	1.46	1.88	0.21	0.24
1996	1.17	1.82	-0.39	0.00	0.15
1997	0.65	3.46	-2.61	-0.28	0.14
1998	6.24	6.42	-0.16	0.13	-0.08
1999	1.57	1.23	0.57	-0.13	0.09
2000	1.95	5.50	-1.86	-0.60	0.22
2001	2.97	6.87	-0.78	-0.86	-0.08
2002	0.13	0.85	-0.33	0.25	-0.32
2003	2.29	5.05	-2.67	0.53	-0.25
2004	7.28	11.76	-4.04	0.26	-0.98
2005	4.73	7.02	-1.96	0.21	-0.19
2006	2.90	13.94	-6.89	-0.17	-1.05

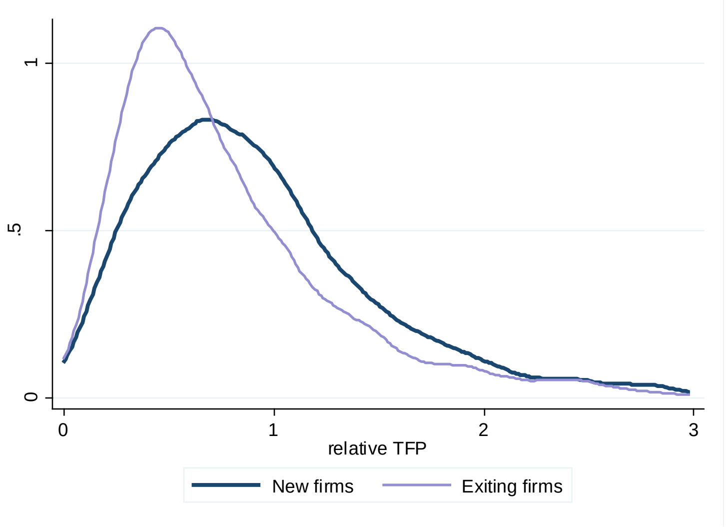
Note: This table presents the decomposition of productivity growth in the manufacturing industry (ΔTFP) into a aggregate firm-level efficiency component ($\Delta\overline{\text{TE}}$), an aggregate within-sector allocative efficiency component ($\Delta\overline{\text{AE}}_{\text{within}}$), an aggregate extensive margin component ($\Delta\overline{\text{EX}}$) and a between-sector allocative efficiency component ($\Delta\overline{\text{AE}}_{\text{between}}$). The aggregate productivity growth has been computed using equation (11) and its components have computed using equations (7) to (10). Because of approximation errors, the sum of the components does not exactly equal changes in aggregate TFP.

Table A.3: Decomposition of the manufacturing industry TFP growth (Foster Haltiwanger and Krizan's method)

	$\Delta \ln TFP$	within	reallocation	cross term	net entry
1991	-1.08	-3.54	-1.63	3.02	0.87
1992	0.99	-2.25	-0.80	3.30	0.83
1993	-3.01	-4.35	-2.25	3.14	0.48
1994	4.65	2.19	-0.03	2.73	-0.22
1995	3.60	1.35	-1.31	2.64	0.87
1996	1.90	-0.65	-0.65	2.67	0.52
1997	2.29	-1.19	0.66	2.70	0.13
1998	6.25	3.73	-0.44	2.69	0.31
1999	0.68	-0.38	-2.11	2.87	0.31
2000	1.93	0.92	-2.53	3.31	0.28
2001	0.92	-0.21	-3.35	4.54	-0.05
2002	0.69	-2.19	-0.62	3.32	0.20
2003	3.59	-1.06	0.67	2.94	1.08
2004	8.01	4.61	0.13	3.18	0.17
2005	4.27	2.06	-1.03	3.01	0.14
2006	3.96	0.14	-0.09	0.03	0.00

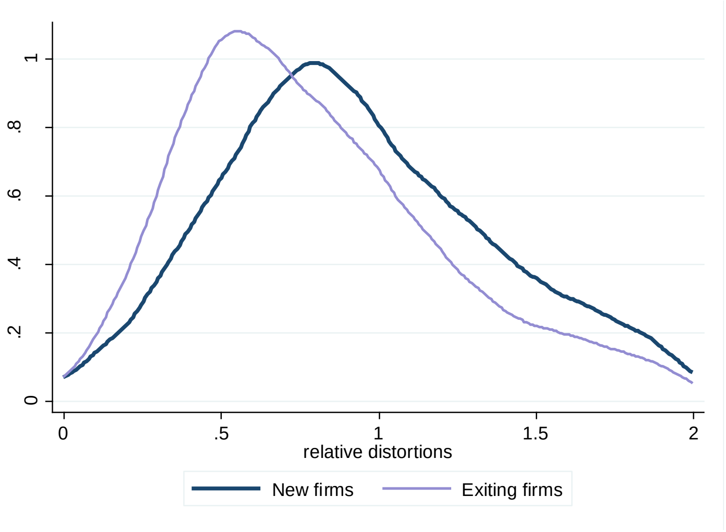
Note: This table presents the decomposition proposed by Foster et al. (2001). The components are computed as described in equation (2).

Figure A.2: New firms vs exiting firms: relative TFP density



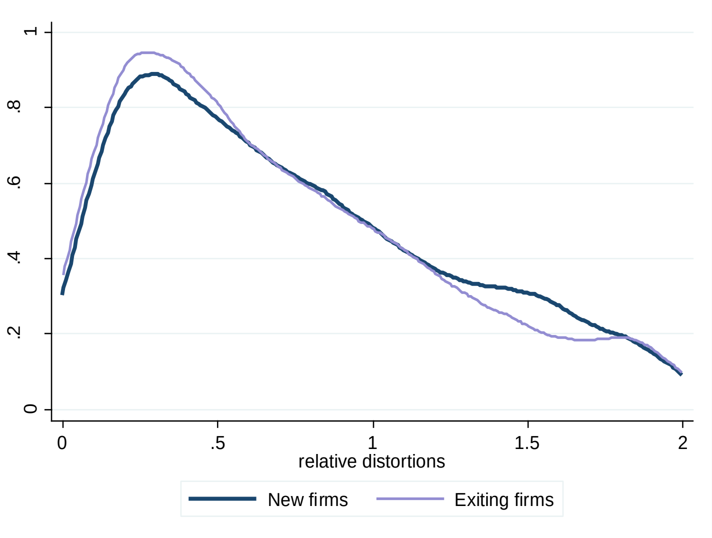
Note: computed in 2000 for the whole manufacturing industry, relative to the median TFP value

Figure A.3: New firms vs exiting firms: relative labor distortions density



Note: computed in 2000 for the whole manufacturing industry, relative to the median distortion value

Figure A.4: New firms vs exiting firms: relative capital distortions density



Note: computed in 2000 for the whole manufacturing industry, relative to the median distortion value

Figure A.5: Aggregate productivity growth: comparison with FHK approach

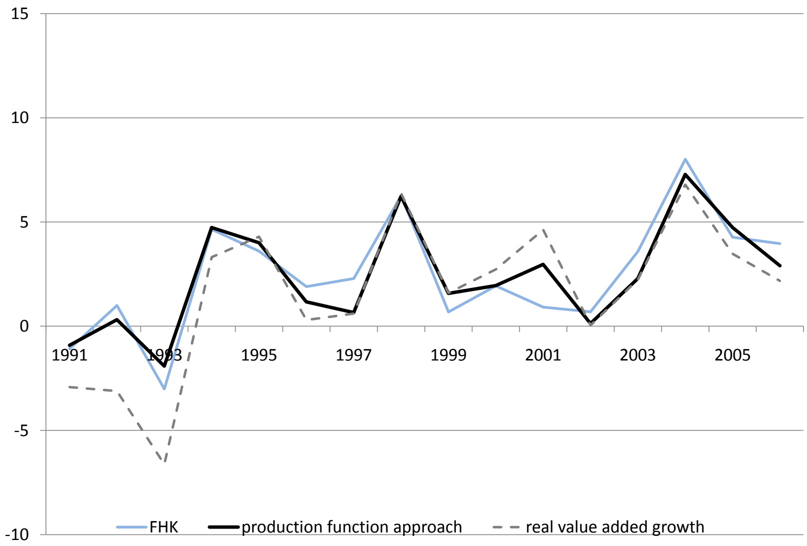


Figure A.6: Firm-level efficiency and allocative efficiency (elasticity of substitution equal to 6)

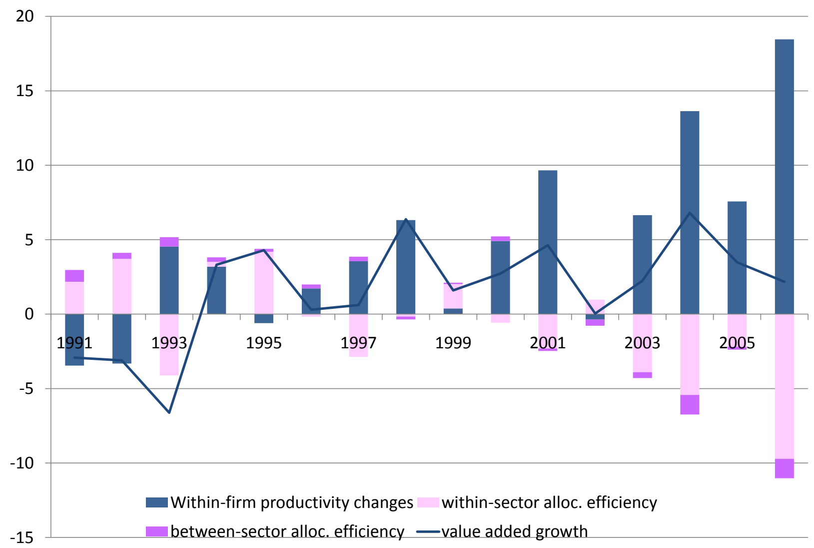
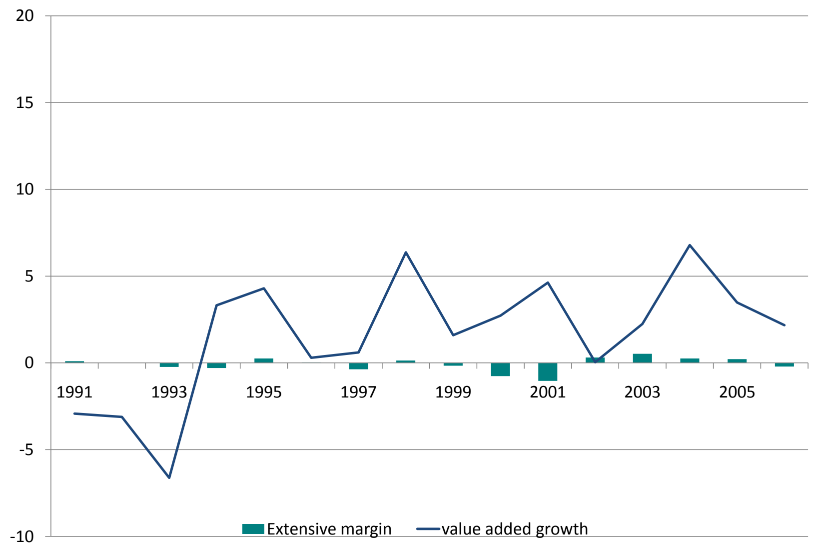


Figure A.7: The extensive margin (elasticity of substitution equal to 6)



Chapter 3

Matching frictions, unemployment dynamics and the cost of business cycles[1]

3.1 Introduction

In a very famous and controversial article, Lucas (1987) argues that the costs of business cycles are negligible: using empirically plausible values for risk aversion, he shows that individuals would only sacrifice a mere 0.008% of their consumption to get rid of all aggregate variability in consumption. This result is completely at odds with the neoclassical synthesis and the Keynesian legacies, but also, to a lesser extent, with a recent literature that documents far larger costs of business cycles (Beaudry and Pages (1999), Krebs (2007), Storesletten et al. (2001), Barlevy (2003), De Santis (2007) and Krusell et al. (2009)).

Our objective in this paper is to show that the matching unemployment theory may also lead to sizable welfare costs of business cycles. Due to strong non-linearities, average employment and therefore average consumption are lowered by the mere process of alternate expansions

[1]This chapter, co-authored with Jean-Olivier Hairault and François Langot, has been published in the *Review of Economic Dynamics*, October 2010.

and contractions: the losses during recessions outweigh the gains during economic booms. The welfare cost of unemployment fluctuations stemming from these non-linearities has been ignored despite the high interest in the recent literature with matching frictions for the cyclical volatility of unemployment (Shimer (2005) and Hall (1991)) and for the design of the optimal monetary policy (Blanchard and Gali (2008), Faia (2009) and Sala et al. (2008)). By filling this gap, our paper adds a new dimension to the analysis of the matching model with aggregate uncertainty.

Following Cole and Rogerson (1999), our analysis is firstly based on the reduced form of the matching model where the job finding and separation rates are considered as exogenous stochastic processes. Focusing on the stock-flow unemployment dynamics allows us to emphasize that the core mechanism underlying our result does not depend on the structural matching considered. Moreover, it clearly unveils the very different implications of fluctuations in the job finding rate and the separation rate for the costs of business cycle. During booms, the increase in the job finding rate is partly offset by the decrease in unemployment. In recessions, the decrease in the job finding rate is amplified by the increase in unemployment. Because of this asymmetry, average unemployment is increased by fluctuations in the job finding rate. Conversely, fluctuations in the job separation rate lower the average unemployment rate[2]. However, we show that the impact of the volatility in the separation rate on the business cycle cost is necessarily one order of magnitude lower than that of the volatility in the job finding rate for the same coefficient of variation. Although Fujita and Ramey (2008) have recently shown that the relative contribution of the separation rate to the unemployment volatility is higher than initially established by Shimer (2004), the impact of job separation fluctuations on the average unemployment rate, and then on the business cycle costs, is intrinsically limited. For the observed process of the job finding rate in the US economy, we show that the asymmetry quantitatively matters and generates sizable employment loss.

