The Macroeconomic Impact of Firm-level Heterogeneity

Tom Schmitz

TESI DOCTORAL UPF / 2015

Director de la tesi: Jaume Ventura

Departament d'Economia i Empresa

Für Michael

Acknowledgements

First of all, I want to thank my advisor, Jaume Ventura. It has been very rewarding to learn from his profound knowledge of economics and his outstanding ability to see the essential behind the superfluous. I am indebted to Jaume as well for helping me to get into conferences and summer schools, and for employing me as his Research and Teaching Assistant. Teaching his classes brought about some of the most enjoyable moments of the last five years.

Andrea Caggese, Jordi Galí and Gino Gancia helped me a lot during the Job Market, both with their letters and with their comments on my paper. Vasco Carvalho, whether in Barcelona or Cambridge, has always given me good advice and support. Christian Fons-Rosen took a lot of time to help me with econometrics, interviews and presentations. Among the many other professors from whom I learned in Barcelona, I want to single out Thijs van Rens, Kurt Schmidheiny, Antonio Ciccone, Alberto Martín, Fabio Canova, José-Luis Peydró, Fernando Broner, Paula Bustos and Giacomo Ponzetto. And without Marta Araque, Laura Agustí, Mariona Novoa and Esther Xifré, it is doubtful whether things would have worked out.

I also want to thank Markus Brunnermeier for inviting me for a short, but insightful visit to Princeton in April 2015. Finally, I will always be grateful to Jean-Yves Chevalier and Jean-Pierre Noreck from the Lycée Henri IV in Paris for everything that they taught me.

Over the course of the past five years, I have been very lucky to make many great friends. The best of them all was Michael Anreiter, to whom this thesis is dedicated. Mapís, Oriol and Jagdish were always there for me when I needed them. Gene, Ciccio and Bruno made traveling to Mallorca and Boston seem almost like a fun school trip. I was also lucky enough to meet Hrvoje, Stefan, Nuno, Tomás, Francisco, Andrei, Emma, Christoph, Giulia, Flo, Patrick, Thomas, Niklas, José, Miguel Angel, Kiz, Fran, Felipe, María, Cristina, Ana, Marco, Lorenzo, Ricardo, Esteban, Pau, Mahesh, Peter, Niek, Sonja and Miguel Karlo. Outside of the PhD bubble, I became friends with Iñigo, Edu, Victor, Cristina Ortíz and the whole team from the Red Cross shelter in Nou Barris, especially Ferrán, Sonia, Albert, Vero and Carolina. Just when I thought that there was nothing new for me in Barcelona anymore, I met Inês. Tu és a mais bonita, simpática, divertida e, acima de tudo, a pessoa mais importante da minha vida. And I cannot say how happy I am that you are already part of the family, just like my parents, my grandparents, my uncle Georges and my brothers Lars, Yves and Rick, who have been a constant source of support from Luxembourg.

Finally, I want to thank some people whom I never met in person. Lucien Favre arrived at Borussia Mönchengladbach on February 14, 2011, to save us from relegation and to bring us back to international football after decades of absence. Seeing us play the way we did in these five years helped me to get over many spells of research frustration. Merci beaucoup, Lucien. Finalmente, querría darles las gracias al Cholo Simeone y a Ada Colau. Me habeís recordado durante estos años en España que nunca hay que resignarse, incluso aunque la tarea parezca imposible. ¡Sí se puede!

Abstract

In this thesis, I show that macroeconomic models which take into account firm- and industry-level heterogeneity provide a more precise understanding of productivity growth dynamics than models with just one firm or industry.

The first chapter is dedicated to cyclical fluctuations in innovation, and highlights the role of the size distribution of innovating firms. Small and young innovating firms are especially important for innovation and productivity growth, but I show that they also suffer most from economic and financial crises. As a result, crises persistently shift the size distribution of innovating firms to the right, and this persistently depresses productivity growth. The size distribution also creates a link between growth and volatility at the industry level.

In the second chapter, I analyse how firm-level innovation decisions and aggregate productivity growth are affected by expansions in international trade. I show that the growth effects of trade depend on whether its expansion is due to falling trade costs for products which were already tradable, or to the possibility to trade products which were not traded before.

Resumen

En esta tesis muestro que los modelos macroeconómicos que consideran la heterogeneidad a nivel empresarial e industrial arrojan un conocimiento más preciso sobre la dinámica de crecimiento de la productividad que los modelos que consideran una sola empresa o industria.

El primer capítulo está dedicado a las fluctuaciones cíclicas en la innovación y destaca el papel de la distribución del tamaño de las empresas innovadoras. Empresas jóvenes y de pequeño tamaño son especialmente importantes para la innovación y el crecimiento de la productividad. Sin embargo, también muestro que éstas son las que más sufren las crisis económicas y financieras. En consecuencia, las crisis desplazan de forma persistente la distribución de tamaño de las empresas hacia la derecha y esto a su vez, genera cambios persistentes en el crecimiento de la productividad. La distribución de tamaño también genera un vínculo entre el crecimiento y la volatilidad a nivel industrial.

En el segundo capítulo analizo cómo afecta la expansión del comercio internacional a las decisiones de I+D+i a nivel empresarial y eventualmente, a la productividad agregada. Muestro que el impacto del comercio internacional sobre el crecimiento depende de si la expansión se debe a un descenso de los costes comerciales de los productos ya intercambiados o a la posibilidad de intercambiar productos que no formaban parte de la cesta de intercambio anteriormente.

Preface

Macroeconomic aggregates, such as the Gross Domestic Product (GDP) of a country, are determined by a huge number of decisions taken by different people, firms and institutions. Analysing these individual decisions with mathematical models is already difficult. However, summing up the interdependent decisions of agents with very different characteristics, possibilities and objectives in order to get a model of aggregate economic activity is even more daunting, and quickly becomes a task of insurmountable complexity.

Macroeconomic models therefore traditionally abstract from taking into account differences between people or firms, and instead try to construct direct relations between aggregate variables (for example, between aggregate consumption and aggregate income, or between aggregate investment and the aggregate real interest rate). This approach dates back to the founding fathers of Macroeconomics, John Maynard Keynes (1936) and John Hicks (1937), and is still largely prevalent today.[1]

Over the last decades, however, advances in computing power and data collection have provided more and more precise empirical evidence on differences in the behaviour of individual people or firms. These new findings exposed the shortcomings of the aggregate approach, and stimulated a lot of theoretical research aiming to identify tractable ways to incorporate micro-level heterogeneity into macroeconomic models.

Empirical studies have shown, for example, that different industries have different rhythms of productivity growth (an hypothesis first formulated by Baumol (1967)) and react differently to the business cycle (Abraham and Katz (1986)). The analysis of large firm-level datasets has uncovered a very high dispersion of sales, investment or productivity even among firms operating in the same industry (Syverson (2011)). Several theoretical models managed to model these heterogeneities in a parsimonious way and to show how they matter for the evolution of aggregate variables. For instance, Melitz (2003) and Hsieh and Klenow (2009) have shown that the allocation of resources across firms with different productivities is a key determinant of the aggregate productivity level. More recently, Gabaix (2011) has pointed out that the size distribution of firms is so heavily skewed that idiosyncratic shocks to the largest firms can have a sizeable impact on aggregate output.

In this thesis, I contribute to the effort of incorporating heterogeneity into macroeconomic

[1] Keynes and Hicks made direct assumptions about these relationships (assuming, for example, that aggregate consumption is an increasing function of aggregate income). Most modern theories instead aim to derive them from the maximization problem of an aggregate, "representative" consumer or firm.

analysis, focusing on endogenous growth theory. The first endogenous growth models (due to Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992)) have shown how private investment in Research and Development (R&D) is determined, and how it drives innovation, technological progress and aggregate productivity growth. In these models, R&D investment is carried out by a single representative firm. However, subsequent empirical evidence has shown that there is great heterogeneity in the R&D decisions of different firms, across and within industries. Furthermore, the seminal contribution of Klette and Kortum (2004) has provided a tractable way to introduce these heterogeneities into a standard endogenous growth model. I build on these insights in order to analyse two concrete cases in which accounting for heterogeneity is crucial to understand aggregate outcomes.

In Chapter 1, I show that economic and financial crises hit the most innovative firms the hardest, and that this amplifies their long-run impact. Precisely, I provide a new theoretical model in which I assume, in line with empirical evidence, that small (and young) firms have a relatively higher innovation capacity than large (and old) ones. My model predicts that small firms however also reduce R&D more than large ones in an economic and financial crisis. Such crises therefore lead to a persistent rightward shift in the size distribution of innovating firms, which persistently depresses productivity growth. Furthermore, my model predicts that there is a positive correlation between the share of innovations done by small firms and the average productivity growth and volatility of an industry. Firm-level evidence from Germany suggests that small firms indeed reduced their R&D more than large ones during the Great Recession. The model's aggregate and industry-level predictions are also consistent with a line of stylized facts.

In Chapter 2, I analyse how international trade affects innovation dynamics in a heterogeneous-firm model with tradable and nontradable products. My model predicts that a fall in trade costs for tradable products increases creative destruction and productivity growth, shifting the size distribution of innovating firms in the tradable sector to the left. However, when some nontradable products become tradable, the effects are quite different: creative destruction and productivity growth increase for the newly tradable products, but fall for all others. The aggregate effects and the reaction of the size distribution are ambiguous. The model's predictions are in line with industry-level evidence on productivity growth in the United States.

Contents

Chapter 1

Fluctuations in R&D Investment and Long-run Growth: The Role of the Size Distribution of Innovating Firms

1.1 Introduction

Private firms' R&D investments create a link between short-run and long-run economic developments. Indeed, while R&D is a key determinant of the long-run productivity level, it also reacts systematically to business cycle fluctuations.

Figure 1.1 illustrates this second point by plotting the deviations of quarterly real private R&D investment and of quarterly real GDP in the United States from their Hodrick-Prescott filtered trends. It shows that fluctuations in both variables are positively correlated over the last decades. In particular, R&D generally falls below trend during recessions. These decreases in R&D slow down innovation, and may therefore adversely affect the level of productivity and output even long after the actual recession has ended.

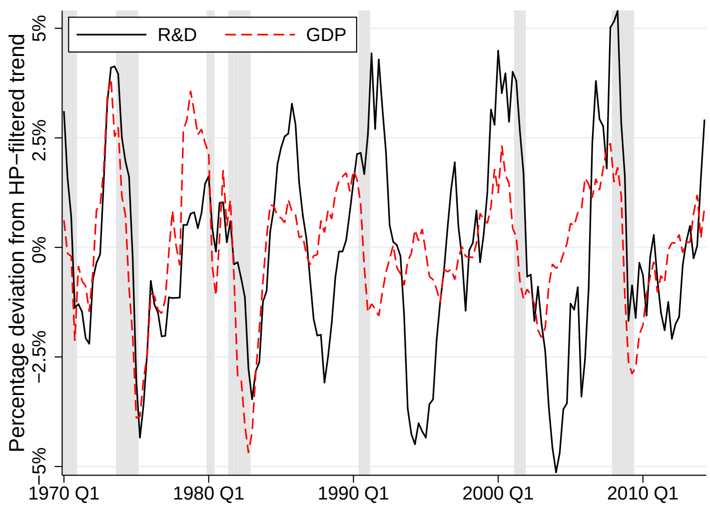
The procyclicality of R&D, widely documented in empirical studies,[1] has generated a large body of research because it challenges some important theoretical priors. For instance, Joseph Schumpeter, who introduced innovation into economic analysis, had argued that recessions prepare the ground for new innovation waves.[2] Other researchers have stressed that the op-

[1] See for example Comin and Gertler (2006), Barlevy (2007), and Ouyang (2011), among many others.

[2] Schumpeter saw short-run innovation fluctuations and long-run growth as inseparable parts of the "capitalist development process". In *The Theory of Economic Development* (Schumpeter (1934)), he argued that crises create the conditions for new innovation waves by lowering factor prices and creating a stock of idle resources. Matsuyama (1999) and Francois and Lloyd-Ellis (2003) formalised this idea in general equilibrium models.

portunity cost of allocating resources away from current production falls in recessions, giving firms an incentive to do more R&D (Aghion and Saint-Paul (1993)). Therefore, the recent literature has mainly focused on identifying mechanisms which can overturn these counter-cyclical forces and explain the procyclicality observed in the data. The leading explanations to date include financial constraints that bind more in recessions (Stiglitz (1993), Aghion et al. (2010), Garcia-Macià (2014)) and the procyclicality of private profits from R&D (Comin and Gertler (2006), Barlevy (2007)).

Figure 1.1: R&D and GDP fluctuations in the United States, 1970-2014



Source: Bureau of Economic Analysis (BEA), National Income and Product Accounts (NIPA). Shaded areas indicate recessions according to the National Bureau of Economic Research (NBER). Time series are computed as percentage deviations from a trend calculated using a Hodrick-Prescott filter with smoothing parameter $\lambda = 1600$.

This theoretical literature can explain why R&D falls during recessions, and provides a framework for analysing the long-run effects of this fall. However, it has, to the best of my knowledge, not yet taken into account firm-level heterogeneity. In particular, it has not considered differences between small and large firms. This is potentially an important omission, as empirical studies suggest that there are substantial differences between the type and the amount of R&D done by small and large firms. For instance, Cohen and Klepper (1996) argue that small innovating firms generate more innovations per dollar of R&D spent. Akçiğit and Kerr (2010) show that small innovating firms spend relatively more on R&D than large ones, and that their innovations are more likely to be major technological advances. These findings suggest that the size distribution of innovating firms matters for aggregate outcomes, and therefore as well for business cycle dynamics.

I explore this hypothesis in a new partial equilibrium endogenous growth model with heterogeneous firms. My model describes an industry in which a continuum of firms produces a fixed mass of differentiated products and improves their productivity through R&D. This basic setup builds on Akçiğit and Kerr (2010) and Klette and Kortum (2004). However, these papers focus on a balanced growth path, while I analyse how R&D dynamics are affected by two types of aggregate shocks: shocks to the level of spending on all products of the industry (aggregate demand shocks) and shocks to firms' ability to borrow (financial shocks).

Firms can improve productivity through two types of R&D. Radical R&D creates radical innovations, enabling the innovating firm to increase the frontier productivity for a product which it does not produce yet (and therefore allowing it to displace that product's previous producer). Incremental R&D creates incremental innovations, enabling the innovating firm to increase the frontier productivity of products which it already produces.

This innovation structure has two crucial features. First, radical innovation triggers creative destruction, and this shapes the size distribution. Firms grow (or enter the industry) doing radical innovation, and they shrink (and eventually exit) when displaced by the radical innovations of others. Second, I assume that the relative innovation capacity of small firms is higher than the one of large firms. Firm size is measured as the number of products produced. Thus, when there are no shocks and firms realise all their innovation opportunities, overall productivity growth in the industry is higher if the fixed mass of products is produced by a large mass of small firms rather than by a small mass of large firms.[3]

My model delivers three main predictions.

Prediction 1. On average, small firms reduce R&D more than large ones after a negative aggregate shock.