However, this reduced form approach suffers from several shortcomings. First, beyond the

[2]During booms, the decrease in the job separation rate is amplified by the increase in employment. In recessions, the increase in the job separation rate is compensated by the decrease in employment.

employment loss, one needs a better founded criterion in order to evaluate the business cycle cost. Secondly, the reduced form analysis supposes that business cycles do not have any effect on the average job finding rate: the counterfactual stabilized job finding rate is assumed to be equal to the average job finding rate. Finally, in the reduced form approach, stabilization policy cannot be explicitly addressed, as there is no operative way of getting rid of the cycles. A structural matching model is then necessary to take into account how the shocks hitting the fundamentals of the economy affect the job finding rate[3]. However, since Shimer (2005), it is well known that the standard matching model fails to generate realistic fluctuations in the job finding rate. The studies aiming at elucidating the Shimer puzzle emphasize different mechanisms and none seems to close the debate[4]. Costain and Reiter (2008) have recently emphasized that no calibration permits the standard matching model to be consistent both with business cycle facts and with the effects of labor market policies. They show that sticky wages can improve the matching model's performance by making the firm's flow of surplus more pro-cyclical, but Pissarides (2007) emphasizes that the observed proportionality between wages in new matches and labor productivity is then not replicated. Despite these remaining quantitative shortcomings, we believe it is interesting to investigate the welfare implication of unemployment volatility in the matching model as it adds a normative dimension to solving the Shimer puzzle.

We choose to study the business cycle cost implied by a matching model with wage rigidity as suggested initially by Hall (1991) in a first response to the puzzle. This framework fits perfectly well with our objectives: it allows us to replicate all the volatility in the job finding

[3]The Lucas' approach can be considered as sharper than ours because it is model-free. But it cannot address the stabilization issue.

[4]Hall (1991) and Shimer (2005) initially proposed a rigid wage approach, which has been reconsidered more recently by Hall and Milgrom (2008). This hypothesis increases the sensitivity of the model to productivity shocks, but at the expense of the observed flexibility in wages (Pissarides (2007)). Hagedorn and Manovskii (2008) show that a higher parametrization for the utility of being unemployed and a lower bargaining power for workers enable the standard matching model to yield realistic fluctuations in the unemployment rate. Whereas Hornstein et al. (2007) introduce investment-specific technological shocks, Kennan (2006) emphasizes the role of procyclical informational rents: the gain that firms obtain by being more informed than workers increases in booms. The inclusion of turnover costs (Pissarides (2007), Mortensen and Nagypal (2003) and Silva and Toledo (2008)) and match-specific technological change (Costain and Reiter (2008)) have also been investigated.

rate[5], but also to reveal, in a very transparent way, the basic non-linearity embedded in the matching model. We show that the matching model naturally predicts that the job finding rate is a concave function of the vacancy-unemployment ratio. This could imply that fluctuations decrease the average job finding rate. But as the vacancy-unemployment ratio is a convex function of productivity, the impact of productivity fluctuations on the average job finding rate is a priori ambiguous. These non-linearities come from the decreasing marginal returns in the matching function. The matching elasticity to vacancy plays a crucial role in the size of the business cycle costs: the lower the elasticity, the lower the average job finding rate and the higher the business cycle costs. The internal mechanisms of the matching model then lead to quite strong additional business cycle costs when the vacancy elasticity of the matching function is sufficiently low.

In order to show the robustness of our results, we propose two very different exercises. We firstly consider the flexible wage approach proposed by Hagedorn and Manovskii (2008). The employment loss is of the same magnitude as in the rigid-wage model, since the unemployment dynamics and the decreasing marginal returns are key features of any matching model. Secondly, using the panel data constructed by Bassanini and Duval (2006) over the period 1982-2003 for 20 OECD countries, we provide evidence that countries in which productivity is more volatile have also a higher level of unemployment.

All in all, the matching model with aggregate productivity shocks may challenge Lucas's controversial view on welfare costs of business cycles. For the US economy, these costs reach 0.55% of permanent consumption for a calibration we consider as empirically relevant. However, this result is obtained under several conditions which deserve to be emphasized. Firstly, it implies to considering an expanded definition of unemployment which induces a higher volatility along the cycle, but also a higher level of structural unemployment. Secondly, the employment loss must be costly enough in terms of consumption: it implies that the surplus from market production is high enough. Typically, the high level of home production implied

[5] Note that we choose to replicate all the volatility in the job finding rate only with productivity shocks, even if it is now well-established that the volatility generated by these shocks only is lower (Pissarides (2007) and Mortensen and Nagypal (2003)).

by the calibration proposed by Hagedorn and Manovskii (2008) for solving the Shimer puzzle leads to less significant welfare costs than the rigid-wage approach.

Finally, we explicitly address the policy issue. Besides stabilizing the job finding rate, our theory suggests dampening the impact of the job finding rate fluctuations on welfare, by subsidizing employment in order to increase the average job finding rate. Note that such subsidies in our model, characterized by a suboptimal level of employment, permit to reach both the stabilization and the resource allocation objectives.

The literature following Lucas (1987) has mostly focused on the consequences of business cycles on the volatility of individual consumption. More precisely, because business cycles amplify individual income risks, they could generate higher welfare losses than Lucas's predictions when financial markets are incomplete. However, in as far as individual income fluctuations are transitory, the costs of the business cycle are still low, even negligible (Krusell and Smith (1999)), mainly because consumption can be smoothed through capital accumulation. On the other hand, when individual income variations are more persistent, the cost of business cycles becomes more substantial, at least 1% percent of the permanent consumption (Beaudry and Pages (1999), Krebs (2007), Storesletten et al. (2001), De Santis (2007), Krusell et al. (2009)). Reis (2009) shows that the persistence also plays a key role when only aggregate shocks are considered. In the literature, very few papers focus on the consequences of business cycles for average consumption. This idea has been sketched out by **?**; they argue that rather than steadying economic activity at its average level, stabilization would prevent output from deviating from its potential level. Ramey and Ramey (1993) explore this mechanism in a model where firms have to pre-commit to a specific technology before starting production. Stabilization may also increase welfare through its effect on capital accumulation (Matheron and Maury (2000), Epaulard and Pommeret (2003) and Barlevy (2003)). Finally, a very recent paper by Jung and Kuester (2009) emphasizes the non-linear relation between unemployment and the job finding rate. They, however, propose a quite complex matching model with capital accumulation, liquidity constraints and human capital, which prevents them from clearly evaluating the basic non-linearities embedded in the matching approach. Moreover, all these additional features bring very small business cycle costs. This shows a clear difference from our paper: we aim at unveiling in the basic matching model the mechanisms which lead to business cycle costs, namely the stock-flow unemployment dynamics and the congestion effects due to decreasing marginal returns in the matching function. Although the implied business cycle costs are weaker than in recent studies (Krebs (2007), Reis (2009) and Barlevy (2003) for instance), the effect of aggregate fluctuations on the employment level adds to the list of factors missing from Lucas's argument, and would indicate more substantial welfare costs from business cycles.

The paper is organized as follows. Section 2 uses a reduced form of the matching model to investigate the consequences of the non-linearity in the unemployment dynamics. Section 3 takes into account the non-linearity in the job finding rate dynamics embedded into the matching model. The fourth section is devoted to robustness tests. The last section concludes.

3.2 Asymmetry in the unemployment dynamics: a reduced form approach

Labor market frictions naturally generate asymmetries in the unemployment dynamics. Because of these asymmetries, aggregate fluctuations may have an impact on average unemployment. In this section, following Cole and Rogerson (1999), our theoretical framework is based on the reduced form of the matching model. The use of a reduced form model has the advantage of emphasizing that these asymmetries are present in any labor market matching model. We first present our theoretical framework, and then analyze the unemployment dynamics. Finally, a preliminary quantitative evaluation of the business cycle costs is proposed.

3.2.1 Framework

We consider unemployment dynamics as the result of exogenous job separation and job finding fluctuations. By shutting down any non linearities[6] that may affect the job finding and separation rates, the reduced form model allows us to focus on the non linearity embedded in the unemployment dynamics.

Unemployment. The unemployment dynamics arise from the entries to and exits from employment. The former are determined by the job finding rate p, the latter by the separation rate s.

$$u_{t+1} = s_t(1 - u_t) + (1 - p_t)u_t \tag{1}$$

Shocks. The economy is hit only by aggregate shocks which generate some fluctuations in the job finding rate p_t and in the separation rate s_t. The job finding rate and separation rate are exogenous with respect to u. This key assumption derives from the matching theory. The aggregate shocks affect linearly both the job finding rate and the separation rate which are assumed to follow an AR(1) process[7]:

$$p_t = (1 - \rho_p)\bar{p} + \rho_p p_{t-1} + \varepsilon_t^p \tag{2}$$

$$s_t = (1 - \rho_s)\bar{s} + \rho_s s_{t-1} + \varepsilon_t^s \tag{3}$$

The shocks ε^p and ε^s have a zero-mean and a standard deviation equal to σ_{ε^p} and σ_{ε^s} respectively. $\bar{p}$ and $\bar{s}$ denote the average job finding rate and the average separation rate respectively.

Business cycle costs. Aggregate shocks may cause average unemployment to differ from the level it would have reached in an economy without any shocks. Following the convention in the literature, we refer to the latter as the stabilized economy. How the stabilized economy is

[6]We will show in the structural approach (Section 3) that the average job finding rate can be increased or decreased by aggregate fluctuations. The job finding rate is a concave function of the vacancy-unemployment ratio, but, as the vacancy-unemployment ratio is a convex function of productivity, the effect of productivity fluctuations on the average job finding rate is a priori ambiguous. The reduced form case with zero effect can then be considered as an instructive benchmark.

[7]We assume symmetrical shocks in order to identify the endogenous asymmetries generated by equation (1). We show in Appendix D that our results are not sensitive to this hypothesis.

reached is not explicitly presented. In particular, nothing is said in this section on the design and the efficiency of stabilization policies.

The cost of fluctuations is the cost of being in an economy hit by aggregate shocks, rather than being in an economy without aggregate shocks. In the former economy, the job finding rate and the separation rate fluctuate around their means, whereas in the latter they are set forever at their average value $\bar{p}$ and $\bar{s}$. The stabilized unemployment (or the structural unemployment) is equal to:

$$\bar{u} = \frac{\bar{s}}{\bar{s} + \bar{p}}$$

The percentage of aggregate employment lost in the business cycle is then given by:

$$\lambda_u = \frac{1 - \bar{u}}{1 - \mathbb{E}(u)} - 1 = \frac{\mathbb{E}(u) - \bar{u}}{1 - \mathbb{E}(u)}$$

with $\mathbb{E}(u)$ the unconditional expectation of unemployment. Traditionally, since Lucas (1987), the business cycle cost is defined as the percentage of the consumption flow that agents would accept sacrificing in order to get rid of aggregate fluctuations. To what extent employment losses are transformed into welfare costs depends on the links between employment, income and consumption. This issue requires a more structural approach and we will show in Section 3 that the employment loss can be considered as a good approximation of the welfare costs in a matching economy. But it is already straightforward that the employment loss λ_u is of the same order of magnitude as the welfare cost in an economy populated by risk-neutral agents without savings.

3.2.2 The analysis of the non-linearities in the unemployment dynamics

By considering equation (1), it is fairly intuitive that shocks on the job finding rate and on the separation rate have non linear effects on unemployment. The impact of these shocks depends on the level of unemployment. During booms, the decline in unemployment offsets the increase in the job finding rate. The small search pool in booms implies that a higher job finding rate increases job creations less than it otherwise would. On the contrary, as unemployment is higher during a recession, the job finding rate shocks have a greater impact in recession; the decline in the job finding rate is magnified by the increase in unemployment. As a result, fluctuations in the job finding rate tend to increase average unemployment.

Conversely, as the impact of the job separation rate shocks depends on the level of employment, the fluctuations in the separation rate lead unemployment to decrease more in booms than to increase in recession: fluctuations in the job separation rate tend to reduce average unemployment.

In this section, we assess precisely these different effects. First, as it is traditionally done in the matching approach with aggregate shocks (see for instance Hall (2005)), a steady state analysis of equation (1) is conducted and so fluctuations in the conditional steady states are considered. This analysis delivers very easily the basics of the non-linearity embedded in the unemployment dynamics. This simple framework highlights the importance of the volatility of the job finding and separation rates for the size of business cycle costs, but also the less expected role played by the structural unemployment rate. Secondly, taking into account the inertia embedded in equation (1), we derive the full non-linear properties of the unemployment dynamics and show that the persistence of the aggregate shocks also matters.

3.2.2.1 Steady state analysis

Let us assume for now that the speed of convergence of unemployment is infinite: fluctuations cause unemployment to jump directly from one conditional steady state to another. A conditional steady state unemployment corresponds to the level $\widetilde{u}_i$ toward which the unemployment rate would converge if the separation and job finding rates forever keep the same value p_i and s_i, i.e. if the economy remains in the same state i. Let us define π_i the unconditional probability of being in state i. The value taken by p and s in each state i and the probability associated π_i define the Markov chains associated with p and s, consistently with equations (2) and (3). The average job finding rate $\bar{p}$ is therefore equal to $\sum_i \pi_i p_i$ and the average separation rate $\bar{s}$ to $\sum_i \pi_i s_i$. Jointly, they determine the structural unemployment rate $\bar{u}$. On the other hand, as unemployment is assumed to jump directly from one conditional steady state to another, the average unemployment in this economy is then equal to the average of the conditional steady states:

$$\tilde{u} = \sum_i \pi_i \widetilde{u}_i$$

The non-linearity embodied in equation (1) implies that average unemployment $\tilde{u}$ has no reason to coincide with structural (stabilized) unemployment $\bar{u}$.

Job finding rate shocks

To understand the specific role of the non linearity in job findings, let us assume first that the separation rate is constant and equal to its mean $\bar{s}$. This non-linearity implies that $\widetilde{u}_i$ is a convex function of the state-dependent job finding rate p_i:

$$\widetilde{u}_i = \frac{\bar{s}}{\bar{s} + p_i} \tag{4}$$

Because unemployment is a convex function of the job finding rate, the average unemployment is higher than the structural (stabilized) unemployment $\bar{u}$. For uniformly small deviations, using a second order Taylor expansion of equation (4), this gap can be written as [8]:

$$\tilde{u} - \bar{u} \approx \bar{u}\ (1 - \bar{u})^2 \left(\frac{\sigma_p}{\bar{p}}\right)^2 \tag{5}$$

A mean-preserving increase in the volatility widens the gap between the stabilized and fluctuating unemployment rates. The more volatile the economy, the greater the business cycle cost. Equation (5) indicates that the unemployment gap also depends on the structural unemployment rate $\bar{u}$. This result is important as it generates strong interactions between structural and cyclical unemployment. More particularly, a higher value of $\bar{u}$ imply a more convex economy and leads to higher business cycle costs[9]. Furthermore, this suggests that labor market

[8] See Appendix A for the derivation.

[9] This holds as long as $\bar{u} \leq 1/3$.

institutions affect the costs of fluctuations as they have an impact on the average job finding and separation rates.

Job separation rate shocks

So far, the separation rate was assumed to be constant. However, it appears clearly from equation (1) that the asymmetry in the unemployment dynamics could also come from fluctuations in the separation rate. The resulting unemployment gap would then read:

$$\widetilde{u} - \bar{u} \approx -\bar{u}^2(1-\bar{u})\left(\frac{\sigma_s}{\bar{s}}\right)^2 \tag{6}$$

with σ_s the unconditional standard deviation of the separation rate. Contrary to the job finding rate case, fluctuations in the separation rate tend to reduce average unemployment. Job separations decrease more in expansion than they increase in recessions. As shown by equation (6), the unemployment gap depends again on the average job finding and separation rates: a higher structural unemployment rate leads to more business cycle gains[10]. Further, the asymmetry embodied in equation (1) translates into average unemployment only if the separation rate is volatile enough.

3.2.2.2 Considering unemployment inertia

All previous calculations have been made with the assumption that unemployment jumps directly to the conditional steady states. Under this assumption, average unemployment in the business cycle economy is equal to the average of the steady states. Because of unemployment inertia, the asymmetry embodied in the conditional steady state does not necessarily manifest in average unemployment. This asymmetry affects the average unemployment rate only if the aggregate shocks are persistent enough.

To see in a synthetic formula the role of the mean, the volatility and the persistence of aggregate shocks, let us solve equation (1). Again, we first compute average unemployment in the case where only the job finding rate is fluctuating. As shown in Appendix B, the

[10]This holds for $\bar{u} \leq 2/3$.

additional unemployment created by business cycles can then be approximated by [11]:

$$\mathbb{E}[u] - \bar{u} \approx (1-\bar{u})^2 \frac{\bar{s}\ \rho_p}{1-\rho_p(1-\bar{s}-\bar{p})} \left(\frac{\sigma_p}{\bar{p}}\right)^2 \tag{7}$$

If p_t was not serially correlated ($\rho_p = 0$), fluctuations in the job finding rate would not affect average unemployment: $\mathbb{E}[u] = \bar{u}$. When aggregate shocks are persistent, average unemployment in the business cycle economy is no longer equal to stabilized unemployment. In line with our intuition, equation (7) also shows some interactions between the job finding rate volatility, its persistence and the structural unemployment rate. An increase in the variance of the shocks raises average unemployment all the more so when the average job finding rate is low and the persistence of the shocks is high.

Symmetrically, in the case where only the separation rate is fluctuating, the gain of fluctuations positively depends on the volatility and on the persistence of the separation rate:

$$\mathbb{E}[u] - \bar{u} \approx -\bar{u}^2 \frac{\bar{p}\ \rho_s}{1-\rho_s(1-\bar{s}-\bar{p})} \left(\frac{\sigma_s}{\bar{s}}\right)^2 \tag{8}$$

Again, the structural unemployment rate interacts with the volatility and the persistence of the shocks. However, as shown by equations (7) and (8), the employment gains are potentially smaller than the employment losses for the same coefficient of variation and the same level of autocorrelation, since the unemployment rate is one order of magnitude lower than the employment rate. Although Fujita and Ramey (2008) have recently established that the relative contribution of the separation rate to the unemployment volatility is higher than initially shown by Shimer (2004), the impact of job separation fluctuations on the average unemployment rate, and then on the business cycle costs, remains intrinsically limited[12].

[11] The approximation consists of neglecting moments of order above 2. This is line with our approach : as we want to understand how symmetrical shocks can yield non symmetrical effects on unemployment, we disregard in particular the consequences of non-zero skewness

[12] The variance decomposition proposed by Fujita and Ramey is based on the linearization of steady state unemployment: $\sigma_u^2 \approx (1-\bar{u})^2\bar{u}^2\left[\left(\frac{\sigma_p}{\bar{p}}\right)^2 + \left(\frac{\sigma_s}{\bar{s}}\right)^2 - 2\rho_{s,p}\frac{\sigma_s}{\bar{s}}\frac{\sigma_p}{\bar{p}}\right]$. If the separation rate and the job finding rate were characterized by the same coefficient of variation, the variance decomposition indicates that the separation rate would contribute to half of the volatility in the unemployment rate, whereas the impact of

This is why we disregard job separation fluctuations in the sequel.

3.2.3 Quantifying the employment loss

To investigate whether the observed fluctuations in the job finding affect average unemployment and hence the costs of business cycles, it is necessary to estimate the AR(1) process[13] described by equation (2). As equation (7) gives only an approximation of average unemployment, we resort to simulations to obtain a more accurate estimate of the costs of business cycles. Consistently with the AR(1) estimation, we simulate job finding rate shocks in order to obtain artificial series. We then use them to simulate equation (1), which allows us to compute the average unemployment rate in the business cycle economy and the business cycle costs[14].

3.2.3.1 Data

The behavior of the job finding rate over the business cycle is still a debated subject. It especially depends on the underlying conception of “unemployment". Contrary to Shimer (2005), Hall (2005) uses a measure of unemployment expanded to include “discouraged workers" and “marginally attached workers". Although those workers are classified as being out of the labor force, their behavior is close to that of workers classified as unemployed. Blanchard and Diamond (1990) advocate that the hazard rate for workers who are not in the labor force but “want a job" is similar to that of the unemployed workers. Their transition rate to employment is significantly higher than that of the other workers out of the labor force. For Canada, the transition rate to employment is 12.4% for marginally attached workers vs 3.5% for other non-participants and 23% for unemployed workers (Jones and Riddell, 2006). Although our benchmark result is based on Hall's data, we also provide the results obtained when Shimer's

fluctuations in the separation rate on the average unemployment rate would be much weaker than that of the job finding rate. In Hairault et al. (2008), we checked that the business cycle gains implied by the observed volatility of the job separation rate is not significantly different from zero.

[13] It must be emphasized that considering a log-normal distribution for p would have led to very similar results. See Appendix D for more details.

[14] Note that we also need mean separation rate for simulating equation (1). It is estimated at 0.031 in Hall's case and at 0.035 in Shimer's case.

approach is used[15], in order to emphasize the sensitivity of the employment loss to this choice. The Hall (2005) and Shimer (2005) measures of the job finding rate both exhibit pro-cyclicality. The job finding rate plunges at each recession and recovers at each expansion (Figures A.1 and A.2, Appendix C). Both measures show a downward trend in the 1970s and in the early 1980s. As some of these movements could be due to factors unrelated to business cycles, we focus hereafter on series detrended by a low frequency filter.

Table 1: Job finding rate statistics

	Hall data	Shimer data
Mean $\bar{p}$	0.285	0.607
Standard deviation σ_p	0.069	0.100
Coefficient of variation $\frac{\sigma_p}{\bar{p}}$	0.242	0.165
Autocorrelation	0.913	0.903

Note: Quarterly average of monthly data. Sample covers 1948q3-2004q3 for Hall (2005) and 1951q1-2003q4 for Shimer (2005). Following Shimer (2005), both sets of data are detrended with a HP smoothing parameter of 10^5.

We then estimate the parameters characterizing the process of the job finding rate as described by equation (2). As expected, the “expanded job finding rate" has a lower mean than Shimer (2005)'s measure (Table 1). More importantly, including low-intensity job seekers also implies a higher coefficient of variation.

3.2.3.2 Simulation results

Table 2: Average unemployment and job finding rate fluctuations

	Hall data	Shimer data
Unemployment	10.22%	5.50%
Stabilized unemployment	9.83%	5.38%
Cost of fluctuations λ_u	0.44%	0.13%

[15] Note that for Shimer's series, we use the job finding and separation rates and not the job finding and separation probabilities as the latter are not consistent with the stock flow equation for unemployment (equation (1)): Shimer (2005) takes into account the probability of finding another job within the month.

Table 2 presents the employment loss due to business cycles in the US economy for the two measures of the job finding rate. Non-linearities in the unemployment dynamics are enough to generate sizable costs of business cycles. In particular, these costs are between one and two orders of magnitude greater than the costs found by Lucas (1987)[16]. The method chosen to measure the job finding rate has strong consequences on the costs of fluctuations. When some non-employed job seekers (Hall's method) are taken into account, fluctuations in the job finding rate induce a 0.44% employment loss. Shimer's measure leads to lower business cycle costs: if the job finding rate is computed using only transitions from unemployment, the cost of business cycles reduces to 0.13%. Such a result was expected, as Shimer's job finding rate series display both a lower volatility and a higher mean (Table 1), two characteristics that we identified as cost-reducing.

3.3 Endogenizing the job finding rate: a structural approach

The previous section showed that the observed volatility and persistence in the job finding rate lead to sizable costs of business cycles. The structural matching model is a candidate for generating such costs. However, Shimer (2005) shows that the canonical matching model fails to generate realistic fluctuations in the job finding rate. The standard deviation of the job finding rate is much greater in the data than in the model (Shimer (2005)). An increase in labor productivity increases the expected profit from a filled job, and thus firms tend to open more vacancies. But there are internal forces in this framework which partially offset the initial increase in expected profits and then dampen the incentives for vacancy creations. There already exist in the literature different approaches which solve the Shimer puzzle[17]. Do we care about identifying the mechanism at the origin of the high fluctuations in the job finding rate? From the analysis conducted in the first part, it could be tempting to say that it

[16] At this stage, we agree that the two approaches are not directly comparable. We will see in the next section that the employment loss is not so far from a cost expressed in terms of consumption.

[17] See for instance Hall (1991), Hall and Milgrom (2008), Pissarides (2007), Hagedorn and Manovskii (2008), Hornstein et al. (2007), Kennan (2006), Mortensen and Nagypal (2003), Silva and Toledo (2008) and Costain and Reiter (2008).

is enough to know that at least one theory is able to replicate the job finding rate dynamics. Actually, we do care. Indeed, the results obtained in the first part are derived from a model in which the job finding rate is exogenous, and in which fluctuations are neutral regarding the average job finding rate. To assess the costs of fluctuations, one must take into account the consequences of stabilization on the average job finding rate. If productivity shocks and the job finding rate are linearly related, the results found in the previous section should *a priori* be close to the endogenous job finding rate case. But if not, the average job finding rate can then be affected by business cycles. Why do we suspect the presence of a non-linear effect of business cycles on the job finding rate? The job finding rate is a non-linear function of the labor market tightness which also depends non-linearly on productivity changes. To show and quantify these different effects, a structural model is then required and the costs of business cycles could differ according to the model specification. This last statement is all the more true as the cost of business cycles will rely on a welfare criterion consistent with the structural model. In particular, more attention must be paid to the cyclical behavior of vacancies.

The choice of the theoretical model is then potentially crucial. We choose to study the business cycle cost implied by a canonical matching model with wage rigidity (Hall (1991)). As in the previous section, we leave aside the fluctuations in the separation rate so as to keep the analysis as tractable as possible[18]. This approach allows us to focus on the implications of the basic non-linearity introduced by the matching function. We then present different calibrations of the matching function elasticity in order to unveil these implications. Each replicates the job finding rate process (standard deviation and mean)[19]. The implied business cycle costs are not necessarily identical and equal to that obtained in the reduced-form part. In a robustness analysis, considering a flexible wage framework (Hagedorn and Manovskii (2008)), we show that these mechanisms are intrinsic to the matching model, and then not specific to the rigid wage approach.

[18]Moreover, it is well-known that job creation conditions are unaffected when job separations are endogenous (Pissarides (2007)).

[19]The persistence would be naturally matched by that of productivity shocks.

3.3.1 A canonical matching model

The model considered hereafter is a version of the matching model *à la* Pissarides with aggregate uncertainty and exogenous separation.

3.3.1.1 Matching technology

Output per unit of labor is denoted by y_t and is assumed to follow a first-order Markov process according to some distribution $G(y, y') = Pr(y_{t+1} \leq y'|y_t = y)$. To hire workers, firms must open vacancies at unit cost κ. Jobs and workers meet pairwise at a Poisson rate $M(u, v)$, where $M(u, v)$ stands for the flows of matches and v the number of vacancies. This function is assumed to be strictly increasing and concave, exhibiting constant returns to scale, and satisfying $M(0, v) = M(u, 0) = 0$. Under these assumptions, unemployed workers find a job with a probability $p(\theta) = M(u, v)/u$ that depends on the ratio of vacancies to unemployment ($\theta = v/u$). The probability of filling a vacancy is given by $q(\theta) = M(u, v)/v$. Hereafter, we impose that the matching function is Cobb-Douglas: $M(u, v) = \varphi u^{1-\alpha} v^{\alpha}$ with $0 < \alpha < 1$. The unemployment dynamics in the economy (equation (1)) are similar to equation (1), except that the job finding rate is now endogenous. Equations (2) and (3) define the job finding rate and the job filling rate respectively which depends[20] on the current productivity state y:

$$u' = s(1-u) + (1 - p(\theta_y))u \tag{1}$$

$$p(\theta_y) = \varphi \theta_y^{\alpha} \tag{2}$$

$$q(\theta_y) = \varphi \theta_y^{\alpha-1} \tag{3}$$

3.3.1.2 Workers

Workers are risk neutral. They have no access to financial markets. This simplifying assumption is not restrictive as our results do not rely on the impossibility of individuals smoothing their income. As agents are risk-neutral, the excessive volatility of consumption implied by

[20] Throughout the paper the notation x_y indicates that a variable x is a function of the aggregate productivity level y and $\mathbb{E}_y$ is the expected value conditional on the current state y.

this assumption is not captured in the welfare calculations. The focus is here on the impact of business cycles on average consumption, which is not affected by smoothing behaviors. Workers can either be employed or unemployed. Employed workers receive wage w until their job is destroyed (at rate s); we do not take into account on-the-job search and voluntary quits. An unemployed worker gets an unemployment benefit[21] z which is equally financed by workers through lump-sum taxes. We choose to consider that the disutilities of working and of not working are both equal to the same value χ. It is then straightforward to derive the representative agent intertemporal preferences:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (c_t - \chi)$$

where $\mathbb{E}_0$ denotes the expectation operator conditional on information at time 0 and β the discount factor.