This holds for financial shocks because small firms have lower cash holdings, and for aggregate demand shocks because they have lower profit margins. The combination of small firms' greater sensitivity to negative aggregate shocks and their higher relative innovation capacity generates two novel predictions, showing that the size distribution of innovating firms matters for aggregate dynamics.

Prediction 2. A financial shock leads to a persistent rightward shift in the size distribution of innovating firms, and therefore persistently depresses productivity growth.

[3] This assumption is motivated by the aforementioned empirical evidence, showing that small firms spend relatively more and more productively on R&D and are relatively more likely to generate major technological advances than large firms. These differences are probably strongest between firms which are both small and young, and firms which are both large and old. In my model, size and age are strongly positively correlated, but as I cannot observe firm age in the dataset used for my empirical analysis, I only refer to size in this chapter.

A financial shock does not affect large firms, but forces many small ones to abandon radical R&D, as their cash holdings are insufficient to pay the R&D cost. Thus, entry falls, creative destruction slows down and the size distribution shifts persistently to the right. As large firms have a relatively lower innovation capacity than small ones, productivity growth remains depressed even once the actual shock has vanished. An aggregate demand shock amplifies these effects, as it limits the amount of cash available to finance R&D.

Prediction 3. There is a positive correlation between the share of innovations done by small firms, average productivity growth and volatility at the industry level.

More precisely, any change in an industry characteristic which shifts the size distribution of innovating firms to the left (without having a direct effect on productivity) ends up increasing the average rate of productivity growth, as there are now more innovative small firms, but also its volatility, as these small firms are more sensitive to negative aggregate shocks.

The firm-level and aggregate predictions of my model are consistent with empirical evidence. Using a German firm-level dataset, I show that during the 2007-2009 Great Recession (which may be interpreted as a joint aggregate demand and financial shock), the median small firm reduced R&D more than the median large firm, even though its sales fell less. As a result, a significant gap in the relative R&D intensity of small and large firms opened up in the crisis years, in line with Prediction 1.

Prediction 2 is consistent with the experience of several developed countries during the Great Recession. For instance, both in Germany and in the United States, there was a (small) rightward shift in the size distribution of innovating firms. Furthermore, empirical studies indicate a persistent fall in productivity growth after the recession in many OECD countries.[4] Finally, I show that as implied by Prediction 3, the average and the standard deviation of productivity growth are positively correlated across industries in the United States manufacturing sector. The (scarce) industry-level data on the size distribution of innovating firms is consistent with the hypothesis that this new stylized fact is linked to differences in the size distribution.

The remainder of this chapter is structured as follows. Section 1.2 describes my model's assumptions and Section 1.3 derives and discusses its predictions. Section 1.4 presents the empirical evidence and Section 1.5 concludes. Section 1.6 contains the Appendix.

[4] In the United States, there is a debate about whether this slowdown is due to the Great Recession. Fernald (2014) and Hall (2014) dispute this and claim that productivity growth fell even before the recession started. In other OECD countries, however, the direct impact of the Great Recession on productivity growth appears more significant (Ball (2014)).

1.2 A model of R&D investment with heterogeneous firms and aggregate shocks

The seminal model of Klette and Kortum (2004) extended the baseline endogenous growth models to allow for rich firm dynamics. In their model, innovations are homogenous and innovation capacity is exactly proportional to firm size. Thus, the size distribution of innovating firms does not affect aggregate productivity growth. Akçiğit and Kerr (2010) and Acemoğlu and Cao (2010) depart from this model by assuming that small firms' and/or entrants' innovations are more productive than those of large incumbents. Then, the size distribution matters, and size-dependent subsidies or taxes affect aggregate productivity growth.[5]
My model builds on these contributions, but analyses the dynamics triggered by aggregate shocks instead of just considering a balanced growth path.[6] Sections 1.2.1 to 1.2.4 lay out the model's assumptions.

1.2.1 The basic structure

I model an industry in partial equilibrium. Time is discrete ($t \in \mathbb{N}$), and money is the unit of account. The industry produces a fixed set of differentiated products, indexed on the interval $[0,1]$. The next two subsections describe the demand and the supply side of the model.

1.2.1.1 Aggregate demand

Aggregate demand is defined as the amount of money spent on all products of the industry. I assume aggregate demand follows an exogenous Markov chain process $(S_t)_{t\in\mathbb{N}}$ which can take two values, S_H and S_L (with $S_H > S_L$). In every period t, a representative consumer allocates aggregate demand across the products of the industry. The consumer aims to maximise his utility, given by

$$C_t = \exp\left(\int_0^1 \ln c_t(j)\, dj\right), \tag{1.1}$$

where $c_t(j)$ denotes the quantity of product j consumed in period t. He takes the prices of differentiated products and the total amount of money spent as given.

[5] Acemoğlu et al. (2013) analyse this in greater detail, focusing on the allocation of resources for innovation.

[6] My model is related to Ateş and Saffie (2014), who show that firm-level heterogeneity in innovation matters for an economy's response to a sudden stop. In their model, there are both high productivity entrants (in limited supply) and low productivity ones. This heterogeneity dampens the effects of a sudden stop: while entry decreases, the average productivity of entrants increases. The size distribution of innovating firms, which is crucial for my results, plays no role in their model.

1.2.1.2 Firms

Differentiated products are produced by a continuum of atomistic firms. Each firm produces a finite number of products, and I denote by $m_{n,t}$ the mass of firms producing n different products in period t. The number of products produced is my measure of firm size. As all firms are innovators, the firm size distribution coincides with the size distribution of innovating firms.[7]
All products of the interval $[0,1]$ are produced, so that in every period t, $\sum_{n=1}^{+\infty} n m_{n,t} = 1$.
Apart from incumbent firms, there is also a mass $m_{0,t}$ of potential entrants which do not produce in period t. The relative mass of potential entrants with respect to incumbents is constant:

$$\frac{m_{0,t}}{m_t} = \psi, \tag{1.2}$$

where $m_t \equiv \sum_{n=1}^{+\infty} m_{n,t}$ is the mass of incumbent firms and $\psi > 0$ is a fixed parameter (which may be interpreted, for example, as an inverse measure of barriers to entry).
Firms produce with constant returns to scale, using labour. For any firm i,

$$\forall j \in [0,1], \; y_t^i(j) = a_t^i(j)\, l_t^i(j), \tag{1.3}$$

where $y_t^i(j)$ stands for the output of product j by firm i in period t, $l_t^i(j)$ for the labour used in production and $a_t^i(j)$ for firm i's productivity for product j. Note that a firm's productivity may differ across products. Labour supply to firms is given by the reduced-form function $L(w_t)$, increasing in the money wage w_t.
Firms maximize the expected net present value (NPV) of profits earned over their existence. They store all the profits they earn (with a constant rate of return set to 0 for simplicity), and only pay dividends upon exit. Potential entrants start out without cash.
Finally, I assume that there is Bertrand competition on the market for each differentiated product. This implies that for any product j, only the firm with the highest productivity produces in equilibrium, and earns positive profits only if its productivity is strictly higher than that of all other firms. Firms therefore have an incentive to improve their productivity through R&D. The next section describes the R&D technologies at their disposal.

1.2.2 R&D investment and innovation

Firms improve productivity through radical or incremental innovations.[8]

[7] In the theoretical sections of this chapter, I therefore use both terms interchangeably. In real-world datasets, however, the two concepts do not coincide, as many firms never innovate.

[8] This follows the terminology of Acemoğlu and Cao (2010).

Radical innovations. In every period t, every firm (whether incumbent or potential entrant) receives a radical innovation opportunity with a fixed probability α. This opportunity allows the firm to increase the frontier productivity for some product j which it currently does not produce if it pays the radical R&D (money) cost f_R. More precisely, if $a_t(j)$ denotes the frontier productivity for product j in period t,[9] a radical innovation enables the firm to produce product j with productivity $\gamma a_t(j)$ $(\gamma > 1)$ from period $t+1$ onwards. The product to which a radical innovation applies is drawn from a pool of "contestable" products, containing one (randomly selected) product of every incumbent firm.

Radical innovation triggers creative destruction, and thereby changes the firm size distribution: by becoming the most productive producer for product j, the innovating firm displaces the incumbent producer and grows at its expense. A firm which loses all its products in this way is assumed to exit forever. Note that the contestable product assumption implies that a firm can neither gain nor lose more than one product per period.

Incremental innovations. Incumbent firms are not limited to radical innovation, but can also improve the frontier productivity of products they already produce by incremental innovation. Paying an incremental R&D (money) cost f_I in period t enables them to increase their productivity for all their non-contestable products by a factor δ $(\delta > 1)$ in period $t+1$.

I assume that the productivity increases implied by radical and incremental innovation hold

$$\alpha \ln \gamma > \ln \delta. \tag{1.4}$$

This inequality states that if all products were subject to radical innovation (would see their productivity increase by a factor γ with probability α), aggregate productivity would increase more than if all products were subject to incremental innovation (would see their productivity increase by a factor δ for sure).[10] In sum, this implies that radical innovation contributes relatively more to productivity growth than incremental innovation, justifying the names of both innovation types.[11]

[9] That is, $a_t(j) = \max_{i \in \mathscr{I}_t} a_t^i(j)$, where $\mathscr{I}_t$ denotes the set of active firms in period t.

[10] As utility is a Cobb-Douglas aggregator, a natural expression for the industry's aggregate productivity (abstracting from labour misallocation, analysed in Section 1.3.1.3) is $A_t = \exp\left(\int_0^1 \ln a_t(j)\right)$. If all products were subject to radical innovation in one period, a fraction α of them would see their productivity increased by a factor γ, and aggregate productivity would be multiplied by γ^α. If all products were subject to incremental innovation, aggregate productivity would be multiplied by δ. $\gamma^\alpha > \delta$ holds if and only if $\alpha \ln \gamma > \ln \delta$.

[11] An example from the car industry may be useful to further highlight the differences between the two innovation types. A recent radical innovation in this industry is the Tesla Roadster (the first long-range all-electrical sports car, developed by a Silicon Valley start-up). This innovation is radical both because it represents a large technological advance and because it has the potential to displace conventional sports car producers. In

For simplicity, I also assume that one period after an innovation is introduced for a given product, imitation allows all firms to produce with a productivity that is arbitrarily close to the one of the innovator. This limits the profits from innovation to one period.[12]

R&D costs for both types of innovation must be paid in period t, while innovations appear only in period $t+1$. This creates a need for finance for all firms which cannot pay R&D costs with their cash holdings. The next section explains the conditions under which firms can borrow to finance R&D.

1.2.3 Financing conditions

Financing conditions exogenously fluctuate between two states of the world. In the "normal" state, firms can borrow as much as needed for R&D projects with a non-negative NPV. In the "crisis" state, firms cannot borrow at all. Thus, if financing conditions are in the crisis state in period t, every firm must satisfy the constraint

$$c_t \geq \mathbf{1}_{R,t} f_R + \mathbf{1}_{I,t} f_I, \tag{1.5}$$

where c_t stands for the firm's cash holdings in period t (after production and before deciding on R&D) and $\mathbf{1}_{R,t}$ and $\mathbf{1}_{I,t}$ are indicator variables for the decisions to do radical R&D (conditional on getting a radical innovation opportunity) or incremental R&D in period t.
To prevent occasionally binding constraints from inducing precautionary savings, I assume that firms can forecast aggregate demand one period ahead. As profits from innovation are limited to one period, this ensures that an R&D project with a positive expected NPV always increases the firm's cash holdings in the period after the payment of the R&D cost.

1.2.4 The innovation capacity of small and large firms

Before proceeding to solve the model, it is useful to shortly discuss the two key features of my model's innovation technologies.

The first key feature is that radical innovation drives firm dynamics and the evolution of the firm size distribution. Firms grow (or enter the industry) doing radical innovation, and they

contrast, GM's introduction of the OnStar system (a teleassistance system which, for example, automatically contacts an emergency hotline in case of an accident) may be considered an incremental innovation: it is not a major technological advance, and improves previously produced cars without displacing other firms.

[12] As the innovator retains an infinitesimal productivity advantage, it remains the only producer of the product as long as it is not displaced by a radical innovation. In order to simplify notation, I assume in the following that imitators can produce with the exact frontier productivity.

shrink (and eventually exit) when displaced by the radical innovations of others. This is a common feature of all models in the tradition of Klette and Kortum (2004), and in line with the extensive empirical evidence on firm-level creative destruction.
The second key feature is that the relative innovation capacity of small firms is higher than that of large firms. Intuitively, this is mainly due to the fact that radical innovation opportunities are independent of firm size, so that small firms can do relatively more radical innovations with respect to their size. This is not compensated by the fact that large firms can do more incremental innovations, because the number of possible incremental innovations increases only linearly with firm size, and incremental innovations improve productivity less than radical ones. More formally, my assumptions imply that if firms realise all their innovation opportunities, the relative contribution of a firm of size n to aggregate productivity growth is $\frac{\alpha \ln \gamma + (n-1) \ln \delta}{n}$.[13] When Equation (1.4) holds, this number is clearly decreasing in firm size n. Therefore, when firms realise all their innovation opportunities, a leftward shift in the firm size distribution increases productivity growth.[14]

This special importance of small firms for innovation is in line with empirical evidence. The early literature, launched by Schumpeter's conjectures (1934, 1942), shows that small firms generate more innovations per dollar of R&D than large ones (Cohen and Klepper (1996)). In the same vein, Kortum and Lerner (2000) argue that young and small firms promoted by venture capital funds account for a disproportionate share of their industry's innovations. Recent empirical studies also assess the relative technological advances represented by small and large firms' innovations, using patent data. Akçiğit and Kerr (2010) show that the average patent filed by a small firm receives more external citations and that major innovations are relatively more frequent in small firms. Ewens and Fons-Rosen (2013) show that upon leaving an established IT firm and creating a small start-up, patents filed by founders are on average of better quality and more likely to represent pioneering innovations than patents filed by their former co-workers. Overall, these results suggest that small firms indeed have a relatively greater innovation capacity than large ones, as implied by my assumptions.[15]

In the next section, I solve for the equilibrium of the model and derive its predictions.

[13] The relative contribution is defined as the industry's productivity growth rate when all firms have size n and realise all their innovation opportunities.

[14] On these matters, my model is similar to Akçiğit and Kerr (2010). They also assume that radical (in their terminology, "exploration") innovation capacity is independent of firm size and show that this implies a relatively greater innovation capacity of small firms.

[15] Direct evidence on the exact contribution of small and large firms to industry productivity growth is more scarce, due to obvious measurement problems. However, Acs and Audretsch (1990) do find that an industry's innovative activity decreases in its concentration level, in line with my assumptions.

1.3 Equilibrium and predictions

1.3.1 Equilibrium and dynamic laws of motion

I first determine the intraperiod equilibrium on product and labour markets. Then, I solve for firms' R&D choices and derive the aggregate and firm-level laws of motion.