3.3.1.3 Aggregate consumption and the welfare cost of business cycles

In our economy, aggregate consumption is equal to the aggregate production net of the vacancy costs.

$$c_t = y_t(1 - u_t) - \kappa\ v_t \tag{4}$$

In this economy, the only sources of fluctuations are the labor productivity shocks. The welfare cost of fluctuations is therefore defined relatively to a counterfactual economy, in which labor productivity remains at its average value. The welfare cost of business cycles λ is defined as the percentage of the consumption flow that the agent would accept sacrificing in order to get rid of aggregate fluctuations:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(1+\lambda)c_t\right] = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \bar{c}_t$$

[21] At this stage, the non market value does not include home production. See the robustness analysis for a discussion of this assumption.

where $\bar{c}_t$ is the level of consumption in the economy without aggregate shocks. It needs to be derived from a counterfactual experiment based on an artificial economy without any shocks. The computation of welfare costs takes explicit account of the transition path to the stabilized economy. However, in order to highlight the welfare cost of business cycles in the matching economy, let us consider the following expression, which leaves aside the transition path[22]:

$$\lambda \approx \frac{\bar{c} - \mathbb{E}(c)}{\mathbb{E}(c)} = \frac{[(1-\bar{u})\bar{y} - \kappa\bar{v}] - \mathbb{E}[(1-u)y - \kappa v]}{\mathbb{E}(c)} \tag{5}$$

Business cycles are costly when they make the production net of the vacancy cost lower than its stabilized level. In order to make a link with the reduced form analysis, the welfare cost of business cycles can be written as follows:

$$\lambda \approx \frac{(\mathbb{E}(u) - \bar{u}) - cov(y, 1-u) + \kappa(\mathbb{E}(v) - \bar{v})}{\mathbb{E}(c)} \tag{6}$$

The first part of the welfare cost of business cycles is the employment loss $(E(u) - \bar{u})$ present in the reduced-form analysis. Let us emphasize that the size of the employment loss is not necessarily of the same magnitude in the structural model, as the job finding rate is now endogenous. The lower the average job finding rate, the higher the employment loss and the higher the welfare cost. The second part $cov(y, 1-u)$ comes from the interaction between productivity and employment. Just as making the job finding rate negatively correlated with the stock of job searchers lowers mean employment, making productivity positively correlated with the stock of workers raises mean output. This explains why a positive covariance leads to decrease the business cycle costs. Finally, the third part shows that more vacancies in the fluctuating economy than in the stabilized one generate higher costs.

3.3.1.4 The value functions

The worker's utility

[22] The transition path does not significantly matter in our economy: it decreases the business cycle cost only very slightly.

Define U_y and W_y to be the state contingent present value of an unemployed worker and an employed worker[23]:

$$U_y = z + \beta\{(1 - p(\theta_y))\mathbb{E}_y[U_{y'}] + p(\theta_y)\mathbb{E}_y[W_{y'}]\} \tag{7}$$

$$W_y = w_y + \beta\{(1 - s)\mathbb{E}_y[W_{y'}] + s\mathbb{E}_y[U_{y'}]\} \tag{8}$$

The firm's surplus

The firm's value of an unfilled vacancy V_y is given by:

$$V_y = -\kappa + \beta\{q(\theta_y)\mathbb{E}_y[J_{y'}] + (1 - q(\theta_y))\mathbb{E}_y[V_{y'}]\} \tag{9}$$

with J_y the state contingent present value of a filled job and $q(\theta_y)$ the probability of filling a vacancy conditionally on the productivity state y. When the job is filled, the firms operate with a constant return to scale technology with labor as only input. The firm's value of a job is given by:

$$J_y = y - w_y + \beta\{(1 - s)\mathbb{E}_y[J_{y'}] + s\mathbb{E}_y[V_{y'}]\} \tag{10}$$

Free entry implies $V_y = 0$ for all y. Therefore, the job creation condition is:

$$\kappa = \beta q(\theta_y)\mathbb{E}_y[J_{y'}] \tag{11}$$

3.3.1.5 Equilibrium

The labor market equilibrium depends on the way the wage is determined in the economy. Though our benchmark is the rigid wage model, we first present the traditional equilibrium with a Nash-bargained flexible wage. It allows us to compare the non-linearities embedded in

[23]For the sake of simplicity, we omit from these equations the disutility of working and not working, the lump-sum tax financing the unemployment benefits and the dividend paid by firms to workers, as these variables are assumed to be identical across individuals.

these two equilibria.

Equilibrium with flexible wages.

When a worker and an employer meet, the expected surplus from trade is shared according to the Nash bargaining solution. The joint surplus S_y is defined by $S_y = W_y + J_y - U_y$. The worker gets a fraction γ of the surplus, with γ her bargaining power. The equilibrium with flexible wages is defined by the job creation condition and the wage rule (equations (12) and (13)), plus equations (1) to (3):

$$\frac{\kappa}{q(\theta_y)} = \beta \mathbb{E}_y \left[y' - w(\theta_{y'}) + (1-s)\frac{\kappa}{q(\theta_{y'})} \right] \tag{12}$$

$$w(\theta_y) = \gamma(y + \kappa\theta_y) + (1-\gamma)z \tag{13}$$

As Shimer (2005) points out, the adjustment of wages is responsible for the insensitivity of the labor market tightness to the productivity shocks. It also makes the interplay of the non-linearities in the model more complex relative to the rigid wage equilibrium, due to the retroaction of wages in the job creation condition (equation (12)).

Equilibrium with rigid wages

Incorporating wage rigidity in the matching model is a natural way to generate enough volatility. Moreover, this allows us to focus on the basic non-linearities introduced by the matching function, present in any matching model.

Following Hall (1991), we consider a constant wage $w_y = w$, $\forall y$. This constant wage is an equilibrium solution if $z \leq w \leq \min \pi_y$, where π_y denotes the annuity value of the expected profit[24]. The wage is set at the Nash bargaining solution relative to the average state of productivity $\bar{y}$. This wage is an equilibrium wage provided it lies in the bargaining set defined by the participation constraints of the firms and the workers.

The rigid wage equilibrium is then defined by substituting equations (12) and (13) by equations

[24]This annuity value is simply computed using the value an employer attaches to a new hired worker who never receives any wage:

$$\widetilde{J}_y = y + \beta(1-s)\mathbb{E}_y[\widetilde{J}_{y'}]$$

The annuity value is then given by $\pi_y = [1 - \beta(1-s)]\widetilde{J}_y$.

(14) and (15), again in addition to the conditions (1) to (3):

$$\frac{\kappa}{q(\theta_y)} = \beta \mathbb{E}_y \left[y' - \bar{w} + (1-s)\frac{\kappa}{q(\theta_{y'})} \right] \quad (14)$$

$$\bar{w} = \gamma(\bar{y} + \kappa\theta_{\bar{y}}) + (1-\gamma)z \quad (15)$$

3.3.2 Non-linearities, welfare cost and employment loss: the rigid wage case

When wages are rigid, the employment loss can be considered as a good approximation of the welfare cost. Indeed, as the production net of vacancy costs is approximately equal to labor earnings (for a discount factor β sufficiently close to 1)[25], it is straightforward to show that:

$$\lambda \approx \frac{\bar{w}(1-\bar{u}) - \mathbb{E}(\bar{w}(1-u))}{\mathbb{E}(\bar{w}(1-u))} = \frac{\mathbb{E}(u) - \bar{u}}{1 - \mathbb{E}(u)} \quad (\beta \to 1) \quad (16)$$

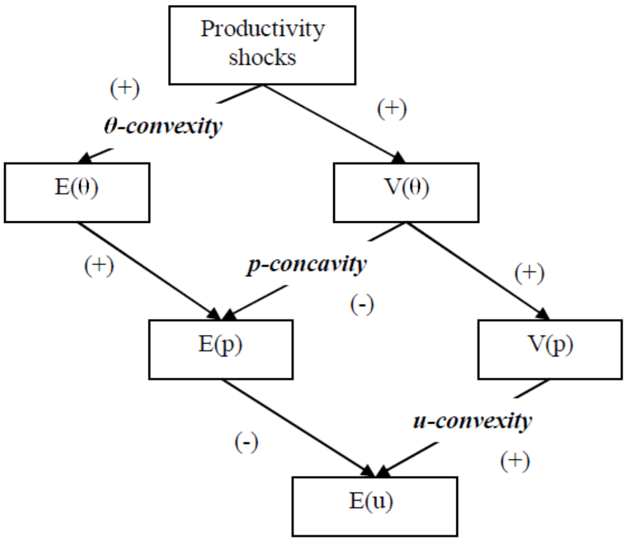
Comparing with equation (6), the cost-decreasing effect of the covariance between productivity and employment and the cost-increasing effect of higher vacancies compensate each other for a discount factor β sufficiently close to 1. More generally, the wage rigidity in the business cycles makes the welfare cost very close to the employment loss.

The total impact of productivity fluctuations on the average unemployment rate is then key to understand the business cycle costs. Figure 1 presents the mapping between labor productivity and unemployment. There is a first source of convexity arising from the unemployment dynamics (u-convexity effect), equation (1). It has been intensively investigated in the previous section. Let us concentrate here on the additional non-linearities that appear once p is endogenous. They determine altogether the average job finding rate and their global impact is not determined a priori.

As the return of an additional vacancy on the job finding rate is decreasing (equation (2)), the average job finding rate in the stabilized economy is potentially higher than in a economy with business cycles (p-concavity effect). The lower the elasticity α of the matching function

[25] In the general case, it is necessary to also take into account the dividends paid by the firms to the workers.

Figure 1: Non-linearities in the structural matching model



to vacancies, the higher the p-concavity effect. Let us emphasize that this effect tends unambiguously to create higher welfare costs as the job finding rate is lowered by the volatility of the labor market tightness, and not by its level[26]. Average unemployment increases without any gains in terms of vacancy reductions.

On the other hand, equation (14), combined with the job filling rate condition (equation (3)), implies that the average labor market tightness is increased by productivity fluctuations. As the matching function exhibits decreasing marginal returns to vacancies, the average duration of a vacancy and hence the cost of opening a vacancy are concave in the labor market tightness. Therefore, the free entry condition is satisfied for greater variations in job creation in booms than in recessions: the labor market tightness is a convex function of productivity (θ-convexity effect). This effect tends to increase the average job finding rate in the fluctuating economy .

[26]In this latter case, a lower labor market tightness could be welfare-improving, if the deterministic equilibrium level of employment was above its optimal value (too many vacancy costs)

From equation (3), it can be seen that this convexity is amplified by a high elasticity α of the matching function to vacancies. *Ceteris paribus*, it leads to lower business cycle costs only in the case where the deterministic equilibrium level of employment is below its optimal value. This is the case in our economy, whatever the parameters we consider, as we impose relatively high unemployment benefits.[27]

Overall, the way the productivity shocks affect the job finding rate on average depends on the degree of decreasing returns in the matching function captured by the elasticity α. To make this point explicit, let us approximate the unemployment rate using the comparative statics of the model without aggregate shocks[28]. It can be shown that:

$$u \approx \overline{u} - \frac{\overline{u}^2}{s} p'(\overline{\theta})\theta'(\overline{y})(y-\overline{y}) + \frac{\overline{u}^2}{s}\left(\frac{\overline{u}}{s}\left[p'(\overline{\theta})\theta'(\overline{y})\right]^2 - \frac{1}{2}\left[p''(\overline{\theta})\theta'(\overline{y}) + p'(\overline{\theta})\theta''(\overline{y})\right]\right)(y-\overline{y})^2 \quad (17)$$

Taking into account that $p = p(\theta(y))$, the average unemployment can then be expressed as follows[29]:

$$\mathbb{E}(u) - \bar{u} \approx \underbrace{\bar{u}\ (1-\bar{u})^2 \left(\frac{\sigma_p}{\bar{p}}\right)^2}_{u\text{-convexity}} - \underbrace{\frac{1}{2}\frac{\bar{u}}{s}\left(\underbrace{p''(\bar{\theta})(\theta'(\bar{y}))^2}_{p\text{-concavity}} + \underbrace{p'(\bar{\theta})\theta''(\bar{y})}_{\theta\text{-convexity}}\right)\sigma_y^2}_{E(p)-\bar{p}} \quad (18)$$

The first part of the welfare cost of business cycle captures the impact of the job finding rate volatility on average unemployment (u-convexity). The second part comes from the impact of the productivity volatility on the average job finding rate. It combines both the p-concavity and the θ-convexity and thus depends on the elasticity α of the matching function to vacancies. This can be shown more formally as follows[30]:

$$p''(\theta)(\theta'(y))^2 + p'(\theta)\theta''(y) = \Gamma(\theta)(2\alpha - 1) \qquad \Gamma(\theta) > 0$$

[27] Note that this is a conservative choice with regard to the size of the business cycle costs.

[28] The comparative static of the model without aggregate shocks can be used to approximate the dynamic stochastic model if the shocks are persistent enough. See Mortensen and Nagypal (2003) for more details.

[29] As $p = \bar{p} + p'(\theta)\theta'(y)(y-\bar{y}) + \frac{1}{2}(p''(\bar{\theta})\theta'(\bar{y}) + p'(\bar{\theta})\theta''(\bar{y}))(y-\bar{y})^2$.

[30] $\Gamma(\theta) = \frac{\varphi\alpha}{(1-\alpha)^2}\left(\frac{\beta}{\kappa(1-\beta(1-s))}\right)^2 \theta^{3\alpha-2}$.

The stabilization of labor productivity either decreases or increases the average job finding rate, depending on the value of the elasticity of the matching function α. The job finding rate is a concave (convex) function of labor productivity if α is below (above) 1/2 and the stabilized job finding rate is then higher (lower) relatively to that of the volatile economy. In the case $\alpha = 1/2$, the θ-convexity and the p-concavity effects exactly compensate each other. In this case, the productivity fluctuations will lead to the same increase in average unemployment as in the reduced form analysis.

These basic non-linearities are not specific to the rigid wage model. It is obvious that the flexible wage equilibrium shares the same fundamental non-linearities, as they stem from the intrinsic characteristics of the unemployment dynamics, of the job finding rate and of the job filling rate which are exactly the same in the two equilibria. However, in the flexible wage case, the θ-convexity effect is modified by the retroaction of the wage in the job creation condition (equations (12) and (13)). Hence, the non-linearity between y and θ_y also depends on the assumption about the wage bargaining process, and more generally on the particular assumptions considered, such as the existence or not of fixed hiring and separation costs. In this sense, the flexible wage equilibrium may introduce non-linearities which are not intrinsic to the matching process. The use of the rigid wage framework allows us to focus on the basic non-linearities which are common to a large class of matching models.

3.3.3 Quantifying the business cycles costs

We first calibrate the rigid wage economy. As equations (16) and (18) give only an approximation of the welfare cost of business cycles, we resort to simulations to obtain an accurate estimate.

3.3.3.1 Calibration

The model is calibrated to match US data. We calibrate the productivity process to match the US labor productivity standard deviation and persistence[31]. The monthly discount rate is set to 0.42%. The job separation rate is set at Hall's estimate for the US economy, 0.031. We choose the elasticity of the matching function α to be 0.5, in Petrongolo and Pissarides (2001) range, in order to start with a structural model as close as possible to the reduced form analysis. Following Shimer (2005), the unemployment benefits are set at 0.4. The scale of the matching function φ is chosen to pin down the US average vacancy-unemployment ratio. The workers' bargaining power γ and vacancy costs κ are then calibrated to reproduce the volatility and the mean of the job finding rate over the cycle[32]. These two targets are computed using Hall (2005)'s measure of the job finding rate[33] (Table 1).

Table 3: Benchmark calibration of the matching model

Parameter	Calibration	Target
Labour productivity		
Average	1	Normalization
Persistence ρ_y	0.90	US data (1951-2003)
Standard deviation σ_y	0.9%	US data (1951-2003)
Discount rate r	0.0042	Corresponds to 5% annually
Job destruction rate	0.031	Hall (2005)
Elasticity of the matching function α	0.5	Petrongolo-Pissarides (2001)
Unemployment benefits z	0.4	Shimer (2005)
Scale of the matching function φ	0.346	Matches US average v-u ratio of 0.72 (Pissarides, 2007)
Cost of vacancy κ	0.240	Matches US average job finding rate of 0.285 and job finding rate volatility of 0.068
Workers' bargaining power γ	0.760	

[31] We use the same data as Shimer (2005), the real output per worker in the non farm business sector, detrended with a HP smoothing parameter of 10^5.

[32] As the model assumes a constant separation rate, we do not aim at replicating all the unemployment volatility. Otherwise, we would overestimate the business cycle costs by imposing too much volatility in the job finding rate in order to replicate all the unemployment volatility.

[33] As noted before, Hall (2005) uses a measure of unemployment expanded to include "discouraged workers" and "marginally attached workers". As workers are considered as homogenous in our matching model, we then assume that they generate the same surplus as the workers who are officially unemployed, which is debatable. If the surplus was inferior, the business cycle costs would be lowered.

3.3.4 Simulation

Table 4: The welfare cost of fluctuations in a matching model with rigid wages

	Business cycles economy				Stabilized economy			Cost of
	$E(u_t)$	$E(p_t)$	σ_p	$E(\theta)$	$\bar{u}$	$\bar{p}$	$\bar{\theta}$	fluctuations
$\alpha = 0.5$	10.22%	0.285	0.069	0.72	9.82%	0.285	0.681	0.44%
$\alpha = 0.4$	10.22%	0.285	0.069	0.72	9.75%	0.288	0.678	0.55%
$\alpha = 0.6$	10.22%	0.285	0.069	0.72	9.89%	0.283	0.685	0.32%

Table 4, Line 1, presents the results for the benchmark calibration of the rigid wage model. These results show that the average unemployment rate is higher in the fluctuating economy. This is also the case for the average labor market tightness, and so for the average vacancies as well. As expected, there is no influence of productivity fluctuations on the average job finding rate as $\alpha = 0.5$, and the welfare cost of business cycles is of the same magnitude as in the reduced form analysis[34]. The size of the business cycle cost is only determined by the u-convexity effect.

We then simulate two other cases: $\alpha = 0.4$ and $\alpha = 0.6$ (last two lines of Table 4). The values of the parameters κ,γ and φ have been changed accordingly in order to still match the job finding rate characteristics[35]. Depending on α, the US welfare cost of fluctuations could reach 0.55% or reduce to 0.32%. Petrongolo and Pissarides (2001) estimate this elasticity to be between 0.3 and 0.5. This suggests that with $\alpha = 0.5$, our benchmark calibration gives a lower bound of the welfare costs of fluctuations. In the more realistic case ($\alpha = 0.4$), the average job finding rate in the fluctuating economy is lower than the value which would be reached in the stabilized economy. A lower elasticity strengthens the p-concavity effect and dampens the θ-convexity effect. In this case, the internal mechanism of the matching model

[34] Note that the cost-decreasing effect of the covariance between productivity and employment, offset by the cost-increasing effect of higher vacancies, generates a small change in the business cycle costs relative to the employment loss (less than 0.06 of a percentage point compared to 0.44).

[35] Note that, in the $\alpha = 0.4$ case, the rigid wage defined at the median productivity is no longer in the bargaining set. The wage is then fixed at its highest value ensuring that the firm's value is still positive ($w = \text{argmin}_y \pi_y$).

leads to quite high business cycle costs, since these costs are increased by one fourth. Note that it occurs only when the labor market tightness θ (and so the job finding rate $p(\theta)$) is volatile enough to make the non-linearity operate. Replicating the volatility of both the labor market tightness and the job finding rate leads to sizable business cycle costs through different channels which are all at work in this structural model.

We believe that these results are quite general and do not depend on the rigid wage framework we consider here[36]. Firstly, the business cycle costs are mainly generated by the u-convexity, which is a key feature of any matching model. Secondly, the non-linearity of the job finding rate due to the matching function is not specific to the rigid wage economy. The flexible wage case actually adds other sources of non-linearity which depend on the specification of the matching model considered. The next section explicitly addresses this robustness issue.

3.4 Robustness

In this section, we evaluate the robustness of our results by undertaking two very different strategies. Firstly, we consider the alternative view to the rigid-wage approach proposed by Hagedorn and Manovskii (2008). Secondly, we propose to test the presence of a convex relationship between productivity shocks and unemployment through a cross-country panel analysis.

3.4.1 A flexible-wage approach

Hagedorn and Manovskii (2008) have recently proposed an alternative response to the Shimer puzzle by suggesting that the problem is more in the way the model is calibrated than in the model itself. The matching framework with flexible wages can generate a realistic volatility in the job finding rate when the value of unemployment and of the workers' bargaining power

[36]Our result suggests another interpretation to the cost of the wage rigidity given by Shimer (2004). This cost is of the same order of magnitude as our measure. He interprets this cost as the result of the non-optimality of the labor market tightness volatility. Our result rather suggests that it comes from the employment loss generated by the volatility of the job finding rate. Note that Shimer (2004) indeed emphasizes that the average output in the rigid wage economy is 0.12 percent below the level in the flexible wage one.

are judiciously calibrated. Considering this alternative view enables us to qualify our results. In this flexible wage framework, the welfare cost of business cycles may differ from the results derived under rigid wages for several reasons. Firstly, the non linearity in the job finding rate is modified due to the retroaction of wages in the job creation condition. Appendix E indeed shows that the condition which ensures that the job finding rate is lowered by business cycles is less stringent in this flexible wage environment. For the same value of α, the flexible wage framework potentially leads to higher business cycle costs. Secondly, the employment loss is no longer necessarily a good approximation of the welfare cost of business cycles. The covariance between productivity and employment and the behavior of vacancies over the cycle matter for the size of the market production loss (equation (6)). Further, the calibration of Hagedorn and Manovskii (2008) adds another difference: they show that the volatility of the job finding rate is high enough only if the value of the non-market activity z is calibrated sufficiently close to the average productivity. This implies calibrating z at a much higher value than its strict interpretation as an unemployment benefit would imply. Following Hagedorn and Manovskii (2008), we consider that z now includes home production l. It leads to a redefinition of aggregate consumption as follows: $c = y(1-u) + ul - \kappa v$. It implies that the business cycle welfare cost is different from the market production loss defined in equation (6) and is then sensitive to the value of l:

$$\lambda_l \approx \frac{(\mathbb{E}(u) - \bar{u})(1-l) - cov(y, 1-u) + \kappa(\mathbb{E}(v) - \bar{v})}{\mathbb{E}(c)} \tag{1}$$

Let us emphasize that, for a given level of z, the value of l matters only for the magnitude of the welfare costs and has no effect on average employment and vacancy rates.

In order to check the robustness of our results, we now calibrate the flexible wage equilibrium, defined by the equations (1), (3) (12) and (13), along the lines of Hagedorn and Manovskii (2008). We consider the same values as in our benchmark economy for all parameters, except for the workers' bargaining power γ, the non-market value z and the vacancy cost κ which are calibrated in order to match the same targets as in the rigid wage case for internal consistency[37].

Table 5: Employment and market production loss with flexible wages

Business cycles economy				Stabilized economy			Employment	Market production
$E(u_t)$	$E(p_t)$	σ_p	$E(\theta)$	$\bar{u}$	$\bar{p}$	$\bar{\theta}$	loss	loss
10.22%	0.285	0.069	0.72	9.74%	0.288	0.697	0.52%	0.49%

Table 5 shows that the effects of the non-linearities present in the matching model appear to be quite robust to variations in the degree of wage rigidity. The employment loss is even larger than in the rigid wage approach (0.52% against 0.44% in Table 4), as the stabilized job finding rate is now higher than its cyclical counterpart for α equal to 0.5. For the calibration considered in this flexible wage framework, the threshold for α under which the job finding rate is higher in the stabilized economy is 0.63 (vs. 0.5 in the rigid wage economy)[38]. However, as wages are now flexible, this employment loss does not necessarily coincide with the market production loss. Actually, the last column in Table 5 shows that these two losses are very close. The interaction between productivity and employment is nearly compensated by the the vacancy gain. This result is not surprising: as the calibration is such that the bargained wage is close to the constant outside opportunity of the employed workers, this flexible wage model "is a close cousin of others that rationalize wage rigidity by dropping Nash wage bargaining" (Hall (2006), p.16).

However, the extent to which the employment (or the market production) loss can be interpreted as a welfare cost crucially depends on the home production value l. The standard calibration of the unemployment benefits for the US economy is between 0.3 and 0.6 (Kitao et al. (2008), Nickell et al. (2005) and Shimer (2005)). For the calibrated value of 0.97 for z, this implies that l is between 0.37 and 0.67, leading to a welfare cost equal to 0.30% and 0.14% respectively. Obviously, a high value of l automatically lowers the costs of business

[37]Note that the targets are different in Hagedorn and Manovskii (2008). The values of κ, γ and z are respectively set at 0.21, 0.033 and 0.97.

[38]See Appendix E.

cycles. This question of the value of non-market activities is undoubtedly key for the welfare costs of business cycles.

Let us emphasize that our results on the employment loss are robust to considering flexible wages. We believe they would also be robust to other matching models solving the Shimer puzzle as the underlying mechanisms are intrinsic to this class of models[39]. On the other hand, our results on the welfare costs are sensitive to the home production value, but this value remains so highly debated among the profession that we consider that this is more a matter of opinion than a matter of fact. Furthermore, when non-market returns are high, the response of unemployment to unemployment insurance is too large (Costain and Reiter (2008)). Hall and Milgrom (2008) also noted that the Hagedorn and Manovskii calibrations imply too high a labor supply elasticity, given empirical estimates.

Moreover, it must be emphasized that we have neglected other dimensions that could have increased our quantitative measure of the business cycle welfare costs. Once the non-market production is considered to be as efficient as the market production, the disutility of “not working" may dominate the disutility of working when job search costs, as well as indirect costs such as psychological damage and skill obsolescence, are taken into account. There are no clear empirical answers on this crucial point. On the other hand, unemployment benefits do not lead to distortive taxation in our theoretical framework. Business cycles, by increasing average unemployment, could imply higher taxes, which would, in turn, weigh employment down. This could have been counted as a cost of business cycles. Another dimension which could magnify these costs is the loss of human capital generated by unemployment spells and reflected in the permanent decrease in wages observed in data (Krebs (2007), Jung and Kuester (2009)). Is this compensated for by more intense human capital accumulation during expansions? All these points would deserve to be addressed to obtain a more general assessment of the welfare cost of unemployment fluctuations.

---

[39] In Hairault et al. (2008), we show that this is true in the case of the fixed cost of recruiting recently introduced by Pissarides (2007).

3.4.2 Unemployment and non-linearity: a test on panel data

The cost of business cycles induced by matching frictions relies crucially on the convexity between unemployment and total factor productivity. Equation (17) can be used as a guideline to gauge whether more volatility is associated with substantially higher unemployment after controlling for other determinants of unemployment. In this section, empirical estimates of this effect, based on cross-country data, are provided[40].

We use the panel data constructed by Bassanini and Duval (2006) over the period 1982-2003 for 20 OECD countries[41], which provide the unemployment rate (consistent with Shimer data), the cyclical component of the total factor productivity process (tfp) and time-varying labor market institutions (LMI) indicators for each country. The LMI aim at accounting for the structural level of unemployment. They include the average unemployment benefit replacement rate, the tax wedge, an indicator of employment protection legislation, an indicator of product market regulation, an indicator of high corporatism in the wage bargaining process and the union density[42]. In order to test whether the relation between the unemployment rate and the total factor productivity is convex, we augment the standard specification of Bassanini and Duval (2006) by introducing the square of the tfp:

$$u_{i,t} = \sum_l d_l X^l_{i,t} + a \ \mathrm{tfp}_{i,t} + c \ (\mathrm{tfp}_{i,t})^2 + \gamma_i + \lambda_t + \epsilon_{i,t} \tag{2}$$

where $X^l_{i,t}$ denotes the LMI l in country i at time t, λ_t the time fixed effect and γ_i is the country fixed effect. The results of this specification are shown in Column (1) and (2) of Table 6. We then include a possible interaction[43] between tfp and time-invariant LMI, consistently with equation (17). Indeed, besides their direct effects on the structural unemployment, LMI may affect the impact of tfp on unemployment. Furthermore, as the non-linearity too may depend

[40] We thank one referee for suggesting this empirical exercise.

[41] Observations for Finland, Germany and Sweden in 1990 and 1991 are removed from the sample as in Bassanini and Duval (2006). We drop New Zealand and Portugal because there are a lot of missing values in the shock series for these two countries.

[42] See Appendix 2 in Bassanini and Duval (2006) for a detailed description of the variables.

[43] We also allow the time fixed effect (λ_t) to depend on the country-specific LMI.

on the structural conditions of the labor market, we include in our estimation the interaction[44] between time-invariant LMI and the square of tfp. We then estimate the following equation:

$$\begin{aligned} u_{i,t} &= \sum_l d_l X^l_{i,t} + a \ \text{tfp}_{i,t} \left(1 + \sum_l b_l(\bar{X}^l_i - \bar{X}^l)\right) + c \ (\text{tfp}_{i,t})^2 \left(1 + \sum_l e_l(\bar{X}^l_i - \bar{X}^l)\right) \\ &+ \lambda_t \left(1 + \sum_l f_l(\bar{X}^l_i - \bar{X}^l)\right) + \gamma_i + \epsilon_{i,t} \qquad (3) \end{aligned}$$

The time-invariant LMI are measured by the difference between the average of the LMI variable l for country i, $\bar{X}^l_i$, and the average across country of $\bar{X}^l_i$, denoted by $\bar{X}^l$. The direct impact of tfp and the square of tfp are respectively measured by a and c, whereas the interaction terms are given by b_l and e_l. The last column of Table 6 presents the results of this specification.