1.3.1.1 The intraperiod equilibrium

The representative consumer maximizes his utility from the consumption of differentiated products defined in (1.1), taking the aggregate amount of money spent S_t and prices as given. This gives the demand functions

$$\forall j \in [0,1], \;\; c_t(j) = \frac{S_t}{p_t(j)}. \tag{1.6}$$

Product market clearing implies $c_t(j) = y_t(j)$ for every j. Thus, defining aggregate output as $Y_t \equiv \exp\left(\int_0^1 \ln y_t(j)\, dj\right)$ gives

$$Y_t = C_t = \frac{S_t}{P_t}, \quad \text{with } P_t = \exp\left(\int_0^1 \ln p_t(j)\, dj\right). \tag{1.7}$$

Bertrand competition implies that the producer of any product sets a price equal to the marginal cost of the second most productive firm.[16] Moreover, imitation implies that the second most productive firm in period t has a productivity equal to the one of the most productive firm in period $t-1$. Therefore,

$$\forall j \in [0,1], \;\; p_t(j) = \frac{w_t}{a_{t-1}(j)}. \tag{1.8}$$

Substituting (1.8) into (1.6) gives $y_t(j) = \frac{S_t a_{t-1}(j)}{w_t}$. Accordingly, the labour demand of the firm producing product j is $l_t(j) = \frac{S_t a_{t-1}(j)}{w_t a_t(j)}$, and the labour market clearing condition is

$$L(w_t) = \frac{S_t}{w_t} \int_0^1 \frac{a_{t-1}(j)}{a_t(j)} dj. \tag{1.9}$$

[16] The equilibrium price is the minimum between the average cost of the second most productive firm (the limit price) and the monopoly price. However, as the demand function defined in (1.6) has a price elasticity of 1, the monopoly price tends towards positive infinity and the limit price is always the equilibrium price.

Equations (1.6) to (1.9) define the equilibrium values of output, prices and wages in period t. They are independent of current R&D decisions, because R&D does not use labour. However, R&D affects future productivities, and thereby future values of output, prices and wages.

1.3.1.2 R&D investment

Innovating firms start by charging a mark-up equal to the factor with which they improve the frontier productivity. Therefore, it is easy to show that a firm improving the productivity of a product by a radical innovation in period $t+1$ earns a profit $\left(1-\frac{1}{\gamma}\right)S_{t+1}$ from this innovation in period $t+1$. Imitation implies that mark-ups and profits fall to 0 in all subsequent periods. Likewise, the profit from producing an incrementally improved product in period $t+1$ is $\left(1-\frac{1}{\delta}\right)S_{t+1}$. Note that profits increase in the size of the productivity improvement and do not depend on the identity of the improved product.

I can now solve for the firms' R&D policy functions. These policy functions depend on two endogenous state variables: the number of products produced by the firm, n_t, and its cash holdings after production, c_t. Even though R&D decisions are binary, they generate a complicated dynamic programming problem. Therefore, I impose three additional parameter restrictions which considerably simplify the problem and deliver an explicit solution.

Restriction 1. $\left(1-\frac{1}{\gamma}\right)S_L - f_R > 0$. The NPV of radical innovation is always positive.

Restriction 2. $\left(1-\frac{1}{\delta}\right)S_H - f_I > 0$. When aggregate demand is high, the NPV of incremental innovation is positive for a firm with one non-contestable product.

Restriction 3. $\left(1-\frac{1}{\gamma}\right)S - f_R > (n_U^* - 1)\left(1-\frac{1}{\delta}\right)S - f_I$ for $S \in \{S_L, S_H\}$.
$n_U^* = \left\lceil \frac{f_R+f_I}{\left(1-\frac{1}{\gamma}\right)S_L - f_R} \right\rceil$, where $\lceil . \rceil$ is the ceiling function, assigning to each positive real number x the smallest natural number larger than x. This restriction implies that the NPV of radical innovation is always greater than the NPV of incremental innovation for firms small enough to be constrained in crisis financing conditions.

I discuss the role of these restrictions when analysing the R&D policy functions they deliver. I first consider a period with normal financing conditions.[17]

[17] I only give an intuitive account of derivations in the main text. Proofs are provided in Appendix Section 1.6.1.1.

R&D decisions for normal financing conditions. Consider a firm with n_t products and cash holdings c_t. Then, if financing conditions are normal in period t,

$$\mathbf{1}_{R,t} = 1,$$

$$\mathbf{1}_{I,t} = 1 \Leftrightarrow n_t \geq n^*_{I,t} = \left\lceil 1 + \frac{f_I}{\left(1 - \frac{1}{\delta}\right) S_{t+1}} \right\rceil.$$

Under normal financing conditions, all firms spend on radical R&D if they get a radical innovation opportunity. This follows immediately from Restriction 1, ensuring that the NPV of a radical innovation is always positive.[18] Firms spend on incremental R&D if and only if its NPV is non-negative. For a firm with n_t products, the NPV of incremental innovation on its $n_t - 1$ non-contestable products is $(n_t - 1)\left(1 - \frac{1}{\delta}\right) S_{t+1} - f_I$. This value increases in n_t (showing that large firms benefit from economies of scope for incremental innovation) and is non-negative if and only if n_t is larger than the threshold size $n^*_{I,t}$ defined above. $n^*_{I,t}$ depends on aggregate demand next period and may therefore take two possible values, denoted by $n^*_{I,H}$ and $n^*_{I,L}$. Restriction 2 implies that $n^*_{I,H} = 2$: when aggregate demand is high and financing conditions are normal, firms realise all their innovation opportunities.[19]

With crisis financing conditions in period t, the desired R&D choices of firms do not change, but they may not be feasible any more for firms with low cash holdings.
n^*_U, defined in Restriction 3, is the threshold size for firms to be forever unconstrained: firms with n^*_U products or more have accumulated enough cash to be unaffected by crisis financing conditions.[20] The threshold size n^*_U is derived by remarking that the lowest possible cash level for a firm with n products is $n\left(\left(1 - \frac{1}{\gamma}\right) S_L - f_R\right)$, as the firm must have earned at least n times the NPV of radical innovation (which is at worst $\left(1 - \frac{1}{\gamma}\right) S_L - f_R$). All firms with less than n^*_U products continue to spend on radical and incremental innovation if and only if their cash holdings are sufficient. Restriction 3 implies that if these firms want to do radical and

[18]This restriction simplifies the firms' problem by excluding situations in which the negative cash flow from a radical innovation must be compared to the positive value it generates by increasing firm size (thus allowing the firm to survive longer and to generate more innovations in the future). It excludes from the outset that radical R&D reacts to demand shocks. This is extreme, but captures the fact that high profit (radical) innovations are less likely to be abandoned than low profit (incremental) ones if future demand is low.

[19]This feature is not important for my results. However, if some small firms with non-contestable products never do incremental R&D, the condition on the relative productivity improvements of radical and incremental innovations given in Equation (1.4) must be strengthened somewhat to ensure that in the absence of negative shocks, small firms still contribute relatively more to productivity growth than large ones.

[20]As the NPV of all realised innovations is positive, and fully collected one period after the R&D investment is made, cash holdings never fall. Therefore, a firm which is unconstrained once is in fact unconstrained forever.

incremental R&D, but have only enough resources to do one of the two, they choose radical R&D.[21] Thus, decisions with crisis financing conditions can be summarized as follows.

R&D decisions for crisis financing conditions. Consider a firm with n_t products and cash holdings c_t. Then, if there are crisis financing conditions in period t,

$$\mathbf{1}_{R,t} = 1 \Leftrightarrow c_t \geq f_R$$

$$\mathbf{1}_{I,t} = 1 \Leftrightarrow n_t \geq n^*_{I,t} \text{ and } \begin{cases} c_t \geq f_R + f_I \text{ or } f_I \leq c_t < f_R & \text{if the firm has a rad. inn. opportunity} \\ c_t \geq f_I & \text{else} \end{cases}$$

Firms' R&D policy functions directly illustrate the model's first important prediction, regarding small and large firms' reactions to a negative aggregate shock.

Prediction 1. On average, small firms reduce R&D more than large ones after a negative aggregate shock.

As firms do not pay dividends before exit, average cash holdings increase in firm size. A financial shock, which forces firms to finance R&D with their cash holdings, therefore especially affects small firms. Aggregate demand shocks also hit small firms (among those with non-contestable products) harder. As they benefit relatively less from economies of scope, their NPV of incremental innovation is low and turns negative if demand falls, while the one of the largest firms always remains positive.

Together with the assumptions ensuring that the relative innovation capacity of small firms is higher than the one of large firms, Prediction 1 is key for my model's aggregate results. Before getting to these, however, I need to complete the model's solution.

1.3.1.3 The dynamic laws of motion

The two previous sections show how equilibrium quantities and prices depend on current productivity, and how firms decide on R&D, which determines future productivity. It is now time to put the pieces together and to determine the model's dynamic laws of motion.

Denote by $u^f_{n,t}$ the fraction of firms of size n which can finance in period t an R&D cost f. By definition, $u^f_{n,t} = 1$ for every n and f if financing conditions are normal in period t. Under crisis financing conditions, $u^f_{n,t}$ is equal to the fraction of firms of size n which have (after production in period t) cash holdings larger or equal to f. Then, the mass of radical innovations

[21] This restriction is not important for my results, but considerably simplifies calculations.

due to R&D in period t is

$$R_t = \alpha \sum_{n=0}^{+\infty} u_{n,t}^{f_R} m_{n,t}. \tag{1.10}$$

The mass of incremental innovations due to R&D in period t is

$$I_t = \alpha \sum_{n=n_{I,t}^*}^{+\infty} (n-1)\left(u_{n,t}^{f_R+f_I} + \max\left(0, u_{n,t}^{f_I} - u_{n,t}^{f_R}\right)\right) m_{n,t} + (1-\alpha) \sum_{n=n_{I,t}^*}^{+\infty} (n-1)\, u_{n,t}^{f_I} m_{n,t}. \tag{1.11}$$

The frontier productivity of radically improved products increases by a factor γ from period t to $t+1$, the one of incrementally improved products increases by a factor δ, and the one of all other products remains unchanged. Therefore, using Equations (1.7) and (1.8), I get a law of motion for the aggregate price level P_t:[22]

$$\frac{P_{t+1}}{P_t} = \frac{w_{t+1}}{w_t} \exp\left(-\left(R_{t-1} \ln \gamma + I_{t-1} \ln \delta\right)\right). \tag{1.12}$$

Using Equation (1.8), I can deduce a law of motion for aggregate output:

$$\frac{Y_{t+1}}{Y_t} = \frac{S_{t+1}}{S_t} \frac{w_t}{w_{t+1}} \exp\left(R_{t-1} \ln \gamma + I_{t-1} \ln \delta\right). \tag{1.13}$$

Finally, defining aggregate productivity as $A_t \equiv \frac{Y_t}{L(w_t)}$, it comes that

$$\begin{aligned} \frac{A_{t+1}}{A_t} &= \exp\left(R_t \ln \gamma + I_t \ln \delta\right) M_t, \\ \text{with } M_t &= \exp\left(\left(R_{t-1} - R_t\right) \ln \gamma + \left(I_{t-1} - I_t\right) \ln \delta\right) \frac{1 - R_{t-1}\left(1-\frac{1}{\gamma}\right) - I_{t-1}\left(1-\frac{1}{\delta}\right)}{1 - R_t\left(1-\frac{1}{\gamma}\right) - I_t\left(1-\frac{1}{\delta}\right)}. \end{aligned} \tag{1.14}$$

Aggregate productivity growth obviously depends on increases in productivity generated by radical and incremental R&D, as captured by the first factor in Equation (1.14). However, there is also a more indirect source of variation, captured by M_t. R&D investment creates mark-up dispersion (unimproved products are sold at marginal cost, improved ones at a mark-up γ or δ) which misallocates labour and depresses aggregate productivity.[23] All else equal, when mark-up dispersion increases between two periods, aggregate productivity falls.

Equations (1.10) and (1.11) show that the firm size distribution determines innovation masses. The firm size distribution itself evolves between periods as a result of radical innovation.

[22]Equation (1.12) is formally derived in Appendix Section 1.6.1.2.

[23]Epifani and Gancia (2011) and Peters (2011) analyse the static and dynamic effects of mark-up dispersion.

Indeed, an incumbent firm producing n products in period t may be in three possible states in period $t+1$.

(a) It may produce $n+1$ products, if it does radical R&D and does not lose its contestable product. By the law of large numbers, the fraction of firms with n products doing radical R&D is $\alpha u_{n,t}^{f_R}$. I denote by d_t the probability that a firm loses its contestable product (in the following, I refer to this as the destruction rate). Receiving a radical R&D investment opportunity and losing a contestable product are independent events, so that finally, a fraction $\alpha u_{n,t}^{f_R}(1-d_t)$ of firms producing n products transitions to producing $n+1$ products.

(b) It may keep producing n products. This happens if it does not do radical R&D and does not lose its contestable product (which is the case for a fraction $\left(1-\alpha u_{n,t}^{f_R}\right)(1-d_t)$ of firms), or if it does radical R&D, but also loses its contestable product (which is the case for a fraction $\alpha u_{n,t}^{f_R} d_t$ of firms).

(c) It may produce $n-1$ products.[24] This happens if it does not do radical R&D and loses its contestable product, which is the case for a fraction $\left(1-\alpha u_{n,t}^{f_R}\right) d_t$ of firms.

Potential entrants cannot lose products by definition, and a fraction $\alpha u_{0,t}^{f_R}$ of them enters each period. Equation (1.15) summarizes the changes in the size distribution between periods t and $t+1$.

$$\begin{aligned} & m_{1,t+1} = \alpha u_{0,t}^{f_R} m_{0,t} + \left(\left(1-\alpha u_{1,t}^{f_R}\right)(1-d_t) + \alpha u_{1,t}^{f_R} d_t\right) m_{1,t} + \left(1-\alpha u_{2,t}^{f_R}\right) d_t m_{2,t} \\ \forall n \geq 2, \quad & m_{n,t+1} = \alpha u_{n-1,t}^{f_R}(1-d_t) m_{n-1,t} + \left(\left(1-\alpha u_{n,t}^{f_R}\right)(1-d_t) + \alpha u_{n,t}^{f_R} d_t\right) m_{n,t} \\ & \qquad + \left(1-\alpha u_{n+1,t}^{f_R}\right) d_t m_{n+1,t} \end{aligned} . \tag{1.15}$$

The destruction rate d_t is endogenous. It is equal to the ratio between the mass of radical innovations and the mass of incumbent firms:[25]

$$d_t = \frac{R_t}{m_t}. \tag{1.16}$$

This completes the description of equilibrium. In the next section, I briefly analyse the model's balanced growth path, which provides a useful starting point before analysing the dynamic effect of aggregate shocks.

[24]For a firm with $n=1$ product in period t, this means exit.

[25]Parameter values need to be restricted such that in every period, $d_t \leq 1$. It can easily be shown that $\alpha(1+\psi) < 1$ is a sufficient condition for this, and I assume from now on that it holds.

1.3.2 The balanced growth path

I define the balanced growth path as the model's solution when financing conditions are normal and aggregate demand is high in every period t.
On the balanced growth path, all firms innovate up to their full innovation capacity. Thus, a fraction α of each size group of firms does a radical innovation every period, the destruction rate is constant, and there is a unique invariant firm size distribution, given by[26]

$$\forall n \geq 1,\ \ m_n = \left(\frac{\psi}{(1-\alpha)(1+\psi)}\right)^2 \left(\frac{1-\alpha(1+\psi)}{(1-\alpha)(1+\psi)}\right)^{n-1}, \tag{1.17}$$

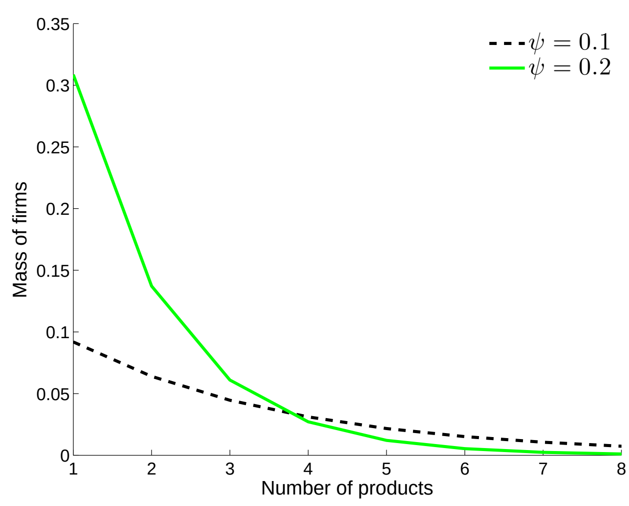
$$m_0 = \frac{\psi^2}{(1-\alpha)(1+\psi)}. \tag{1.18}$$

When the initial firm size distribution in period $t=0$ is the invariant one, all aggregate variables grow at a constant rate throughout. Using Equations (1.10) and (1.11), it is easy to show that innovation masses are constant and hold $R = \frac{\alpha\psi}{1-\alpha}$ and $I = \frac{1-\alpha(1+\psi)}{(1-\alpha)(1+\psi)}$. Therefore,

$$\forall t \in \mathbb{N},\ \frac{Y_{t+1}}{Y_t} = \frac{A_{t+1}}{A_t} = \exp\left(R\ln\gamma + I\ln\delta\right) \quad \text{and} \quad \frac{P_{t+1}}{P_t} = \exp\left(-\left(R\ln\gamma + I\ln\delta\right)\right). \tag{1.19}$$

Wages, labour supply and the efficiency of the labour allocation remain constant over time.

Figure 1.2: The invariant firm size distribution for different values of ψ



The shape of the invariant firm size distribution is determined by parameters, and in partic-

[26]The invariant size distribution is formally derived in Appendix Section 1.6.1.3.

ular, by the relative mass of potential entrants ψ. An increase in ψ shifts the distribution to the left.[27] This is intuitive: a greater mass of potential entrants increases actual entry and accelerates creative destruction. Incumbents are less likely to grow large, and a greater share of products is produced by entrants or small firms. Figure 1.2 illustrates this by plotting the invariant firm size distribution for $\alpha = 0.7$ and two different values of ψ.

As small firms have a higher relative innovation capacity than large ones, differences in the size distribution are related to differences in growth rates. In particular, an industry which for structural reasons (for example, low entry barriers) has a high value of ψ has both a left-skewed size distribution and a high balanced growth rate of productivity. In the next section, I analyse how the size distribution and industry productivity growth react to aggregate shocks.

1.3.3 Impulse responses to financial and aggregate demand shocks

To analyse the effects of an aggregate shock, I first look at impulse responses. That is, I assume that the initial firm size distribution is the invariant one, and that financing conditions are normal and aggregate demand is high from period 0 to period $T-1$ and from period $T+1$ to $+\infty$. In the crisis period T, however, there is a financial and/or aggregate demand shock.

1.3.3.1 Impulse responses to a financial shock

I assume first that there are crisis financing conditions in period T, but aggregate demand remains high. Figure 1.3 shows the impulse responses to this financial shock. The vertical line in Panels 1, 2 and 4 indicates the period in which the shock hits.

Panel 1 shows that the financial shock lowers the masses of radical and incremental innovations on impact, as potential entrants and a fraction of small incumbents do not have sufficient cash to finance R&D.[28] Lower innovation leads to a drop in productivity growth (see Panel 2) and to a permanent loss of industry output. This mechanism has been repeatedly highlighted by representative firm models. However, my model yields an additional, novel amplification and persistence effect that runs through the size distribution of innovating firms.

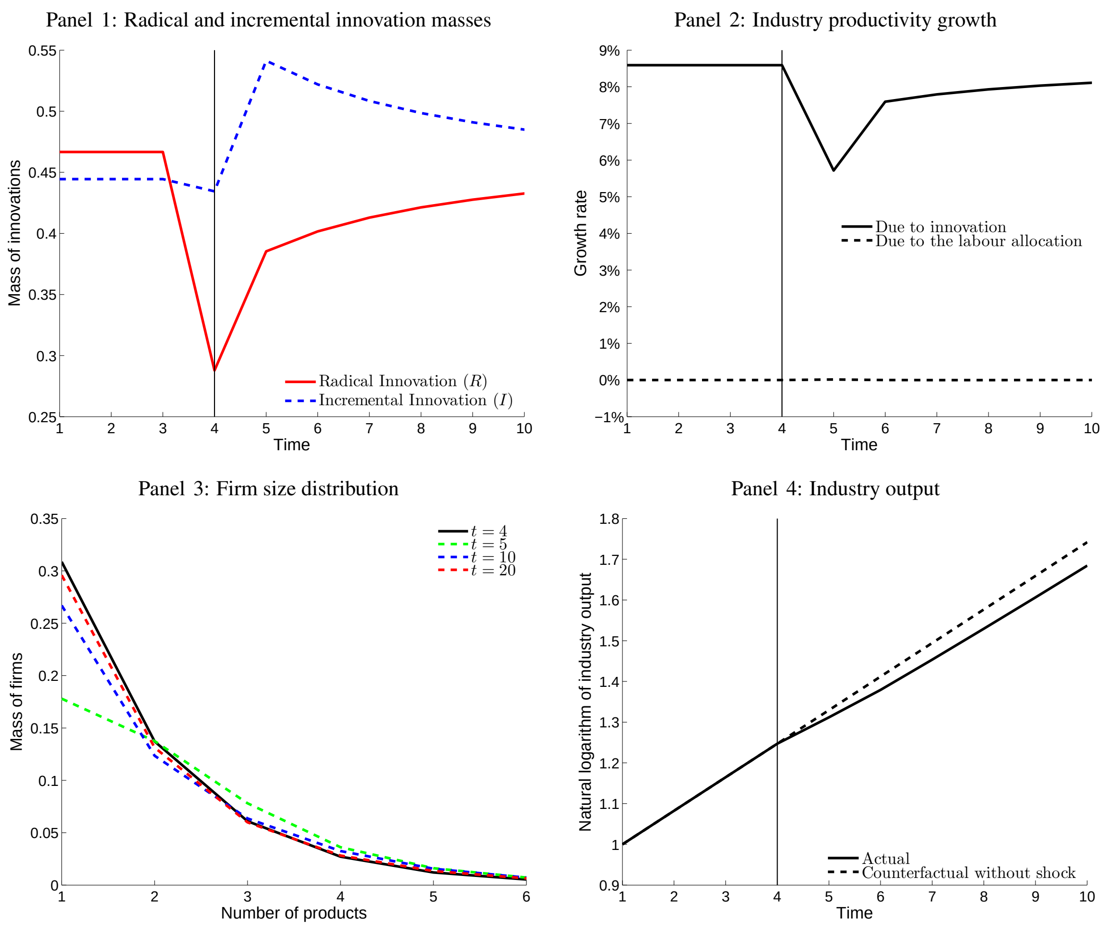
As the financial shock disrupts radical innovation, it keeps potential entrants and a fraction of small firms from expanding and overtaking the products of large firms. Therefore, the firm size distribution shifts to the right, as shown in Panel 3. As small firms' innovation capacity is relatively higher than the one of large firms, this implies that productivity growth remains

[27] See Appendix Section 1.6.1.3 for a proof of this claim.

[28] I determine the fractions $u^f_{n,t}$ of unconstrained firms numerically. Details are provided in Appendix Section 1.6.1.4.

below its balanced growth path level in the period after the shock, even though financial constraints completely disappear. Furthermore, the composition of innovation changes: while radical innovation is depressed (as the overall mass of firms is below its balanced growth path level), incremental innovation is stimulated (as the rightward shift in the size distribution increases the fraction of non-contestable products).

Figure 1.3: Impulse responses to a financial shock



Notes: The crisis shock hits at $T = 4$. Parameter values used for computing impulse responses are given in Table 1.6 in the Appendix.

The shift in the size distribution (and therefore the fall in productivity growth[29]) is persistent over time: even as entry and creative destruction resume in the aftermath of the shock, innovation is stochastic and it therefore takes some time until large firms lose the "excessive" products they were able to maintain in the crisis period. The persistent fall in productivity

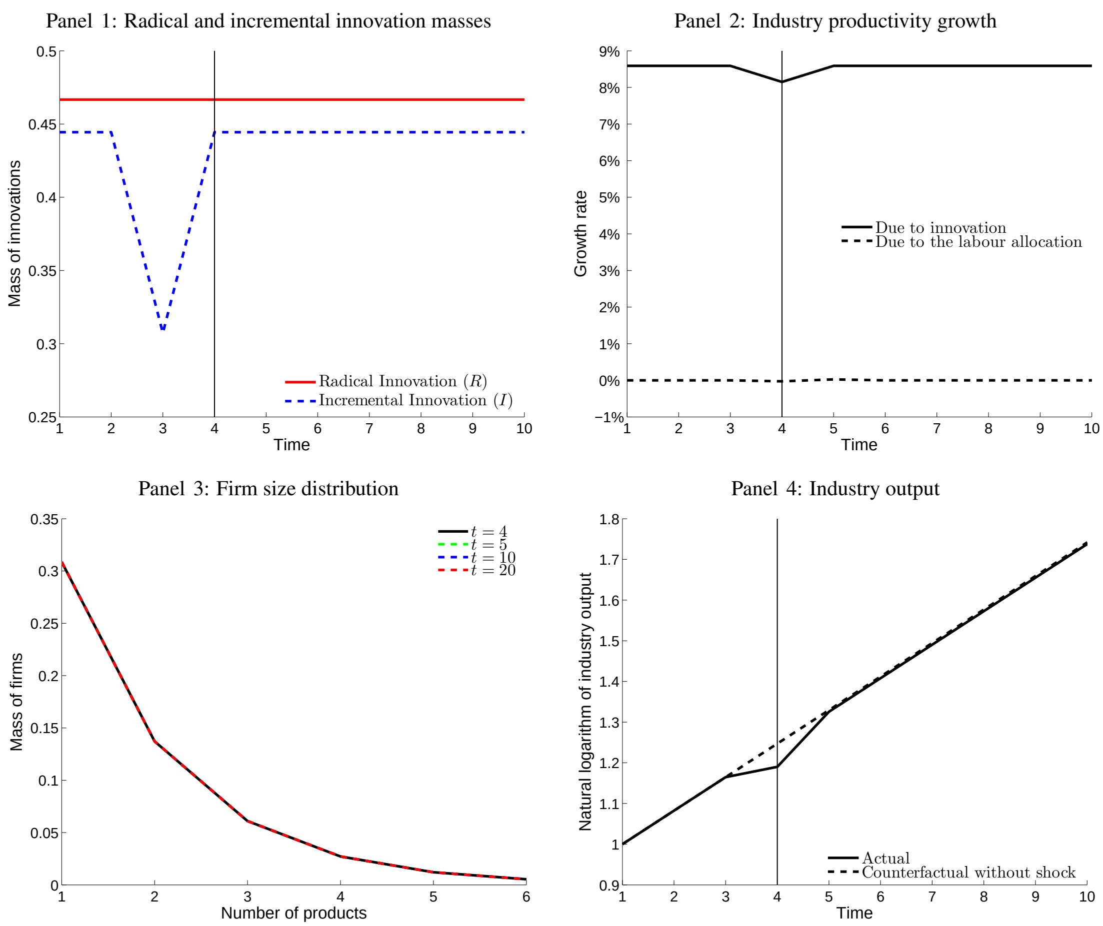
[29]Productivity growth is almost entirely explained by the direct effect of innovation. As shown in Panel 2, changes in the efficiency of the labour allocation are negligible.

growth greatly increases the permanent loss of industry output. This is illustrated in Panel 4, which plots the actual path of output against a counterfactual one that would have prevailed in the absence of any shock. The graph shows that as long as productivity growth is depressed, actual output keeps falling further below counterfactual output.

1.3.3.2 Impulse responses to an aggregate demand shock

I now assume aggregate demand falls to its low level S_L in period T, but financing conditions remain normal. The impulse responses to this shock are shown in Figure 1.4.

Figure 1.4: Impulse responses to an aggregate demand shock



Notes: See Figure 1.3.

Panel 1 shows that radical innovation is unaffected by the shock, as it is profitable even with low aggregate demand. Incremental innovation, however, is not profitable any more for small firms with few products. Therefore, the mass of incremental innovations falls (one period before the shock, as firms can forecast aggregate demand). This fall is transitory: once aggregate demand returns to its balanced growth path level, incremental innovation and industry productivity growth also jump back. The main reason why an aggregate demand shock has no persistent effect on innovation and productivity growth is that it does not affect the firm size distribution (see Panel 3).[30] Thus, it only causes a small permanent loss of industry output through the fall of innovation on impact.

Finally, Panel 4 shows that the aggregate demand shock also has a direct effect on output. Falling demand reduces wages, and this reduces labour supply and output. This effect is reversed in the period after the shock, when demand, wages and labour supply increase again.