We first estimate equation (2) without introducing the square of tfp. The results, shown in Column 1 of Table 6, document the role of institutions and tfp in explaining the heterogeneous dynamics of unemployment across countries. Unemployment decreases with tfp. It also depends positively on the average replacement rate, the tax wedge and product market regulation, and negatively on employment protection and on the degree of corporatism. The proportion of workers affiliated to a trade union (union density) is not significant. These estimates are consistent with the results of Bassanini and Duval (2006). Column 2 gives some first evidence on the mechanism we emphasized in previous sections: there is a convex relation between the unemployment rate and productivity. This implies that the volatility of productivity tends to raise average unemployment and thereby confirms that business cycles can induce an employment loss.

We test in Column 3 whether the non linearity between tfp and unemployment depends on the labor market institutions[45]. The results indicate that unemployment benefits amplify the

[44]From equation (17), the impact of LMI is *a priori* ambiguous. However, focusing on the u-convexity shows that LMI that increase the structural unemployment rate also increase the effect of the convexity.

[45]We cannot estimate the equation (3) with the full set of time-invariant institutions as explanatory variables, due to multi-collinearity problems. As shown in Bassanini and Duval (2006), multi-collinearity arises from the high (cross-country) correlation which exists between several of the policy indicators used as explanatory

Table 6: Estimation results

	(1)		(2)		(3)	
	Coef.	P-value	Coef.	P-value	Coef.	P-value
Direct impact of tfp (a and c)						
tfp	-8.26*	[0.08]	-6.14	[0.20]	-9.86**	[0.02]
$(\text{tfp})^2$			317.19**	[0.03]	294.76**	[0.03]
Direct impact of the LMI(d_l)						
replacement rate	0.16***	[0.00]	0.16***	[0.00]	0.15***	[0.00]
tax wedge	0.36***	[0.00]	0.36***	[0.00]	0.34***	[0.00]
employment protection	-0.64	[0.17]	-0.65	[0.16]	0.13	[0.77]
union density	-0.01	[0.85]	-0.01	[0.84]	0.04	[0.24]
product market regulation	0.92***	[0.00]	0.95***	[0.00]	0.35	[0.10]
high corporatism	-1.48***	[0.00]	-1.52***	[0.00]	-1.46***	[0.00]
Interaction terms tfp-LMI (b_l)						
replacement rate					0.01	[0.68]
high corporatism					1.60	[0.18]
Interaction terms $(\textbf{tfp})^2$**-LMI** (e_l)						
replacement rate					0.09**	[0.05]
high corporatism					-3.59*	[0.08]
Interaction terms time fixed effect-LMI (f_l)						
replacement rate					0.06***	[0.00]
high corporatism					-1.45***	[0.00]
Country fixed effect	yes		yes		yes	
Time fixed effect	yes		yes		yes	
Observations	390		390		390	
R^2	0.90		0.90		0.98	

Note: OLS estimations for (1) and (2) and non-linear least square estimation for (3) (with robust p-values). * significant at 10%; ** significant at 5%; *** significant at 1%

non-linear impact of the productivity volatility, whereas high corporatism tends to dampen this impact: LMI variables that increase average unemployment amplify the convexity that

variables. We consider the same variables as in Bassanini and Duval (2006): the replacement rate and high corporatism.

leads to higher business cycle costs.

All these results suggest that the volatility of productivity can induce an employment loss. Using these estimates, we can derive for each country the employment loss induced by the volatility of productivity. From equation (3), the direct employment loss can be computed as $c\,\sigma^2_{tfp}$. For the US, this estimate leads to an employment loss of 0.10 of a percentage point, consistent with the results we obtain in section 2.3.2 for the same definition of the unemployment rate, i.e. for the Shimer data. This employment loss is estimated to be sizable for countries with a higher volatility of productivity such as Canada (0.20), Finland (0.24) or Ireland (0.32). On the other hand, the total impact of the volatility of productivity, taking into account the interaction of the productivity with the LMI, is given by $c\ \sigma^2_{tfp}\left(1+\sum_l e_l(\bar{X}^l_i - \bar{X}^l)\right)$. The LMI amplify the impact of productivity fluctuations in countries where the structural unemployment rate is relatively high: for instance, in France, the total employment loss is estimated at 0.33 of a percentage point whereas the direct employment loss was equal to 0.08.

3.5 Structural policy as a stabilizer

The matching model displays sizable costs of business cycles. However, we agree that these costs cannot be interpreted as the gain of a stabilization policy since the productivity shocks have been exogenously shut off. It can be argued that the business cycle costs give an upper bound for the benefits of stabilization policies.

The next step could have been to address more explicitly the question of stabilization, especially that of the job finding rate p through fiscal or monetary policy, as the fluctuations in p are welfare-degrading. But studying the design of these macroeconomic policies is clearly beyond the scope of this paper. We rather exploit the fact that the structural and the cyclical dimensions are strongly interrelated in our framework. Section 2 has revealed that the costs of business cycles are sensitive to the mean of the job finding rate. The higher the latter, the lower the cost of fluctuations. Beyond the stabilization of the job finding rate volatility, our theory suggests dampening the impact of its fluctuations by increasing the mean of the job finding rate. However, any public intervention in this direction must take care not to introduce any additional distortions in the economy. Due to the existence of unemployment benefits and a high workers' bargaining power, the rigid wage economy is characterized by a suboptimal level of employment. We therefore consider a constant subsidy to firms financed by means of lump-sum taxes equally paid by employed and unemployed workers, allowing the economy to reach the first best allocation in an environment without shocks. This policy implies no trade-off between the stabilization and the resource allocation objectives. Let us note that this subsidy policy has the advantage of not being state-dependent and therefore does not suffer from a potential lack of information[46].

Leaving aside the welfare gain due to the reduction of the structural distortions, we focus on the additional gain implied by the reduction of the cost of fluctuations. We then consider the subsidized rigid wage economy[47] with and without productivity shocks for our benchmark calibration. Let us note that the mean of the job finding rate at the first best is equal to 0.672 (Table 7), much higher than its value in the economy without subsidies (Table 4). The business cycle costs implied by the productivity shocks are then considerably dampened by the structural policy. Whereas the volatility of the job finding rate is not modified, the mean of the unemployment rate is no longer significantly increased by these fluctuations (Table 7). The structural policy reduces the costs of business cycles by one order of magnitude: these costs drop to 0.03%. This result illustrates the fact that reducing Harberger triangles may lead to dampening the welfare cost of Okun gaps. This is the natural policy implication of the existence of strong non-linearities in the matching model. Sizable business cycle costs do not necessarily imply the need for stabilization policies.

[46] Alternatively, it is possible to think of counter-cyclical subsidies to firms. The government would provide subsidies τ_y to firms in order to fully compensate for productivity shocks. This would make the firm's value of a job independent of the business cycle. However, the implementation of this policy would require more information than is usually available on aggregate productivity shocks.

[47] In the flexible wage framework, there is the same interaction between the structural and the cyclical dimensions. But the subsidy, which must take into account the existence of both the domestic production and the low bargaining power of the workers, would also have an effect on the volatility of the job finding rate. On the other hand, there would be still no conflict between the structural and cyclical objectives. Even in the case where a high level of home production would imply the economy to be without ambiguity in over-employment, decreasing the job finding rate would not inflict significant business cycle costs, given the high level of home production.

Table 7: The welfare cost of fluctuations with structural subsidies

Business cycles economy				Stabilized economy			Cost of fluctuations
$E(u_t)$	$E(p_t)$	σ_p	$E(\theta)$	$\bar{u}$	$\bar{p}$	$\bar{\theta}$	
4.46%	0.672	0.069	3.815	4.42%	0.672	3.771	0.03%

3.6 Conclusion

This paper shows that non-linearities in the unemployment dynamics caused by frictions on the labor market can generate sizable costs of fluctuations. Using a reduced-form model of the labor market, these costs are estimated to be between one and two orders of magnitude greater than those computed by Lucas (1987) and those usually obtained in dynamic stochastic general equilibrium models. We also show in the rigid wage version *à la* Hall (2005) that the internal mechanisms of the matching model matter for the magnitude of business cycle costs as they impact the average job finding rate through different non-linearities. Our results emphasize that the welfare cost of fluctuations does not only depend on the variability of aggregate shocks. The persistence of these shocks, but also the level of structural unemployment have important implications. Furthermore, a high structural unemployment rate magnifies the welfare consequences of the volatility and the persistence of macroeconomic shocks. We then show that an employment subsidy, by increasing the mean job finding rate, acts as a stabilization policy. We believe that our results are robust to other matching models solving the Shimer puzzle as the underlying mechanisms are intrinsic to this class of models. These results also suggest that business cycles may reduce average consumption by more in continental European countries which would then suffer from both higher structural unemployment and more costly unemployment fluctuations. Business cycle costs would not be alike across countries.

Overall, the welfare implications of aggregate fluctuations in the matching model question the optimism of Lucas (1987) about the weakness of business cycle costs. The welfare costs can be still considered as weak, but they are obtained in a canonical framework without taking into account the individual risks associated with aggregate unemployment, which has received more attention in the literature since the seminal work of Krusell and Smith (1999). Unemployment fluctuations could then imply welfare costs through both a decrease in aggregate consumption and an increase in individual consumption volatility. There are no reasons to think that these two dimensions are not cumulative, leading to substantial business cycle costs. The verification of this assertion is left to further research.

Finally, our results give strong support to the recent line of research which takes into account labor market frictions in the design of optimal monetary policy (Blanchard and Gali (2008), Faia (2009) and Sala et al. (2008)). However, they also emphasize that leaving aside the non-linearities arising from the matching frictions leads to ignoring an important source of business cycle costs and then potentially to deriving misleading policy recommendations. From a methodological standpoint, our paper then questions first-order approximations to the equilibrium conditions which are extensively used to approximate welfare up to second order in the optimal policy literature[48].

[48] See Schmitt-Grohé and Uribe (2007) and Faia (2009) for a similar point of view.

Appendix A: The steady state unemployment gap

The conditional steady state can be written as a function of the job finding rate:

$$\widetilde{u}_i = \widetilde{u}(p_i)$$

Let $\nu_i = p_i - \bar{p}$, the conditional unemployment rate is therefore:

$$\widetilde{u}_i = \widetilde{u}(\bar{p} + \nu_i)$$

Because the unemployment rate is a convex function of the job finding rate, volatility in the job finding rate affects average unemployment. The unemployment gap ψ_p between an economy characterized by a stable job finding rate and an economy with a volatile job finding rate can be computed as follows:

$$\sum_i \pi_i \widetilde{u}(\bar{p} + \nu_i) = \bar{u} + \psi_p$$

A second order approximation of the left hand side yields:

$$\sum_i \pi_i \left[\widetilde{u}(\bar{p}) + \nu_i \widetilde{u}'(\bar{p}) + \frac{\nu_i^2}{2} \widetilde{u}''(\bar{p}) \right] \approx \bar{u} + \psi_p$$

Which gives:

$$\begin{aligned} \psi_p &\approx \frac{\sigma_p^2}{2} \widetilde{u}''(\bar{p}) \\ \psi_p &\approx \sigma_p^2 \frac{\bar{s}}{(\bar{s} + \bar{p})^3} \end{aligned}$$

A similar calculation gives for the job separation rate:

$$\begin{aligned} \psi_s &\approx \frac{\sigma_s^2}{2} \widetilde{u}''(\bar{s}) \\ \psi_s &\approx -\sigma_s^2 \frac{\bar{p}}{(\bar{s} + \bar{p})^3} \end{aligned}$$

Appendix B: The unemployment gap: the general case

The unemployment dynamics read:

$$u_{t+1} = \bar{s} + (1 - \bar{s} - p_t)u_t \tag{A.1}$$

Define $\phi_t \equiv 1 - \bar{s} - p_t$ and $\mathbb{E}[\phi] \equiv \bar{\phi} = 1 - \bar{s} - \bar{p}$.

$$\phi_t = \rho_p \phi_{t-1} + (1 - \rho_p)\bar{\phi} - \varepsilon_t^p$$

Where ε_t^p is the innovation of the job finding rate process. It is iid, has mean zero and variance $\sigma_{\varepsilon_p}^2$.

The unemployment dynamics can be written:

$$u_{t+1} = \bar{s} + \phi_t u_t \tag{A.2}$$

A backward substitution gives:

$$\begin{aligned} u_{t+1} &= \bar{s} + \sum_{k=0}^{+\infty} \prod_{j=0}^{k} \phi_{t-j}\bar{s} \\ \mathbb{E}[u_{t+1}] &= \bar{s}\left(1 + \sum_{k=0}^{+\infty} \mathbb{E}[\prod_{j=0}^{k} \phi_{t-j}]\right) \end{aligned}$$

Write ϕ as an infinite moving average:

$$\phi_t = \bar{\phi} - \Sigma_{j=0}^{+\infty} \rho_p^j \varepsilon_{t-j}^p$$

The mean of unemployment can then be written:

$$\mathbb{E}[u] = \bar{s}\left(1 + \mathbb{E}\Sigma_{k=0}^{+\infty}[(\bar{\phi} - \Sigma_{j=0}^{+\infty}\rho_p^j \varepsilon_{t-j}^p)...(\bar{\phi} - \Sigma_{j=0}^{+\infty}\rho_p^j \varepsilon_{t-j-k}^p)]\right)$$

Neglecting moments of order above 2, the mean of unemployment can be approximated by:

$$\mathbb{E}[u] \approx \bar{s}\left(\frac{1}{\bar{s}+\bar{p}} + \frac{\sigma^2_{\varepsilon^p}}{1-\rho_p^2}\sum_{k=2}^{+\infty}(1-\bar{s}-\bar{p})^{k-2}\sum_{i=0}^{k-1}(k-1-i)\rho_p^{i+1}\right)$$

which simplifies to:

$$\mathbb{E}[u] \approx \bar{u} + \bar{s}\frac{\sigma^2_{\varepsilon^p}}{1-\rho_p^2}\sum_{k=1}^{+\infty}\frac{\rho_p k(1-\bar{s}-\bar{p})^{k-1}}{1-\rho_p(1-\bar{s}-\bar{p})}$$

This yields equation (7).