1.3.3.3 Impulse responses to a joint financial and aggregate demand shock

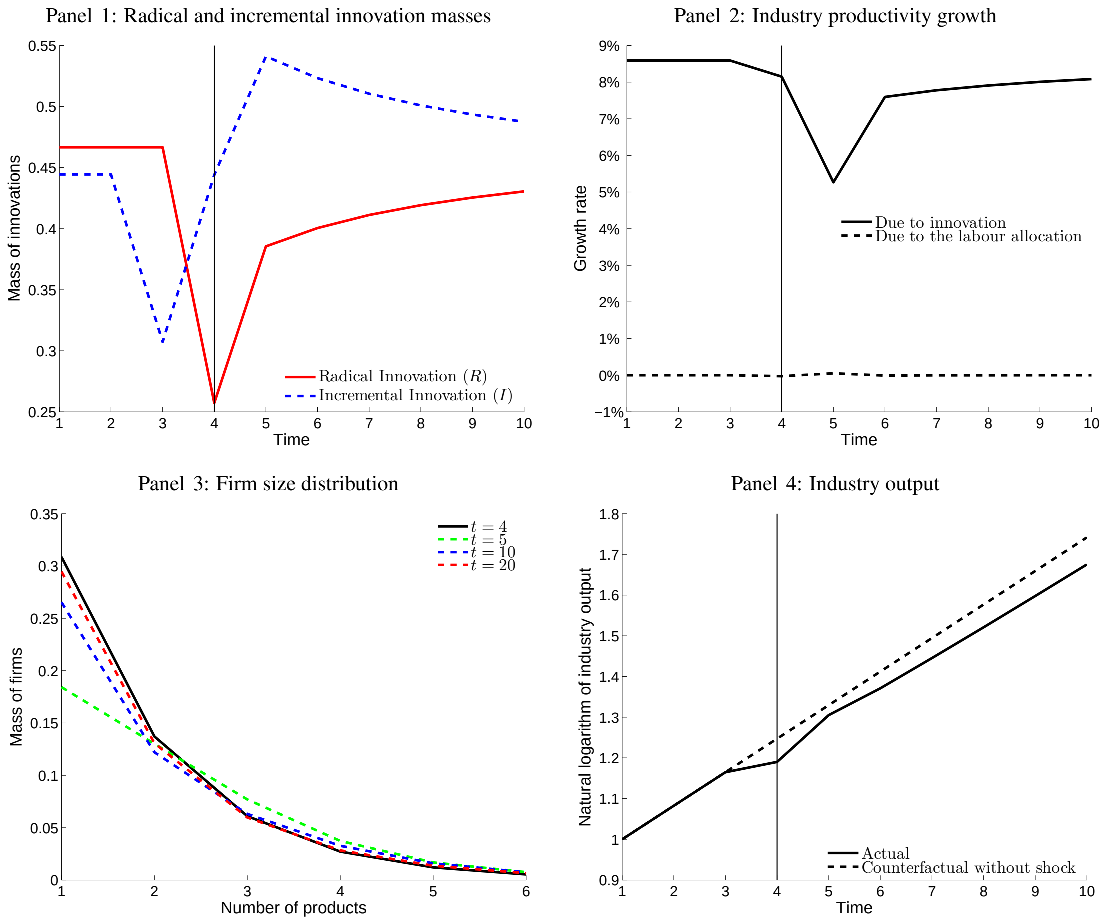
Figure 1.5 shows impulse responses when both a financial and an aggregate demand shock hit in period T. Therefore, it may be interpreted as showing an industry's response to an event such as the 2007-2009 Great Recession.

A joint shock combines the effects of the two shocks taken separately. On impact, incremental innovation falls in anticipation of lower aggregate demand, and radical innovation falls because financial constraints are binding for entrants and a fraction of small firms. Lower radical innovation induces a persistent rightward shift in the firm size distribution which persistently depresses productivity growth. Therefore, even after recovering from the short-run effect of the aggregate demand shock, output growth is depressed (see Panel 4).

It is worth noting that although financial constraints are important in my model, their role is very distinct from the classic financial accelerator mechanism, as described in Bernanke and Gertler (1989) or, in a model with heterogeneous firms and creative destruction, Caballero and Hammour (2005). In these models, financial constraints are permanent. Therefore, a transitory aggregate demand (or productivity) shock, which lowers firms' profits, makes financial constraints more likely to bind in the following periods and persistently depresses investment. In my analysis, the fall in profits does reinforce the effect of the financial shock by increasing the mass of constrained firms in the crisis period. However, once financing conditions have normalised, the fall in profits becomes irrelevant (as firms do not need them any more to finance R&D), and all persistence is due to the shift in the size distribution. In sum, my model emphasizes a persistence mechanism which is both distinct from and complementary to the financial accelerator.

[30] Aggregate demand shocks do not affect the size distribution in my model because their impact is the same for all firms. Indeed, even if an aggregate demand shock would make radical innovation unprofitable, the size distribution would not change, as all firms would stop radical innovation.

Figure 1.5: Impulse responses to a joint aggregate demand and financial shock



Notes: See Figure 1.3.

The main conclusion of this section is summarized in Prediction 2.

Prediction 2. A financial shock leads to a persistent rightward shift in the size distribution of innovating firms, and therefore persistently depresses productivity growth.[31]

In the next section, I compare the reaction of industries with different characteristics to a random sequence of aggregate demand and financial shocks.

1.3.4 Growth, volatility and the firm size distribution

The shape of the firm size distribution in my model is affected by exogenous parameters such as α and ψ, which can be seen as industry characteristics. As noted earlier, an industry with a

[31] This mechanism is a priori not limited to financial shocks. It equally applies to any other transitory shock which disrupts radical innovation in small firms more than in large ones.

higher value of ψ (which could be due to lower barriers to entry) has both a more left-skewed size distribution and higher productivity growth on the balanced growth path. However, the impulse response analysis from the previous section also suggests that productivity growth in this industry reacts more intensely to a negative aggregate shock.[32] These considerations lead to Prediction 3.

Prediction 3. There is a positive correlation between the share of innovations done by small firms, average productivity growth and volatility at the industry level.

Prediction 3 implicitly assumes that all industry characteristics which jointly affect the size distribution and productivity growth have been controlled for.[33] Furthermore, it presumes that industries are hit by the same sequence of shocks, and that these shocks are not too frequent. Indeed, if an industry with a very left-skewed firm size distribution were constantly hit by financial shocks, its average productivity growth could be lowered decisively. However, as financial crises are rare events, assuming infrequent shocks appears reasonable.

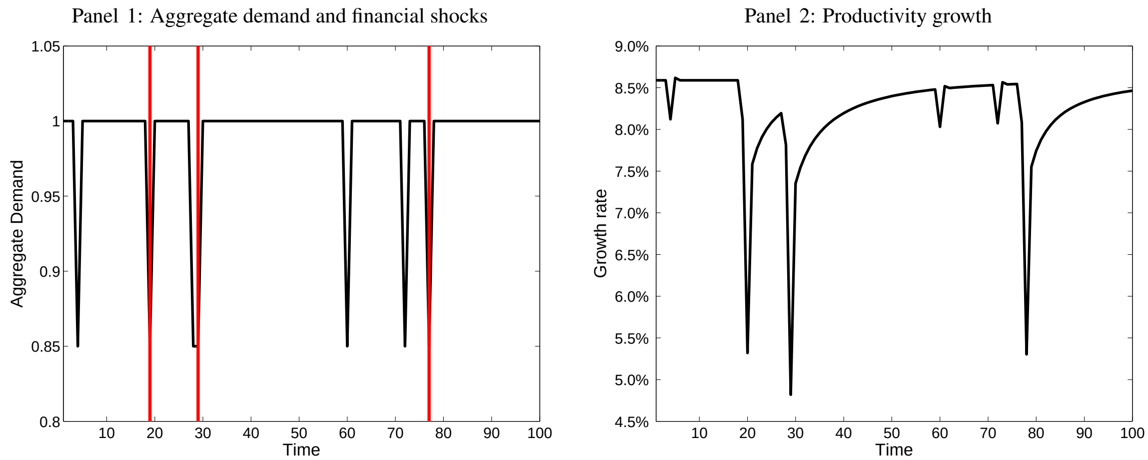
I illustrate Prediction 3 by simulating my model for a random sequence of shocks. I assume that transition probabilities for the aggregate demand Markov chain are $p_{HL} = 0.1$ and $p_{LH} = 0.8$. Financial shocks are correlated with aggregate demand: crisis financing conditions occur with probability $p_F = 0.25$ for any period with low aggregate demand and do not occur if aggregate demand is high.[34] I then draw a sequence of shocks for 100 periods and calculate the model's solution for it (assuming the industry starts on the balanced growth path). Panel 1 of Figure 1.6 provides an example for a sequence of shocks. The solid line plots aggregate demand, while vertical bars mark periods with financial shocks. Panel 2 shows the corresponding path of productivity growth. It is fully in line with the intuitions developed in Section 1.3.3: aggregate demand shocks generate small and transitory falls in productivity growth, while joint financial and aggregate demand shocks generate larger and persistent falls.

[32]This is easy to see for a financial shock, as the industry has a greater mass of constrained firms. For an aggregate demand shock, the analysis is less clear-cut. Indeed, firms that abandon incremental innovation after an aggregate demand shock are small, but they are not the smallest firms (firms with only one product, for instance, never do incremental innovation). However, if a leftward shift in the firm size distribution does not only increase the mass of firms with one product, but also the masses of slightly larger firms, it makes the industry more sensitive to an aggregate demand shock.

[33]For instance, in my model, the parameter α (the probability of a firm to receive a radical innovation opportunity) affects the size distribution, but also has a direct effect on the mass of possible innovations. These and other characteristics may blur the relationship suggested by Prediction 3 if they vary across industries.

[34]I did not need to specify the laws of motion for aggregate demand and financing conditions before, as they are irrelevant for firms' R&D policy functions.

Figure 1.6: An example for a simulation run



I use these simulations to compare two industries which are identical except for their value of ψ. I draw 500 different runs of aggregate shocks and calculate for each run the average and the standard deviation of productivity and output growth over the 100 periods of the simulation for both industries. Table 1.1 reports the average of both indicators over the 500 runs. It shows that both the average and the standard deviation of productivity growth rates are higher in the high ψ industry, which also has on average a more left-skewed size distribution. Industry output growth and volatility are positively correlated as well, as they are mainly driven by productivity. However, volatility differences are smaller, as output volatility is also directly affected by aggregate demand shocks, which are identical for both industries.

Table 1.1: Simulation results

		$\psi = 0.1$	$\psi = 0.2$
Average production share of small firms		20.3%	54.6%
Productivity	Average growth rate	5.5%	8.2%
	Standard deviation	0.2%	0.6%
Output	Average growth rate	5.5%	8.2%
	Standard deviation	2.38%	2.41%

Notes: Small firms are defined as firms with two products or less. Their production share in period t equals $m_{1,t} + 2m_{2,t}$. Each simulation run had 100 periods. Of these, there were on average 11.2 periods with negative demand shocks, and 2.9 periods with financial shocks. All parameter values other than ψ are the same as those used previously, and given in Table 1.6 in the Appendix.

1.3.5 Robustness to different assumptions

Before concluding on the model, it is useful to shortly discuss the role of some assumptions that have not been highlighted so far.

A line of simplifying assumptions exclude aspects which are beyond the scope of my analysis. For instance, considering a partial equilibrium model, and therefore an exogenous aggregate demand process, prevents feedback effects from wages and dividends to aggregate demand. These feedback effects may lead to the clustering of innovations (as has been pointed out by Shleifer (1986)), but they do not create differences between small and large firms. Likewise, my assumptions on the payoffs of R&D investment exclude precautionary effects due to the anticipation of potentially binding financial constraints in the future. This does not imply that these effects are negligible in practice.[35] However, as they do not affect the qualitative predictions of my model, I abstract from them in order to simplify the firms' problem.

Finally, two assumptions play some role in the persistent rightward shift of the firm size distribution after a financial shock, which is key for my model's predictions.
First, I have assumed that the mass of potential entrants is a linear function of the mass of incumbents (see Equation (1.2)). Thus, the fall in the mass of incumbents after the shock also reduces the mass of potential entrants. This reinforces the persistence of the size distribution's rightward shift, but it is not crucial. Indeed, the rightward shift could only be undone if entry would exceed its balanced growth path level in the aftermath of the shock.[36] However, there are no reasons for such an overshooting in my model, not even if the mass of potential entrants was pinned down by a free entry condition.[37]
Second, an important implicit assumption preventing entry from overshooting is that radical innovations cannot be stored: a firm that foregoes a radical innovation in period t cannot realise the same radical innovation in period $t+1$. This assumption can be interpreted as reflecting obstacles to postponing innovation.[38] However, even if it were possible to store

[35] Pérez (2012) and Khan and Thomas (2013) show how they may affect investment dynamics.

[36] Thus, all results are qualitatively unchanged if I assume that the absolute mass of potential entrants is fixed (that is, if I replace Equation (1.2) with $m_{0,t} = \psi$).

[37] After a financial shock (and as long as the latter does not affect demand expectations), potential entrants face the same aggregate conditions than on the balanced growth path, but the mass of incumbents is lower and therefore, the destruction rate d_t is higher for every given mass of entrants (as there are fewer contestable products). Thus, with free entry, entry would actually be persistently depressed after a financial shock.
A free entry condition would, however, change the impulse responses to an aggregate demand shock. As the shock lowers the value of entry, it would now lead to a fall in entry and a persistent rightward shift in the firm size distribution, generating a persistent fall in productivity growth.

[38] For example, small firms may lose their skilled employees if they are prevented from realising their innovation opportunities, and it may be difficult to hire them back later.

radical innovations, the size distribution would still be shifted to the right at least in the period immediately after the shock, lowering productivity growth in that period. Moreover, it is unlikely that all delayed innovations could actually be implemented, as there may be overlaps between them and the new innovations created in period $t+1$. For example, suppose a firm could not increase the frontier productivity of some product j from $a(j)$ to $\gamma a(j)$ in period t. In period $t+1$, another firm may be able to do the same improvement, and as only one firm can actually innovate, the delayed firm may not be able to expand after all. Thus, allowing for innovation postponement cannot undo the persistent shift in the size distribution.

Summing up, I have proposed a new model of R&D fluctuations in which the size distribution of innovating firms plays a key role. Under the assumption that small firms' innovation capacity is relatively higher than the one of large firms, my model generates three main predictions. First, small firms' R&D reacts more to adverse aggregate shocks (Prediction 1). This explains that a financial shock triggers a persistent rightward shift in the firm size distribution, which persistently depresses productivity growth (Prediction 2). It also suggests that all else equal, industries with a more left-skewed size distribution of innovating firms have both higher productivity growth and volatility (Prediction 3). The remainder of the chapter analyses the empirical evidence for these predictions.

1.4 Empirical evidence

1.4.1 Small and large German firms' R&D during the Great Recession

Prediction 1, which states that small firms reduce R&D more than large ones after a negative aggregate shock, is in line with several empirical studies comparing small and large firms' cyclical behaviour. Gertler and Gilchrist (1994) have shown that small manufacturing firms' inventories fall after a negative monetary policy shock, while those of large firms increase. More recently, Fort et al. (2013) have argued that the job creation rate of young and small firms falls more than the one of old and large ones during economic downturns. This differential was particularly large during the 2007-2009 Great Recession.[39] Many studies also document a large fall in entry and firm creation during the Great Recession (OECD (2012), Sedláček and Sterk (2013), Klapper et al. (2014)).

However, only a small number of studies have explicitly considered R&D cyclicality by firm

[39] Krueger and Charnes (2011) and Siemer (2014) also claim that small firms' employment suffered disproportionately during the Great Recession. Moscarini and Postel-Vinay (2012) take a dissenting stand.

size.[40] In the remainder of this section, I contribute to fill this gap by using a German firm-level dataset to compare the response of small and large firms' R&D to the Great Recession, which can be considered as a joint aggregate demand and financial shock.
My dataset comes from the Mannheim Innovation Panel (MIP), an annual survey carried out by the Centre for European Economic Research (ZEW).[41] The survey targets a representative sample of German firms with 5 employees or more, in a broad range of innovating sectors. It asks firms about R&D, total innovation spending (including R&D, implementation and marketing expenses) and some other variables such as sales or employment. The dataset is an unbalanced panel covering the period 1999-2009. A more complete description can be found in Appendix Section 1.6.2.1.

The aggregate evolution of innovation spending in Germany during the Great Recession is in line with my model's Prediction 1: total innovation spending by firms with 50 employees or less fell by 12% from 2007 to 2009, while total innovation spending by firms with 1000 employees or more only fell by 2% (Rammer et al. (2010)). However, this does not automatically imply that small firms have reacted more intensely to the economic and financial crisis. Idiosyncratic or industry-specific sales shocks may have hit small firms harder, and this may explain their larger R&D reductions.
In Table 1.2, I therefore compare the evolution of R&D in the median small firm (with 50 employees or less) with the one in the median large firm (with more than 500 employees) in the MIP dataset, conditioning on the sign of the sales change in order to compare firms affected with similar intensity by the crisis.[42] The table shows that among firms which saw their sales fall during the Great Recession, small firms cut R&D investment considerably more than large ones (especially in 2009). Among firms with increasing sales, small and large firms increased R&D more or less in the same proportions.[43] Furthermore, note that

[40]Paunov (2012) shows that in a panel of firms from eight Latin American countries, young firms were more likely to abandon innovation investment during the Great Recession. She does not find an independent role for size. Hall (2011), on the other hand, presents preliminary evidence for small firms' R&D investment falling more during the same period in the United States. Finally, Aghion et al. (2012) show that R&D investment reacts more to negative sales shocks in financially constrained firms. However, even though most conventional measures suggest that constrained firms are on average smaller than unconstrained ones (Farre-Mensa and Ljungqvist (2013)), Aghion et al.'s proxy for financial constraints is uncorrelated with firm size.

[41]I am grateful to Sandra Gottschalk, Bettina Peters and Christian Rammer of the ZEW for their help with this data.

[42]I report R&D rather than innovation spending because the former is more commonly used in the literature. However, all my results are unchanged when I consider innovation spending instead (see Appendix Section 1.6.2.2). I set the cut-off for large firms to 500 rather than to 1000 employees in order to have more observations for the group of large firms.

[43]Table 1.2 therefore also confirms that R&D is procyclical: in general, firms increase R&D if their sales increase and lower R&D if their sales fall.

74% of the large firms in my sample had falling sales in 2009, while that was only true for 55% of the small ones. Against this backdrop (probably due to the greater exposition of large firms to the collapse in world trade), the divergence in aggregate trends is even more remarkable.

Table 1.2: R&D and sales fluctuations by firm size during the Great Recession

	2008			**2009**		
Median rate of change	**R&D**	**Sales**	**Obs.**	**R&D**	**Sales**	**Obs.**
Firms with decreasing sales						
Small (50 or less emp.)	-8.00	-10.83	87	-36.07	-18.70	201
Large (more than 500 emp.)	0.00	-6.23	45	-7.47	-21.13	76
Firms with increasing sales						
Small (50 or less emp.)	14.59	11.76	196	13.61	10.64	162
Large (more than 500 emp.)	10.97	7.84	76	10.52	6.81	27

Notes: Statistics are computed on the sample of firms which have observations for at least two consecutive years. They exclude firms with R&D investment smaller than 10.000€ in both years, in order not to overweight changes in trivial amounts. Results are unchanged with a threshold of 20.000€. Rates of change for a variable x are computed as $100\frac{x_t - x_{t-1}}{0.5(x_t + x_{t-1})}$, following Davis et al. (1998). Employment refers to the first year of observation (i.e., 2007 for the rate of change between 2007 and 2008). Firms with unchanged sales are included in the increasing sales category.

Table 1.2 suggests that small firms reacted more intensely to the mix of idiosyncratic and common sales shocks received during the Great Recession. As a result, a gap between the R&D intensity (the ratio of R&D to sales) of small and large firms opened up, in disfavour of small firms. Table 1.2 however does not show whether this gap is statistically significant.

I therefore analyse a series of fixed effect regressions which take firms' R&D intensity as their dependent variable. Precisely, I estimate

$$\begin{aligned}\frac{\mathrm{RD}_{it}}{\mathrm{Sales}_{it}} = \quad & \alpha_i + \sum_{\tau=1999}^{2009} \beta_\tau D_t^\tau + \beta_{SE}\mathrm{SE}_{it} + \sum_{\tau=1999}^{2008} \beta_{SE,\tau}\mathrm{SE}_{it} D_t^\tau \\ & + \beta_{ME}\mathrm{ME}_{it} + \sum_{\tau=1999}^{2008} \beta_{ME,\tau}\mathrm{ME}_{it} D_t^\tau + \varepsilon_{it}\end{aligned}, \tag{1.20}$$

where α_i is a firm fixed effect controlling for time-invariant factors affecting R&D intensity. SE_{it} is a dummy for firm i being small, that is, having 50 employees or less in the last period of observation before t,[44] and ME_{it} is a dummy for medium-size firms, with between 51 and 500 employees.[45] D_t^τ is a time dummy for the year τ. I winsorize outliers for R&D intensity (5% of the highest non-zero values for every year) and cluster standard errors at the firm and at the industry-year level. Finally, as many firms never spend on R&D, I estimate Equation (1.20) on the sample of firms which have at least one observation with non-zero R&D.[46]

Table 1.3: Fixed effect regression results

Dependent variable: R&D intensity

Sample	All firms	Res.-int. manuf.
$\text{SE}D^{1999}$	1.35 (0.49)***	2.15 (0.92)**
$\text{SE}D^{2000}$	1.04 (0.48)**	3.27 (0.80)***
$\text{SE}D^{2002}$	2.45 (0.46)***	4.38 (0.80)***
$\text{SE}D^{2003}$	1.39 (0.41)***	3.11 (0.69)***
$\text{SE}D^{2004}$	2.24 (0.44)***	3.20 (0.88)***
$\text{SE}D^{2005}$	1.15 (0.41)***	1.98 (0.66)***
$\text{SE}D^{2006}$	0.93 (0.38)**	2.00 (0.78)**
$\text{SE}D^{2007}$	0.91 (0.38)**	2.49 (0.67)***
$\text{SE}D^{2008}$	0.84 (0.37)**	1.30 (0.60)**
Observations	15206	5729

Notes: *** significant at 1%, ** significant at 5%, * significant at 10%. Estimates for the coefficients on year dummies, the uninteracted small firm dummy and medium-size firms are not shown. The R&D Services industry (NACE Code 73) is left out. There are no estimates for 2001, as the survey for this year did not contain a question on R&D. Cluster-robust standard errors are given in parentheses. There are 4234 clusters at the firm and 343 at the industry-year level in the full sample, and 1681 firm and 80 industry-year clusters in the research-intensive manufacturing sample (industries are classified according to the two-digit NACE Rev. 1.1 classification).

My baseline results are shown in the left column of Table 1.3. To interpret them, note that

[44] In a firm's first period of observation, the dummy's value is determined with respect to present employment.

[45] The presence of firm fixed effects implies that the uninteracted dummies for small and medium-size firms are only identified by firms that switch categories over time. My results are unchanged when I leave these variables out and define permanent dummies for firm size, determined by a firm's employment in the first year of observation (see Appendix Section 1.6.2.2).

[46] This is common practice in firm-level R&D studies (see, for instance, Aghion et al. (2012)).

coefficients are identified by within-firm variation and that Equation (1.20) omits the group of large firms (with more than 500 employees) and the interaction between the year dummy for 2009 and the small firm dummy. Thus, the positive and significant estimates for the coefficients of the interactions for years prior to 2009 and small firm dummies indicate that in 2009, the R&D intensity of small firms fell significantly more (with respect to the within-firm average, and compared to large firms) than in other years. The difference amounts to around 0.9 percentage points.[47] Estimates for medium-size firms (not shown) have the same signs than those for small firms, but are in general insignificant.
Using R&D intensity as a dependent variable controls for industry-level differences in sales shocks, but does not control for potential industry-level differences in the elasticity of R&D with respect to sales.[48] Therefore, as a robustness check, I test whether findings hold in a subset of "research-intensive manufacturing industries",[49] which account for almost two thirds of German R&D spending. Results for this subsample are shown in the right column of Table 1.3, and are in line with the full sample. I also estimated Equation (1.20) with a full set of industry-year dummies instead of only year dummies. Results (shown in Appendix Section 1.6.2.2) are similar to the ones shown in Table 1.3. Finally, I have considered specifying the dependent variable in natural logarithms or omitting firms which are observed only for two years. In each case, results were qualitatively unchanged.[50]

In sum, with respect to the sales shocks received, small German firms on average reduced R&D more than large ones during the Great Recession. This empirical support for Prediction 1, together with the evidence for the higher relative innovation capacity of small firms discussed in Section 1.2.4, makes my model's aggregate and industry-level predictions more plausible. However, aggregate predictions can also be examined directly. I turn to this in the two remaining sections.

1.4.2 Productivity growth after the Great Recession

My model predicts that a financial shock shifts the size distribution of innovating firms persistently to the right and thereby depresses productivity growth (Prediction 2). This is consis-

[47] The gap is also economically significant, as R&D intensities rarely exceed 5%. However, these results obviously do not prove that the Great Recession is the reason for this evolution.

[48] This is the reason for which I exclude R&D Services (NACE Code 73) from the baseline estimation shown in Table 1.3. This industry is composed almost exclusively by small firms, whose R&D intensities fall substantially in 2008 and 2009. While in line with my argument, it would be suspicious if this would drive my results, as they could then be due to a particularity of the R&D services industry.

[49] This category contains Chemicals, Machinery, Electrical and Optical Equipment and Transport Equipment.

[50] Results for these last two specifications are available upon request.

tent with the experience of several OECD countries after the Great Recession.[51] In the United States, slow productivity growth during the recovery has recently received a lot of attention.[52] However, it is unclear whether this is due to the financial shocks of the Great Recession. Indeed, Fernald (2014) argues that the slowdown started already in the mid-2000s and is due to the fading out of the IT-driven productivity improvement wave of the 1990s. In other countries, the case for a slowdown induced by financial disruptions is stronger. Ball (2014) shows the fall in the growth rate of potential output after the Great Recession (driven to an important extent by productivity growth) was largest in the Eurozone periphery (Spain, Portugal, Greece, Ireland), where financial disruptions have arguably been most severe.[53]
Are these slowdowns in productivity growth due to a rightward shift in the size distribution of innovating firms? It is difficult to assess this claim, as there is little information on the distribution's evolution over time. Evidence for Germany and the United States is consistent with a (small) rightward shift. In Germany, the share of small firms (with less than 50 employees) in the population of firms reporting continuous or occasional R&D activities fell from 68.4% in 2008 to 66.3% in 2010 in manufacturing, and from 83.3% to 81.9% in services.[54] In the United States, the share of firms with less than 500 employees in aggregate R&D fell from 21.2% in 2008 to 19.3% in 2010.[55] In sum, aggregate evolutions are consistent with my model, even though they obviously do not provide causal evidence for its mechanisms.

1.4.3 Growth, volatility and the size distribution of innovating firms

Finally, my model predicts that, all else equal, there is a positive correlation between the innovation share of small firms, productivity growth and volatility across industries (Prediction 3). In this section, I show that across manufacturing industries in the United States (which have experienced similar financial and aggregate demand shocks), the average and the standard deviation of productivity growth are indeed positively correlated. The scarce data

[51] My model is set up in partial equilibrium, so its predictions should ideally be assessed with industry-level data. However, industry-level time series data on the size distribution of innovating firms is scarce, and the mechanisms driving Prediction 2 are likely to extend to general equilibrium. I therefore analyse country-level data in this section.

[52] See for instance the cover story of The Economist, July 19th, 2014, entitled "*America's lost oomph*".

[53] According to Ball's estimates, potential growth fell by 0.34 percentage points (pp) in the United States, but by 2.64 pp in Spain, 1.34 pp in Portugal, 4.11 pp in Greece and 4.82 pp in Ireland.

[54] These figures are computed from ZEW estimates based on the MIP (Rammer et al. (2010, 2012), Page 16).

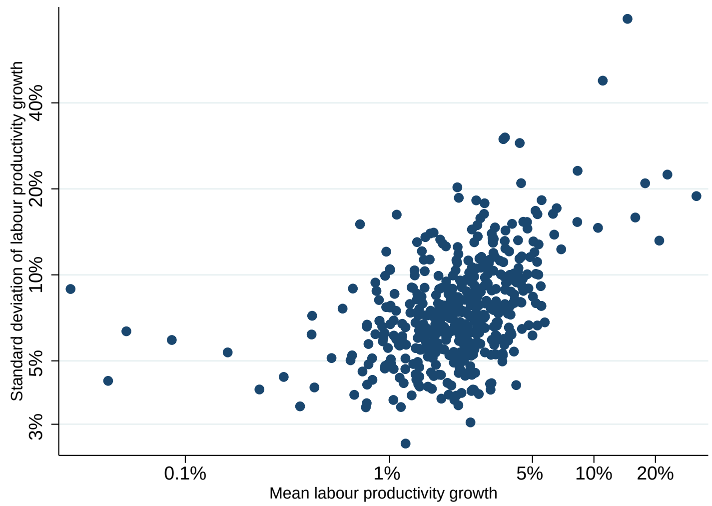
[55] These figures are computed using data from the Business Research and Development and Innovation Survey (BRDIS) carried out by the National Science Foundation (NSF). I thank Raymond M. Wolfe for his help with this data. It is unfortunately difficult to compare these figures to pre-crisis years, as the BRDIS survey was introduced only in 2008 and several changes with respect to its predecessor may have created discontinuities.

on the size distribution of innovating firms suggest this new stylized fact[56] may be linked to the innovation share of small firms, as predicted by my model.

1.4.3.1 Productivity growth and volatility across industries

To analyse industry-level productivity growth, I use the NBER-CES Manufacturing Database (Becker et al. (2013)), and following Acemoğlu et al. (2014), I measure an industry's productivity as labour productivity, that is, real shipments per employee.[57]

Figure 1.7: Productivity growth and volatility in US manufacturing industries, 1972-2009



Notes: Both axes have a logarithmic scale. Every dot corresponds to one NAICS six-digit industry.

Figure 1.7 plots, on a logarithmic scale, the average growth rate of labour productivity between 1972 and 2009 against the standard deviation of productivity growth rates, for all six-digit NAICS manufacturing industries. It indicates a strong positive correlation between both variables.

This positive correlation is robust to the introduction of control variables. In particular, I estimate the model

$$\ln sd_k^{LP} = \beta_0 + \beta_1 \ln g_k^{LP} + \beta_2 \text{Size}_k + \beta_3 \text{Capital Share}_k + \beta_4 sd_k^{CU} + \varepsilon_k, \tag{1.21}$$

[56]To the best of my knowledge, this correlation has not yet been documented in the literature. However, Imbs (2007) has shown that there is a positive correlation between the average and the standard deviation of output growth at the industry level, using panel data from 47 countries.