A similar calculation gives the consequences of job separation rate fluctuations on the average unemployment rate:

$$\mathbb{E}[u] - \bar{u} \approx -\frac{\sigma^2_{\varepsilon^s}}{1-\rho_s^2}\left[\sum_{k=1}^{+\infty}(1-\bar{s}-\bar{p})^{k-1}\sum_{i=0}^{k-1}\rho_s^{i+1} - \sum_{k=2}^{+\infty}\bar{s}(1-\bar{s}-\bar{p})^{k-2}\sum_{i=0}^{k-1}(k-1-i)\rho_s^{i+1}\right]$$

which simplifies to:

$$\mathbb{E}[u] - \bar{u} \approx -\frac{\sigma^2_{\varepsilon^s}}{1-\rho_s^2}\left[\frac{\rho_s}{1-\rho_s}\left(\frac{1}{\bar{s}+\bar{p}} - \frac{\rho_s}{1-\rho_s(1-\bar{s}-\bar{p})}\right) - \frac{\rho_s}{1-\rho_s(1-\bar{s}-\bar{p})}\frac{\bar{s}}{(\bar{s}+\bar{p})^2}\right]$$

This yields equation (8).

Appendix C: Job finding rate data

Figure A.1: Job finding rate - Shimer data

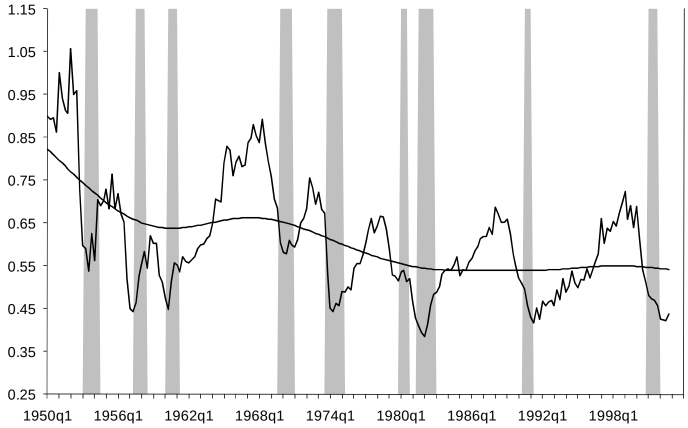
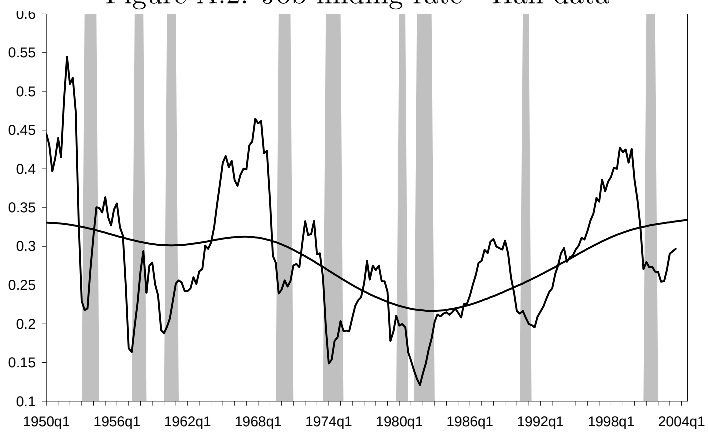


Figure A.2: Job finding rate - Hall data



Appendix D: Alternative specification: log-normal shocks

We estimate the following AR(1) process:

$$ln(p_t) = a_{lp} + \rho_{lp} ln(p_{t-1}) + \varepsilon_t^{lp} \quad (A.3)$$

$$(A.4)$$

Table A.1 shows that the implied characteristics for the level of p is relatively close to those displayed in Table 1. Using the HP filter makes the choice of estimating in log or in level pointless.

Table A.1: Job finding rate fluctuations under log-normality

	Hall data	Shimer data
Estimated process of $ln(p)$		
Standard deviation of ε^{lp}	0.097	0.067
Autocorrelation ρ_{lp}	0.913	0.913
Implied dynamics of p		
Mean $\bar{p}$	0.285	0.607
Standard deviation σ_p	0.068	0.10
Autocorrelation	0.913	0.912
Business cycle cost	0.43%	0.13%

Note: Quarterly average of monthly data. Sample covers 1948q3-2004q3 for Hall (2005) and 1951q1-2003q4 for Shimer (2005). Following Shimer (2005), both sets of data are detrended with an HP smoothing parameter of 10^5.

Appendix E: Non linearities in the endogenous job finding rate: the flexible wage case

In this Appendix, we show that the condition ensuring the concavity of the job finding rate is less stringent when wages are flexible than when wages are rigid. We use the comparative statics of the model without aggregate shocks to approximate the response of the job finding rate.

For a level y of productivity, the equilibrium vacancy-unemployment ratio θ is characterized by the following equation.

$$H(\theta, y) \equiv \frac{\kappa}{q(\theta)} - \Psi\left((1-\gamma)(y-z) - \gamma\kappa\theta\right) = 0$$

where $\Psi \equiv \frac{\beta}{1-\beta(1-s)}$

Let us define

$$G(\alpha) \equiv \frac{\partial^2 p}{\partial y^2} = p''(\theta)(\theta'(y))^2 + p'(\theta)\theta''(y)$$

Using the function $H(\theta, y)$ to compute the implicit derivatives $\theta'(y)$ and $\theta''(y)$, we then obtain:

$$G(\alpha) = \frac{\alpha(1-\alpha)\Psi^2(1-\gamma)^2}{\Gamma(\theta)^2}\left[-\varphi\theta^{\alpha-2} + \frac{\alpha\kappa\theta^{-2}}{\Gamma(\theta)}\right]$$

where $\Gamma(\theta) = \kappa\varphi^{-1}(1-\alpha)\theta^{-\alpha} + \Psi\gamma\kappa$

We deduce that the job finding rate is a concave function of productivity if:

$$G(\alpha) < 0 \quad \Leftrightarrow \quad 2\alpha - 1 - \Psi\varphi\gamma\theta^{\alpha} < 0$$

If $\alpha = 1/2$, then:

$$G(1/2) = -\Psi\varphi\gamma\theta^{1/2} < 0$$

The job finding rate is a strictly concave function of productivity when $\alpha = 1/2$ (i.e at the rigid wage threshold). Because $2\alpha - 1 < 0$ for $\alpha < 1/2$, this restriction is also satisfied for any

$\alpha \in [0, 1/2]$. Then, we deduce that the concavity of the job finding rate is more probable in the case of flexible wages.

Conclusion générale

Les fluctuations conjoncturelles sont traditionnellement étudiées à partir des interactions entre les décisions d'un ménage et d'une entreprise représentatifs. Ce cadre d'analyse à agents représentatifs occulte le rôle des variations du nombre d'entreprises et de travailleurs. Cette thèse illustre l'apport de l'introduction des flux de création et destruction d'emploi et d'entreprises à l'analyse du cycle économique. Nous nous sommes intéressés en particulier à l'impact des imperfections de marché sur les mécanismes de propagation et d'amplification des chocs conjoncturels, ainsi que sur le coût en bien-être des fluctuations. Rendre compte des flux de créations et destructions donne un éclairage nouveaux à ces questions traditionnelles de la macroéconomie.

Dans le premier chapitre, nous analysons les conséquences des frictions financières et montrons que tenir compte des flux de destruction d'entreprises conduit à l'identification d'un nouveau canal d'amplification. Parce qu'elles réduisent la profitabilité des entreprises, les contraintes de crédit accroissent le nombre moyen d'entreprises liquidées. Mais surtout, les contraintes de crédit augmentent le nombre d'entreprises vulnérables à un choc conjoncturel. Suite à un choc de productivité agrégée négatif, le taux de sortie des entreprises s'accroît davantage lorsque les entreprises sont soumises à des contraintes financières, amplifiant ainsi l'effet du choc initial sur la production agrégée. La mise en évidence de ce nouveau canal d'amplification repose sur la modélisation de l'hétérogénéité entre entreprises : les contraintes financières amplifient les fluctuations du taux de sortie en rendant vulnérables des entreprises qui n'auraient pas été conduites à liquider leur activité en l'absence de ces contraintes. Des contraintes de crédit plus importantes conduisent des entreprises de plus en plus productives à liquider leur activité. L'originalité de ce travail est de montrer que les contraintes de crédit ne modifient pas seulement la composition des flux de destruction d'entreprises, mais également leur volatilité. Les simulations du modèle, calibré pour reproduire un coût de défaut égal à 10% des actifs, indiquent que l'accroissement du taux de sortie pourrait être plus deux fois plus importante qu'en l'absence de frictions, réduisant ainsi la production de 0.6 point de pourcentage supplémentaire. Afin d'isoler les effets de ce nouveau mécanisme de ceux bien connus de l'accélérateur financier, les simulations ont été menées en négligeant les effets du choc de productivité sur les fonds propres des entreprises. Une évaluation complète des effets des contraintes financières sur la marge extensive devrait alors tenir compte, non seulement de l'effet de vulnérabilité mis en évidence dans ce chapitre, mais également des effets de la baisse des fonds propres. Car l'accélérateur financier, bien que mis en évidence à la marge intensive (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997), produit également des effets d'amplification similaires à la marge extensive : la baisse des fonds propres s'avère susceptible d'accentuer la hausse du taux de sortie en récession. Une évaluation quantitative complète de la contribution de la marge extensive reste donc à établir. Par ailleurs, ce chapitre appelle des prolongements empiriques. Le modèle délivre en effet des prédictions sur les liens entre taux de sortie et conditions financières qu'il serait intéressant de tester. Tout d'abord, le modèle prédit que le taux de destruction d'entreprises est plus élevé mais aussi plus volatil dans les pays dont les marchés de capitaux sont peu développés. Le modèle prédit de plus que la productivité relative des entreprises liquidées devrait être plus élevée dans ces pays.

Le second chapitre étudie le rôle des frictions affectant le capital et le travail sur la dynamique cyclique de la productivité agrégée. En rendant compte des flux microéconomiques sous-tendant les évolutions macroéconomiques observées, ce chapitre permet d'enrichir la compréhension des fluctuations de la productivité agrégée. Dans cette perspective, nous étudions le rôle des réallocations de facteurs de production, ainsi que ceux des flux de destruction et de création d'entreprises dans la dynamique de la productivité. La croissance de la productivité agrégée ne dépend en effet pas uniquement des variations de la productivité au niveau de chaque entreprise, mais également de l'efficacité dans l'allocation des facteurs de production entre entreprises, ainsi que des changements de composition des flux d'entrée-sortie d'entreprises. Nous proposons alors une nouvelle méthode permettant d'isoler les effets de chacune de ces composantes sur la croissance de la productivité agrégée. A la différence des décompositions existantes, la décomposition mise en oeuvre dans ce chapitre repose sur une fonction de production agrégée construite à partir des fonctions de production microéconomiques. En outre, la mesure des changements d'efficacité allocative que nous proposons repose sur un critère théoriquement fondé. Cette mesure est directement liée à l'écart aux conditions du premier ordre déterminant l'allocation optimale des ressources. Plus encore que dans le premier chapitre, la présence d'hétérogénéité entre entreprise est ici cruciale. En l'absence d'hétérogénéité entre entreprises, la question de l'efficacité de l'allocation des ressources entre entreprises se vide de sens. La productivité agrégée n'est en effet affectée que par les frictions qui distordent l'allocation des ressources entre entreprises hétérogènes. Estimée sur données françaises sur la période 1991-2006, la décomposition révèle que la contribution qu'ont les flux d'entrée-sortie d'entreprises à la dynamique de la productivité agrégée est négligeable. Les principaux déterminants de la productivité agrégée sont les changements de productivité au sein de chaque entreprise, ainsi que les changements dans l'efficacité de l'allocation des ressources entre entreprises. Ces résultats remettent ainsi en question la place centrale occupée par l'entrée et la sortie d'entreprises parmi les déterminants microéconomiques des fluctuations de la productivité agrégée. Alors que la littérature sur l'effet de *cleansing* souligne l'impact positif des destructions d'entreprises sur la productivité agrégée pendant les récessions, nous montrons qu'en plus d'avoir un faible impact sur la productivité agrégée, leur contribution est sensiblement procyclique. Finalement, ce serait plutôt l'effet des réallocations entre entreprises existantes qui tend à accroître la productivité agrégée pendant les périodes de récessions. Plus précisément, nous trouvons que l'efficacité allocative s'accroît pendant les périodes de récession, limitant ainsi la procyclicité de la productivité agrégée.

Les résultats sur la contribution des entrées-sorties d'entreprises et de l'efficacité allocative contrastent avec ceux obtenus à partir de décompositions usuelles. Il convient toutefois de préciser que notre décomposition, à l'instar des décompositions usuelles, ne permet d'évaluer pas d'évaluer les effets de long-terme des entrées-sortie sur la productivité agrégée ; seuls les effets sur la productivité contemporaine sont captés par notre mesure. A long terme, la contribution de la marge extensive est susceptible d'être plus élevée si le taux de croissance de la productivité des entreprises nouvellement créées est supérieur, à long-terme, à celui des entreprises détruites. Enfin, ce travail appelle des recherches plus approfondies quant aux sources de frictions les plus importantes pour la dynamique de la productivité agrégée. En effet, notre approche permet de capter les effets de l'ensemble des frictions auxquelles sont exposées les entreprises, mais ne nous permet pas de discriminer entre elles. Ce travail mériterait donc d'être prolongé par une évaluation de la contribution respective des différentes sources de frictions que sont les coûts d'ajustement, les contraintes financières et les coûts d'appariement.

Dans le dernier chapitre, nous évaluons l'impact des fluctuations conjoncturelles sur le bien-être des ménages en présence de frictions sur le marché du travail. Nous nous intéressons en particulier aux frictions intrinsèquement liées aux flux de création et destruction d'emploi : les frictions d'appariement. Etudier l'effet des fluctuations conjoncturelles à partir des flux d'emplois et non directement à partir des variations nettes d'emploi nous a permis de mettre en évidence une dimension supplémentaire dans l'analyse du coût des fluctuations. Celles-ci sont coûteuses pour les ménages non seulement du fait de l'incertitude qu'elles font peser sur leur consommation, mais également du fait des conséquences indirectes qu'elles exercent sur les niveaux moyens de consommation. Cet effet de la volatilité sur la moyenne de la consommation résulte des non-linéarités dans l'interaction entre stock et flux d'emploi. En particulier, nous montrons que les fluctuations du taux de retour à l'emploi observées aux Etats-Unis ont fait augmenter d'environ un point de pourcentage le niveau de chômage moyen. Nos résultats suggèrent aussi que les fluctuations conjoncturelles seraient plus dommageables pour les pays à fort taux de chômage structurel, ce dernier étant défini comme le taux de chômage en l'absence de fluctuations. Ceci révèle ainsi l'existence d'une forte interaction entre les dimensions cycliques et structurelles du chômage. Les conséquences de cette hausse du taux de chômage moyen sur le bien-être des ménages dépendent évidemment de ses effets sur le niveau de consommation. A l'extrême, si l'on considère, comme Hagedorn and Manovskii (2008), que la production domestique est aussi productive que la production de marché, alors l'effet des fluctuations sur le niveau de chômage moyen n'est pas dommageable pour les ménages. Plus généralement, le coût en bien-être des fluctuations dépend de la valeur relative de la production domestique ainsi que de la désutilité associée au non-emploi. Notons également que les implications de politique économique de ce travail sont potentiellement importantes. Ce chapitre invite notamment à réévaluer les politiques optimales de stabilisation en tenant compte des non-linéarités induites par les frictions d'appariement. En générant un coût en bien-être des fluctuations non négligeable, la présence des ces frictions est susceptible de modifier le poids attribué à la stabilisation de la production dans les règles de politique monétaire optimale.