[57]Further details on the variables and measures used are given in Appendix Section 1.6.2.3.

where sd_k^{LP} stands for the standard deviation of labour productivity growth rates in industry k and g_k^{LP} for their average. I consider three control variables: industry size (measured by the natural logarithm of average value added over the period), the capital share (measured by the average capital share in the industry's production function[58]) and the standard deviation of capacity utilisation (sd_k^{CU}).[59] Table 1.4 shows that larger industries have lower volatility, while the other two control variables are insignificant. The correlation between productivity growth and volatility remains positive and strongly significant.

Table 1.4: Productivity growth and volatility: Regression evidence

Dependent variable: ln (Standard deviation of productivity growth)

ln (Average productivity growth)	0.268^{***}	0.282^{***}
	(0.040)	(0.039)
ln (Average value added)		-0.162^{***}
		(0.017)
St. Dev. Capacity Utilisation		0.567
		(0.637)
Capital share		0.141
		(0.206)
Constant	-1.552^{***}	-0.386^{*}
	(0.154)	(0.219)
R^2	0.206	0.362
Observations	457	457

Notes: *** significant at 1%, ** significant at 5%, * significant at 10%. Productivity is measured by real shipments per employee. Average value added of the industry is given in millions of dollars. Robust standard errors are given in parentheses.

Results are also unchanged when considering the measure of total factor productivity (TFP)

[58] Becker et al. (2013) describe the measurement of the capital share in more detail.

[59] The labour productivity measure used may overstate the volatility of productivity growth in industries with a volatile capacity utilisation rate. Data on capacity utilisation is provided monthly by the Federal Reserve (accessible at http://www.federalreserve.gov/RELEASES/G17/caputl.htm) for 22 manufacturing sectors. I take averages of the monthly data to obtain an annual time series, and assign to every industry in the NBER-CES database the standard deviation of annual capacity utilisation of the corresponding sector.

provided in the NBER-CES database (calculated as a Solow residual from a neoclassical production function) instead of labour productivity, omitting petrol industries, weighting observations by industry size or using the average yearly growth rate of productivity (i.e., the annualized growth rate of productivity between 1972 to 2009) instead of the average of annual growth rates.[60] I now briefly analyse whether the link between growth and volatility can be related to differences in the size distribution of innovating firms.

1.4.3.2 The role of the size distribution of innovating firms

Unfortunately, data on the size distribution of innovating firms at the industry level is scarce. Table 1.5 provides partial evidence for some US industries, showing the share of Small and Medium Enterprises (SMEs) in industry R&D (a proxy for small firms' share of innovation) as well as average productivity growth and volatility between 1978 and 2009.

Table 1.5: SME's share in R&D, productivity growth and volatility in selected industries

		SME share in R&D			**Productivity (1978-2009)**	
	NAICS code	**1997**	**2004**	**2010**	**Avg. Growth**	**Volatility**
Computer and Electronic Products	334	7%	13%	13%	10.6%	8.4%
Chemicals	325	3%	8%	10%	2.4%	4.4%
Transportation Equipment	336	0%	2%	5%	2.5%	5.0%
Computer Systems Design	5415	n.a.	n.a.	51%	4.9%	5.5%

Notes: SME shares in R&D are measured as shares of domestic R&D financed with company or other non-federal funds carried out by firms with 500 employees or less (except for 2010, where other non-federal funds are excluded). Data comes from the NSF. For 1997, the share for the Computer and Electronic Products industry is an average between the industries Office, computing, and accounting machines and Electrical Equipment. Productivity is calculated with the NBER-CES Manufacturing Database (as real shipments per employee) for the three manufacturing industries, and with the BEA Industry Accounts (as real value added per employee) for Computer Systems Design.

The table indicates that industries such as Computer and Electronic Products or Computer Systems Design, which have larger shares of R&D carried out by SMEs, also had a higher average and standard deviation of productivity growth. Traditional manufacturing industries such as the car industry (Transportation Equipment) and the chemical industry, where R&D

[60] These results are available upon request. Note, however, that both TFP and labour productivity are imperfect measures, as they confound the effect of mark-ups and physical productivity.

is dominated by large firms, have grown slower and fluctuated less.[61] My model suggests that this pattern may be due to industry differences in exogenous fundamentals (for example, barriers to entry associated with the production technology) which explain that small firms are more prevalent in some industries than in others. This, in turn, explains that these industries have both higher productivity growth and volatility.

1.5 Conclusions

The model developed in this chapter shows that the size distribution of innovating firms matters for the analysis of R&D fluctuations over the business cycle and across industries. Indeed, while small firms have a relatively higher innovation capacity than large ones, they also reduce R&D more after a negative aggregate shock. Therefore, financial shocks shift the size distribution of innovating firms persistently to the right, and this persistently lowers productivity growth. Furthermore, the share of small firms in innovation, productivity growth and volatility are positively correlated at the industry level. The empirical evidence is consistent with the model's predictions.

Exploring the quantitative implications of my model is beyond the scope of this chapter. However, after relaxing a series of assumptions made to preserve analytical tractability, my model could be calibrated to firm-level data in order to provide an estimate for the productivity loss triggered by a financial shock. Such a quantitative model could also be used to study the effect of government policies, such as countercyclical R&D subsidies targeted at small firms.

[61] Note that I do not control for capacity utilisation (which is more volatile in manufacturing, and especially in the car industry) in Table 1.5. Doing so would probably accentuate volatility differences even further.

1.6 Appendix to Chapter 1

1.6.1 Theoretical Appendix

1.6.1.1 The firm's dynamic programming problem

The firm's problem An incumbent with n_t products ($n_t \geq 1$) and cash holdings c_t solves

$$\max_{\left(\mathbf{1}_{R,s},\mathbf{1}_{I,s}\right)_{s\geq t}} E_t\left(\sum_{s=t}^{+\infty}\mathbf{1}_{E,s}\tilde{c}_s\right),$$

where $\mathbf{1}_{E,s}$ is an indicator function for the firm still existing in period s and $\tilde{c}_s$ is its cash flow in that period. Cash flows depend on R&D decisions as described in the main text. For instance, if the firm realises a radical innovation opportunity in period s, this yields cash flows of $-f_R$ in period s and $\left(1-\frac{1}{\gamma}\right)S_{s+1}$ in period $s+1$.

Maximization faces two constraints. First, in periods with crisis financing conditions, Equation (1.5) must hold. This makes cash holdings a state variable. Second, the firm can only collect cash flows as long as it is exists ($\mathbf{1}_{E,s}=1 \Leftrightarrow n_s \geq 1$). The number of products produced by the firm follows the stochastic process

$$n_{s+1} = \begin{cases} n_s - 1 & \text{with probability } d_s\left(1-\alpha\mathbf{1}_{R,s}\right) \\ n_s + 1 & \text{with probability } \alpha\mathbf{1}_{R,s}\left(1-d_s\right) \quad \text{if } n_s \geq 1. \\ n_s & \text{else} \end{cases}$$

Once a firm falls to zero products, it exits forever. Potential entrants in period t solve the same problem, but for them, $\mathbf{1}_{E,t}=1$ (they exist in period t even though they do not produce). If they receive a radical innovation opportunity and decide to pay the radical R&D cost, they enter and produce one product in the next period. Otherwise, they stay out forever.

A basic property of the value function I define V as the value of an incumbent's problem in period t, after earning current profits and before learning whether it receives a radical innovation opportunity. V depends on the endogenous state variables n_t and c_t, but also on current and future financing conditions, future values of aggregate demand and the path of future destruction rates $(d_s)_{s\geq t}$. Future destruction rates depend on the joint distribution of products and cash across firms (which determine the mass of contestable products and the mass of constrained firms if a financial shock hits). Thus, the problem must in principle be solved with the Krusell and Smith (1998) algorithm. However, the parameter restrictions made in the main text simplify the problem substantially.

First, notice that the value function V is non-decreasing in n_t and in c_t. Indeed, a firm with n_t products can mimic any strategy of a firm with a lesser number of products (while the reverse is not possible) and therefore must achieve on expectation at least the value of that firm.[62] The same observation holds for cash levels.

Policy functions To simplify notation, I regroup all state variables which are exogenous to the firm in the (infinite-dimensional) vector Φ_t and denote by $NPV_{R,t}$ and $NPV_{I,t}(n_t)$ the NPV delivered by radical or incremental innovation paid for in period t for a firm with n_t products.[63] Consider first a period t in which financing conditions are normal. The Bellman equation for an incumbent's problem is

$$V(n_t, c_t, \Phi_t) = (1-\alpha)\max(V_1, V_3) + \alpha \max(V_1, V_2, V_3, V_4),$$

where V_1 to V_4 are the values associated to the firm's choices in period t. V_1 is the expected value of the firm when doing neither radical nor incremental R&D:

$$V_1 = d_t E_t(V(n_t - 1, c_t, \Phi_{t+1})) + (1-d_t) E_t(V(n_t, c_t, \Phi_{t+1})).$$

V_2 is the expected value when doing radical, but no incremental R&D:

$$V_2 = NPV_{R,t} + d_t E_t(V(n_t, c_t + NPV_{R,t}, \Phi_{t+1})) + (1-d_t) E_t(V(n_t + 1, c_t + NPV_{R,t}, \Phi_{t+1})).$$

V_3 is the expected value when doing incremental, but no radical R&D:

$$\begin{aligned} V_3 = \; & NPV_{I,t}(n_t) + d_t E_t(V(n_t - 1, c_t + NPV_{I,t}(n_t), \Phi_{t+1})) \\ & + (1-d_t) E_t(V(n_t, c_t + NPV_{I,t}(n_t), \Phi_{t+1})). \end{aligned}$$

and V_4 is the expected value when doing both radical and incremental R&D:

$$\begin{aligned} V_4 = \; & NPV_{R,t} + NPV_{I,t}(n_t) + d_t E_t(V(n_t, c_t + NPV_{R,t} + NPV_{I,t}(n_t), \Phi_{t+1})) \\ & + (1-d_t) E_t(V(n_t + 1, c_t + NPV_{R,t} + NPV_{I,t}(n_t), \Phi_{t+1})). \end{aligned}$$

With probability $1-\alpha$, the firm receives no radical innovation opportunity and therefore only

[62] If the optimal strategy of the smaller firm involves incremental R&D, the value of the larger firm is strictly higher than that of the smaller one, as it earns more profits from incremental innovation.

[63] That is, $NPV_{R,t} = \left(1 - \frac{1}{\gamma}\right) S_{t+1} - f_R$ and $NPV_{I,t}(n_t) = (n_t - 1)\left(1 - \frac{1}{\delta}\right) S_{t+1} - f_I$.

has the choice between V_1 and V_3. Clearly, V_3 is larger than V_1 if and only if $NPV_{I,t}(n_t) > 0$ (which is true if n_t is larger or equal to the threshold size $n^*_{I,t}$ defined in the main text). In this case, incremental innovation has a positive NPV and increases cash next period for sure. Both effects have some non-negative value for the firm (the second because V is non-decreasing in cash), so doing incremental R&D is better than not doing it. Likewise, $n_t \geq n^*_{I,t}$ also implies $V_4 > V_2$. As radical innovation always delivers a positive NPV and increases both the cash holdings and the number of products produced next period, it is always true that $V_2 > V_1$ and $V_4 > V_3$. Thus, with normal financing conditions, all firms do radical R&D if they receive a radical innovation opportunity,[64] and do incremental R&D if and only if $n_t \geq n^*_{I,t}$.
When there are crisis financing conditions, firms' R&D decisions remain unchanged as long as they can finance them with their cash holdings. Firms that would want to do both radical and incremental R&D, but only have enough cash for one of the two, prefer to maintain radical R&D because it delivers both a higher NPV (because of Restriction 3) and an increase in expected firm size. Thus, firms do radical R&D if and only if they get a radical innovation opportunity and their cash holdings exceed f_R, and do incremental R&D if $n_t \geq n^*_{I,t}$ and their cash holdings after the radical innovation decision exceed f_I.

1.6.1.2 Proof of Equation (1.12): the law of motion of the aggregate price level

Substituting Equation (1.8) into Equation (1.7) yields $P_t = \exp\left(\int_0^1 \ln \frac{w_t}{a_{t-1}(j)} dj\right)$. From period t to period $t+1$, the price of a mass R_{t-1} of products is multiplied by $\frac{w_{t+1}}{\gamma w_t}$, the price of a mass I_{t-1} by $\frac{w_{t+1}}{\delta w_t}$ and the price of the remainder by $\frac{w_{t+1}}{w_t}$. Therefore,

$$\begin{aligned} P_{t+1} &= \tfrac{w_{t+1}}{w_t} \exp\left(\int\limits_0^1 \left(\ln p_t(j) + R_{t-1} \ln\left(\tfrac{1}{\gamma}\right) + I_{t-1} \ln\left(\tfrac{1}{\delta}\right)\right) dj \right) \\ &= P_t \tfrac{w_{t+1}}{w_t} \exp\left(-\left(R_{t-1}\ln\gamma + I_{t-1}\ln\delta\right)\right). \end{aligned}$$

1.6.1.3 The invariant firm size distribution

On the balanced growth path, for every $n \geq 2$, the law of motion in (1.15) can be rewritten as

$$m_{n+1} = \left(1 + \frac{\alpha(1-d)}{(1-\alpha)d}\right) m_n - \frac{\alpha(1-d)}{(1-\alpha)d} m_{n-1}.$$

The characteristic equation of this sequence has roots 1 and $\frac{\alpha(1-d)}{(1-\alpha)d}$. Therefore, there exist

[64] This also holds for potential entrants, which cannot do incremental R&D and whose outside option when not taking the radical innovation opportunity (staying out forever) has value 0.

two real numbers C_1 and C_2 such that

$$\forall n \geq 1, \quad m_n = C_1 + C_2 \left(\frac{\alpha (1-d)}{(1-\alpha) d} \right)^{n-1}.$$

As $\sum_{n=1}^{+\infty} m_n$ must be finite, $C_1 = 0$. This implies $C_2 = \frac{\alpha}{(1-\alpha)d} m_0$. Furthermore, note $d = \frac{\alpha(m_0+m)}{m} = \alpha(1+\psi)$ to get

$$\forall n \geq 1, \quad m_n = \frac{m_0}{(1-\alpha)(1+\psi)} \left(\frac{1-\alpha(1+\psi)}{(1-\alpha)(1+\psi)} \right)^{n-1}.$$

m_0 is pinned down by the condition that all products must be produced, that is, $\sum_{n=1}^{+\infty} n m_n = 1$. An increase in ψ shifts the invariant distribution to the left, as claimed in Footnote 27.

$$\forall n \geq 1, \quad \frac{\partial m_n}{\partial \psi} = \psi \frac{(1-\alpha(1+\psi))^{n-2}}{((1-\alpha)(1+\psi))^{n+1}} \left(2(1-\alpha) - \frac{(n+1)\psi}{1+\psi} \right).$$

This has the sign of $2(1-\alpha) - \frac{(n+1)\psi}{1+\psi}$. It is easy to show that $\frac{\partial m_1}{\partial \psi} > 0$, using the condition $\alpha(1+\psi) < 1$. For all other n, the sign of the derivative is indeterminate. However, the term defining the derivative's sign is decreasing in n, and $\lim_{n \to +\infty} 2(1-\alpha) - \frac{(n+1)\psi}{1+\psi} = -\infty$. Thus, there is a threshold size $n^* \geq 2$ at which the sign of $\frac{\partial m_n}{\partial \psi}$ changes from positive to negative.