Bibliography

Andolfatto, D. (1996). Business cycles and labor-market search. *American Economic Review*, 86(1):112–32.

Andrade, G. and Kaplan, S. (1998). How costly is financial (not economic) distress? evidence from highly leveraged transactions that became distressed. *Journal of Finance*, 53(5):1443–1493.

Arellano, C., Bai, Y., and Zhang, J. (2009). Firm dynamics and financial development.

Baily, M., Bartelsman, E. J., and Haltiwanger, J. (2001). Labor productivity: Structural change and cyclical dynamics. *The Review of Economics and Statistics*, 83(3):420–433.

Barlevy, G. (2003). Credit market frictions and the allocation of resources over the business cycle. *Journal of Monetary Economics*, 50(8):1795–1818.

Bartelsman, E. J., Haltiwanger, J. C., and Scarpetta, S. (2009). Cross-country differences in productivity: The role of allocation and selection. NBER Working Papers 15490.

Bassanini, A. and Duval, R. (2006). Employment patterns in oecd countries: Reassessing the role of policies and institutions. OECD Economics Department Working Papers 486.

Basu, S. and Fernald, J. G. (2002). Aggregate productivity and aggregate technology. *European Economic Review*, 46(6):963–991.

Basu, S., Pascali, L., Schiantarelli, F., and Serven, L. (2009). Productivity, welfare and reallocation: Theory and firm-level evidence. NBER Working Papers 15579.

Beaudry, P. and Pages, C. (1999). The cost of business cycles and the stabilization value of unemployment insurance. *European Economic Review*, 45(8):1545–1572.

Bernanke, B. and Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review*, 79(1):14–31.

Bernanke, B. S., Gertler, M., and Gilchrist, S. (1999). *Handbook of Macroeconomics*, volume 1, chapter The financial accelerator in a quantitative business cycle framework, pages 1341–1393. Elsevier, 1 edition.

Bilbiie, F., Ghironi, F., and Melitz, M. (2007). Endogenous entry, product variety, and business cycles. NBER Working Papers.

Blanchard, O. and Gali, J. (2008). Labor markets and monetary policy: a new-keynesian model with unemployment. NBER WP 13897.

Braun, M. and Larrain, B. (2005). Finance and the business cycle: International, inter-industry evidence. *Journal of Finance*, 60(3):1097–1128.

Bricongne, J.-C., Fontagné, L., Gaulier, G., Taglioni, D., and Vicard, V. (2009). Firms and the global crisis: French exports in the turmoil. Working papers 265, Banque de France.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly Journal of Economics*, 121(2):541–585.

Buera, F. and Moll, B. (2011). Aggregate implications of a credit crunch.

Bunel, M., Duhautois, R., and Gonzalez, L. (2009). Types de fusions-acquisitions et évolution de l'emploi des entreprises restructurées. *Travail et Emploi*, 117.

Caballero, R. J. and Hammour, M. L. (1994). The cleansing effect of recessions. *American Economic Review*, 84(5):1350–68.

Carlstrom, C. T. and Fuerst, T. S. (1997). Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review*, 87(5):893–910.

Chang, Y., Kim, S., and Schorfheide, F. (2010). Labor-market heterogeneity, aggregation, and the lucas critique. NBER Working Papers 16401, National Bureau of Economic Research, Inc.

Chari, V., Kehoe, P., and McGrattan, E. (2007). Business cycle accounting. *Econometrica*, 75(3):781–836.

Claessens, S., Kose, M. A., and Terrones, M. E. (2009). What happens during recessions, crunches and busts? *Economic Policy*, 24:653–700.

Clementi, G. and Palazzo, D. (2010). Entry, exit, firm dynamics, and aggregate fluctuations. Unpublished manuscript.

Cogley, T. and Nason, J. (1995). Output dynamics in real-business-cycle models. *American Economic Review*, 85(3):492–511.

Cole, H. and Rogerson, R. (1999). Can the mortensen-pissarides matching model match the business-cycle facts. *International Economic Review*, 40(4):933–1959.

Cooley, T. F. and Quadrini, V. (2001). Financial markets and firm dynamics. *American Economic Review*, 91(5):1286–1310.

Costain, J. and Reiter, M. (2008). Business cycles, unemployment insurance, and the calibration of matching models. *Journal of Economic Dynamics and Control*, 32:1120–1155. forthcoming.

Davis, S. J. and Haltiwanger, J. (1990). Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. *NBER Macroeconomics Annual*, 5:123–186.

De Santis, M. (2007). Individual consumption risk and the welfare cost of business cycles. *American Economic Review*, 97:1488–1506.

Del Gatto, M., Ottaviano, G., and Pagnini, M. (2008). Openness to trade and industry cost dispersion: Evidence from a panel of italian firms. *Journal of Regional Science*, 48(1):97–129.

Dell'Ariccia, G., Detragiache, E., and Rajan, R. (2008). The real effect of banking crises. *Journal of Financial Intermediation*, 17(1):89–112.

Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy*, 90(5):881–94.

Dixit, A. K. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy*, 97(3):620–38.

Dunne, T., Roberts, M., and Samuelson, L. (1989). The growth and failure of u.s. manufacturing plants. *The Quarterly Journal of Economics*, 104(4):671–98.

Eisfeldt, A. and Rampini, A. (2006). Capital reallocation and liquidity. *Journal of Monetary Economics*, 53(3):369–399.

Elsby, M., Hobijn, B., and Sahin, S. (2008). Unemployment dynamics in the oecd. NBER Working Papers 14617, National Bureau of Economic Research, Inc.

Epaulard, A. and Pommeret, A. (2003). Recursive utility, growth, and cost of volatility. *Review of Economic Dynamics*, 6(2):672–684.

Faia, E. (2009). Ramsey monetary policy with labour market frictions. *Journal of Monetary Economics*, 56.

Fisher, F. (1969). The existence of aggregate production functions. *Econometrica*, 37(4):553–577.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Foster, L., Haltiwanger, J. C., and Krizan, C. J. (2001). Aggregate productivity growth. lessons from microeconomic evidence. In *New Developments in Productivity Analysis*, NBER Chapters, pages 303–372. National Bureau of Economic Research, Inc.

Frisch, R. (1933). *Essays in Honor of Gustav Cassel*, chapter Propagation and Impulse Problems in Dynamic Economics.

Fujita, S. and Ramey, G. (2008). The cyclicality of separation and job finding rates. *International Economic Review*. forthcoming.

Godement (2001). *Analyse Mathématique I*. Springer.

Griliches, Z. and Regev, H. (1995). Productivity and firm turnover in israeli industry: 1979-1988. *Journal of econometrics*, 65:175–203.

Hagedorn, M. and Manovskii, I. (2008). The cyclical behaviour of equilibrium unemployment and vacancies revisited. *American Economic Review*, 98(4):1692–1706.

Hairault, J., Langot, F., and Osotimehin, S. (2008). Unemployment dynamics and the cost of business cycles. WP 3840, IZA.

Hall, R. (2005). Employment efficiency and sticky wages: Evidence from folows in the labor market. *Review of Economics and Statistics*, 87(3):397–407.

Hall, R. (2006). Work-consumption preferences and employment volatility. Unpublished.

Hall, R. and Milgrom, P. (2008). The limited influence of unemployment on the wage bargain. *American Economic Review*, 98(4):1653–1674.

Hall, R. E. (1991). Invariance properties of solow's productivity residual. NBER Working Papers 3034.

Hennessy, C. A. and Whited, T. M. (2007). How costly is external financing? evidence from a structural estimation. *Journal of Finance*, 62(4):1705–1745.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–50.

Hornstein, A., Krusell, P., and Violante, G. (2007). Modelling capital in matching models: Implications for unemployment fluctuations. mimeo,.

Houthakker, H. (1955). The pareto distribution and the cobb-douglas production function in activity analysis. *Review of Economic Studies*.

Hsieh, C. and Klenow, P. (2009). Misallocation and manufacturing tfp in china and india. *Quarterly Jounral of Economics*, 124(4).

Jung, P. and Kuester, K. (2009). The (un)importance of unemployment fluctuations for welfare. Federal Reserve Bank of Philadelphia Working Paper.

Justiniano, A., Primiceri, G. E., and Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, 57(2):132–145.

Kennan, J. (2006). Private information, wage bargaining and employment fluctuations. NBER Working Papers.

King, R. and Rebelo, S. (1999). *Handbook of Macroeconomics*, volume 1, chapter 14, pages 927–1007. Elsevier, 1 edition.

Kitao, S., Ljungqvist, L., and Sargent, T. (2008). A life cycle model of trans-atlantic employment experiences. NYU mimeo.

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of Political Economy*, 105(2):211–48.

Krebs, T. (2007). Job displacement risk and the cost of business cycles. *American Economic Review*, 97(3):664–686.

Kroszner, R. S., Laeven, L., and Klingebiel, D. (2007). Banking crises, financial dependence, and growth. *Journal of Financial Economics*, 84(1):187–228.

Krusell, P., Mukoyama, T., Sahin, A., and Smith, A. (2009). Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12:393–402.

Krusell, P. and Smith, A. (1999). On the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 2(1):245–272.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–70.

Lee, Y. and Mukoyama, T. (2008). Entry, exit and plant-level dynamics over the business cycle.

Long, J. and Plosser, C. (1983). Real business cycles. *Journal of Political Economy*, 91(1):39–69.

Lucas, R. (1987). *Models of Business Cycles.* Blackwell Publishing.

Lucas, R. E. (1977). Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5(1):7–29.

Malinvaud, E. (1993). A framework for aggregation theories. *Ricerche Economiche*, 47:107–135.

Matheron, J. and Maury, T. (2000). The welfare cost of fluctuations in ak growth models. mimeo, University of Paris.

McDonald, R. L. and Siegel, D. R. (1985). Investment and the valuation of firms when there is an option to shut down. *International Economic Review*, 26(2):331–49.

Merz, M. (1995). Search in the labour markert and the real business cycle. *Journal of Monetary Economics*, 36(2):269–300.

Miao, J. (2005). Optimal capital structure and industry dynamics. *Journal of Finance*, 60(6):2621–2659.

Mortensen, D. and Nagypal, E. (2003). More on unemployment and vacancy fluctuations. *Review of Economic Dynamics*, 61:397–415.

Mortensen, D. T. (1982). Property rights and efficiency in mating, racing, and related games. *American Economic Review*, 72(5):968–79.

Nataf, A. (1948). Sur la possibilité de construction de certains macromodèles. *Econometrica*, 16:232–244.

Nickell, S., Nunziata, L., and Ochel, W. (2005). Unemployment in the oecd since the 1960s, what do we know. *Economic Journal*, 115:1–27.

Ouyang, M. (2009). The scarring effect of recessions. *Journal of Monetary Economics*, 56(2):184–199.

Petrin, A. and Levinsohn, J. (2005). Measuring aggregate productivity growth using plant-level data. Technical report.

Petrin, A., Reiter, J., and White, K. (2011). The impact of plant-level resource reallocations and technical progress on u.s. macroeconomic growth. *Review of Economic Dynamics*, 14(1):3–26.

Petrongolo, B. and Pissarides, C. (2001). Looking into the black box: a survey of the macthing function. *Journal of Economic Literature*, 39:716–741.

Pissarides, C. (2007). The unemployment volatility puzzle: Is wage stickiness the answer? The Walras-Bowley lecture, North American Summer Meetings of the Econometric Society, Duke University.

Pissarides, C. A. (1985). Short-run equilibrium dynamics of unemployment vacancies, and real wages. *American Economic Review*, 75(4):676–90.

Ramey, G. and Ramey, V. (1993). Technology commitment and the cost of economic fluctuations. *NBER Working Paper No. W3755*, pages 433–480.

Reis, R. (2009). The time-series properties of aggregate consumption: implications for the costs of fluctuations. *Journal of European Economic Association*, 7.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous plants. *Review of Economic Dynamics*, 11(4):707–720.

Sala, L., Soderstrom, U., and Trigari, A. (2008). Monetary policy under uncertainty in an estimated model with labour market frictions. *Journal of Monetary Economics*, 55:983–1006.

Samaniego, R. M. (2008). Entry, exit and business cycles in a general equilibrium model. *Review of Economic Dynamics*, 11(3):529–541.

Scarpetta, S., Hemmings, P., Tressel, T., and Woo, J. (2002). The role of policy and institutions for productivity and firm dynamics: Evidence from micro and industry data. OECD Economics Department Working Papers.

Schmitt-Grohé, S. and Uribe, M. (2007). Optimal, simple, and implementable monetary and fiscal rules. *Journal of Monetary Economics*, 54:1702–1725.

Shimer, R. (2004). The consequences of rigid wages in search models. *Journal of the European Economic Association*, 2(2-3):469–479.

Shimer, R. (2005). The cyclical behaviour of equilibrium unemployment and vacancies. *American Economic Review*, 95(1):25–49.

Shimer, R. (2007). Reassessing the ins and outs of unemployment. Nber working papers, National Bureau of Economic Research, Inc.

Silva, J. and Toledo, M. (2008). Labour turnover costs and the behavior of vacancies and unemployment. *Macroeconomic Dynamics*. forthcoming.

Slutsky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica*, 5:105–146.

Solow, R. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3):312–320.

Stokey, N., Lucas, R. E., and Prescott, E. (1989). *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.

Storesletten, K., Telmer, C., and Yaron, A. (2001). The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk. *European Economic Review*, 45:1311–1339.

Sylvain, A. (2003). Lois de mortalité et durées de vie des équipements dans l'industrie. *Bulletin de la Banque de France*, (111).

Wolfers, J. (2003). Is business cycle volatility costly? evidence from surveys of subjective wellbeing. *International Finance*, 6(1):1–26.