1.6.1.4 Tracking the fraction of unconstrained firms

In all periods with crisis financing conditions, I need to compute the fractions of firms of each size class which have cash holdings exceeding f_I, f_R and $f_I + f_R$. These fractions are key for determining aggregate innovation masses (Equations (1.10) and (1.11)) and the evolution of the firm size distribution (Equation (1.15)). I calculate these fractions by tracking the joint distribution of size and cash for firms which may be constrained under crisis financing conditions (that is, for firms with cash holdings lower than $f_I + f_R$).[65]

The evolution of the mass of firms which are in state (n_t, c_t) in period t (that is, which produce n_t products and have cash holdings c_t after production) depends on the current state of financing conditions. When financing conditions are normal, there are four different cases:

1) A fraction $\alpha(1-d_t)$ of firms transitions to state $\left(n_t+1, c_t + NPV_{R,t} + \mathbf{1}_{n_t \geq n^*_{I,t}} NPV_{I,t}(n_t)\right)$, where $\mathbf{1}_{n_t \geq n^*_{I,t}}$ is an indicator function for the firm producing $n^*_{I,t}$ products or more. These firms gain one product by radical innovation, do not lose their contestable product and in-

[65] Recall that as cash holdings never fall, all firms which pass this threshold are forever unconstrained.

crease their cash holdings with the cash flow from radical innovation (and incremental innovation, if they are large enough to do incremental R&D).
2) A fraction αd_t transitions to state $\left(n_t, c_t + NPV_{R,t} + \mathbf{1}_{n_t \geq n^*_{I,t}} NPV_{I,t}(n_t)\right)$.
3) A fraction $(1-\alpha)(1-d_t)$ transitions to state $\left(n_t, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} NPV_{I,t}(n_t)\right)$.
4) The remaining fraction $(1-\alpha) d_t$ transitions to state $\left(n_t - 1, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} NPV_{I,t}(n_t)\right)$, or exits if $n_t = 1$.
Finally, a mass $\alpha m_{0,t}$ of entrants transition to state $(1, \tilde{c}_{R,t})$.

Under crisis financing conditions, there is no entry, and transitions of incumbent firms depend on their cash holdings. If $c_t < f_R$, firms cannot do radical innovation, and there are just two cases for firms in state (n_t, c_t) in period t.
1) A fraction $(1-d_t)$ transitions to state $\left(n_t, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I} NPV_{I,t}(n_t)\right)$.
2) The remaining fraction d_t transitions to state $\left(n_t - 1, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I} NPV_{I,t}(n_t)\right)$, or exits if $n_t = 1$.

Finally, if $c_t \geq f_R$, firms can do radical innovation, and there are again four cases for firms in state (n_t, c_t) in period t.
1) A fraction $\alpha(1-d_t)$ transitions to state $\left(n_t + 1, c_t + NPV_{R,t} + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I + f_R} NPV_{I,t}(n_t)\right)$.
2) A fraction αd_t transitions to state $\left(n_t, c_t + NPV_{R,t} + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I + f_R} NPV_{I,t}(n_t)\right)$.
3) A fraction $(1-\alpha)(1-d_t)$ transitions to state $\left(n_t, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I} NPV_{I,t}(n_t)\right)$.
4) The remaining fraction $(1-\alpha) d_t$ transitions to state $\left(n_t - 1, c_t + \mathbf{1}_{n_t \geq n^*_{I,t}} \mathbf{1}_{c_t \geq f_I} NPV_{I,t}(n_t)\right)$, or exits if $n_t = 1$.
Using these transition laws, I can track the joint distribution of size and cash holdings over time and calculate the fractions of potentially constrained firms of each size class.

To determine the joint distribution of size and cash for potentially constrained firms on the balanced growth path, I could start from an arbitrary distribution and iterate the laws of motion described above until convergence. However, there is a more simple computation method, taking into account firm age.
I denote by $m_n(c, \mathfrak{a})$ the mass of firms with n products, cash holdings c and age $\mathfrak{a}$. Firms of age 1 (last period's entrants) must necessarily have one product and cash holdings NPV_R. Therefore, $m_1(NPV_R, 1) = \alpha m_0$. From this, I deduce the masses of firms of age 2, using the same transition processes as before and taking into account that all firms of age 2 must have had age 1 last period. Iterating forward, I calculate the distribution of size and cash for every age $\mathfrak{a}$, and get the unconditional distribution by integrating over this last dimension.

1.6.1.5 Parameter values for simulations

For Figures 1.3 to 1.6, I assume $\psi = 0.2$. All other parameter values are given in Table 1.6.

Table 1.6: Parameter values

γ	1.16	f_R	0.1
δ	1.03	f_I	0.025
S_H	1	α	0.7
S_L	0.85	$L(w)$	$\sqrt{w}$

1.6.2 Data Appendix

1.6.2.1 The MIP survey and my dataset

The ZEW sends the MIP survey questionnaires every year to a representative panel of German firms with at least 5 employees in "innovating" industries.[66] The panel is updated every two years, with firms exiting the target population being replaced by new ones to maintain representativeness.

While all firms in the panel receive the survey every year, they unfortunately do not all respond. The dataset is therefore a highly unbalanced panel, with many firms responding in some years but not in others. The selection bias arising from voluntary participation, however, does not appear to be very important. Non-respondent surveys conducted by telephone have shown that the share of innovating firms among responding firms is only slightly lower than in the entire sample (Peters (2009)).

In my analysis, I consider the period 1999-2009. Earlier data, for 1992-1998, is available but partial and more noisy. Data for later years had not yet been released at the time of writing. The original dataset contains 55889 observations for 20144 different firms. I do several adjustments on the raw data, listed below.

1. I delete observations which do not belong to industries targeted by the MIP, using the definition of the MIP's scope provided in Rammer et al. (2011). I also drop observations with missing or incorrect NACE industry codes and, following a usual convention in firm-level

[66] This includes manufacturing, mining and some services (wholesale trade, transport, software development and services to firms). It excludes agriculture, construction, retail trade, real estate and health care.

studies, the financial industry (NACE codes 650 to 672) and the energy and water industry (NACE codes 400 to 410).

2. I delete observations with missing or zero values for employment (variable bges) or sales (variable um).[67]

3. I delete several outliers. In some firms, the unit of reference changes over time (a subsidiary may report sometimes its own values, and sometimes values for the mother). This creates large jumps in firm-level time series. To avoid them, I drop for every firm observations with sales or employment which are two times larger or two times smaller than the firm's median for the respective variable. This excludes around 5% of all observations. Note that this does not eliminate outliers for firms with only one observation. However, these are irrelevant for my analysis, as they are not used by definition in a regression with firm fixed effects. I additionally control for outliers by winsorizing the variables used in regressions as indicated in the main text.

After these adjustments, the dataset contains 41874 observations for 16506 different firms. Table 1.7 shows some summary statistics for the firms which are effectively used in my analysis, that is, the firms which have at least two observations and at least one positive observation for R&D. This sample has a considerably smaller number of observations, which is due to the fact that many firms never spend on R&D. Note that there have been some changes in the scope of the survey over time, so that cross-sectional statistics in different years are not directly comparable. For instance, the fall in median R&D is just due to the fact that the sample has been enlarged over the years to better track small firms with lower R&D budgets.

The MIP survey asks firms about their R&D spending ("Forschungs- und experimentelle Entwicklungsausgaben") and innovation spending ("Innovationsausgaben"). R&D is defined as "*the systematic creative work for the enlargement of current knowledge and the use of the so gained knowledge to develop new products or processes*" (Rammer et al. (2011), my translation). While this variable is widely used in the empirical literature, innovation spending is a less common concept. It covers, on top of R&D, several types of implementation spending.[68] In the stylized framework of my model, innovation spending and R&D investment coincide.

[67] I also fill in some missing values. If a firm reports in a survey that it never did R&D over the last three years (variable fuekon equal to 0) and has a missing value for R&D in that year, I replace the missing value by a 0. I do the same for non-innovating firms which have a missing value for innovation spending. Non-innovating firms are firms which have had no product or process innovation, no current innovation project and no interrupted innovation project in the last three years (variables pd, pz, pa and pn all equal to 0) or firms which have not spent on innovation projects (variable iagesno equal to 1).

[68] This includes buying machines to produce a new product, marketing expenses, patent or licence purchases, spending on continuing education of employees, etc.

In reality, innovation spending may capture some innovative activities ignored by the standard R&D measure. As indicated in the main text and shown in the next section, my results do not depend on using one measure or the other.

Table 1.7: Summary statistics for the final sample

	Obs.	**% Small Firms**	**Median R&D**	**Median R&D int.**
1999	967	31.2%	102’285 €	0.92%
2000	1199	38.4%	51.130 €	0.50%
2002	1541	42.6%	85’000 €	1.03%
2003	1397	41.4%	100’000 €	1.03%
2004	1608	44.4%	80’000 €	1.01%
2005	1693	45.2%	60’000 €	0.78%
2006	1745	44.9%	65’000 €	0.76%
2007	2053	45.2%	50’000 €	0.63%
2008	1757	46.6%	50’000 €	0.60%
2009	1827	47.9%	50’000 €	0.66%
Total	**15787**	**43.6%**	**51’129 €**	**0.71%**

Notes: All statistics refer to the sample of innovating firms with at least two observations. Small firms are firms with 50 employees or less. R&D int. stands for R&D intensity. In 2001, the survey did not contain a question about R&D.

1.6.2.2 Additional tables

This section contains the additional tables referred to in the main text in Section 1.4.1.
Table 1.8 shows the median rates of change of innovation spending and sales by firm size during the Great Recession, conditional on the sign of sales changes.[69]
Table 1.9 shows the results of the baseline estimation for R&D intensity when using industry-year fixed effects (at the NACE two-digit level) or permanent size dummies.
Finally, Table 1.10 shows the results when Equation (1.20) is estimated using innovation intensity as the dependent variable.

[69] The general pattern of Tables 1.2 and 1.8 is not changed when considering mean instead of median changes. These tables are available on request.

Table 1.8: Innovation spending and sales fluctuations by firm size during the Great Recession

	2008			**2009**		
Median rate of change	**Inn. Sp.**	**Sales**	**Obs.**	**Inn. Sp.**	**Sales**	**Obs.**
Firms with decreasing sales						
Small (50 or less emp.)	-28.57	-8.70	167	-34.63	-18.56	362
Large (more than 500 emp.)	-22.22	-6.88	53	-7.79	-22.67	101
Firms with increasing sales						
Small (50 or less emp.)	0.00	10.91	349	0.00	8.97	285
Large (more than 500 emp.)	0.00	6.45	95	3.93	7.69	36

Notes: See Table 1.2. Firms with trivial innovation spending (less than 10.000€) are excluded.

Table 1.9: R&D intensity regression robustness checks

Dependent variable: R&D intensity

Sample	All firms	Res.-int. manuf.	All firms	Res.-int. manuf.
SED^{1999}	1.41 (0.56)**	2.28 (1.01)**	1.40 (0.48)***	2.01 (0.92)**
SED^{2000}	1.22 (0.49)**	3.50 (0.83)***	1.10 (0.46)**	2.98 (0.81)***
SED^{2002}	2.33 (0.48)***	4.63 (0.83)***	2.44 (0.44)***	4.08 (0.77)***
SED^{2003}	1.36 (0.42)***	3.16 (0.73)***	1.34 (0.40)***	2.83 (0.70)***
SED^{2004}	1.93 (0.47)***	3.23 (0.96)***	2.29 (0.44)***	3.14 (0.86)***
SED^{2005}	0.84 (0.40)**	1.84 (0.69)***	1.16 (0.40)***	1.78 (0.64)***
SED^{2006}	0.80 (0.42)*	1.96 (0.79)**	0.92 (0.37)**	1.76 (0.77)**
SED^{2007}	0.91 (0.40)**	2.65 (0.67)**	0.99 (0.35)***	2.44 (0.60)***
SED^{2008}	0.89 (0.39)**	1.42 (0.62)**	0.88 (0.36)**	1.35 (0.60)**
Specification	Industry-year fixed effects		Permanent firm size dummies	
Obs.	15787	5729	15206	5729

Notes: See Table 1.3. Note that the R&D services industry is included in the regressions with industry-year fixed effects (as those should account for industry-level differences in the elasticity of R&D with respect to sales).

Table 1.10: Panel regressions with innovation intensity

Dependent variable: Innovation intensity

Sample	All firms	Res.-int. manuf.
$\mathrm{SE}D^{1999}$	$2.27\,(0.61)^{***}$	$1.99\,(1.24)$
$\mathrm{SE}D^{2000}$	$1.45\,(0.53)^{***}$	$3.26\,(1.00)^{***}$
$\mathrm{SE}D^{2001}$	$1.79\,(0.55)^{***}$	$1.74\,(1.07)$
$\mathrm{SE}D^{2002}$	$1.71\,(0.61)^{***}$	$2.65\,(0.99)^{***}$
$\mathrm{SE}D^{2003}$	$2.37\,(0.56)^{***}$	$3.49\,(0.96)^{***}$
$\mathrm{SE}D^{2004}$	$2.65\,(0.49)^{***}$	$2.88\,(1.01)^{***}$
$\mathrm{SE}D^{2005}$	$1.82\,(0.45)^{***}$	$1.57\,(0.81)^{*}$
$\mathrm{SE}D^{2006}$	$1.00\,(0.51)^{**}$	$0.93\,(0.96)$
$\mathrm{SE}D^{2007}$	$1.65\,(0.45)^{***}$	$1.71\,(0.83)^{**}$
$\mathrm{SE}D^{2008}$	$0.95\,(0.47)^{**}$	$-0.08\,(0.82)$
Obs.	23554	7022

Notes: See Table 1.3. There are now 6351 firm-level and 384 industry-year level clusters in the full sample (2011 and 88 in the research-intensive manufacturing sample). The sample now includes all firms which have at least one non-zero observation for innovation spending.

1.6.2.3 Industry-level productivity data

In Section 1.4.3.1, I measure industries' labour productivity as real shipments per employee. Real shipments are defined as $\frac{\text{vship}}{\text{piship}}$ (using the variable names of the NBER-CES database) and employees are given by emp. I limit my analysis to the period 1972-2009, as data on capacity utilisation, a control variable in my regression, is only available from 1972 onwards. As regressions are specified in natural logarithms, I omit industries which had negative average labour productivity growth from 1972 to 2009. These industries are Manufactured Home (Mobile Home) Manufacturing (NAICS Code 321991), Prefabricated Wood Building Manufacturing (321992), All Other Leather Good and Allied Product Manufacturing (316999), Coffee and Tea Manufacturing (311920) and Prefabricated Metal Building and Component Manufacturing (332311).

Chapter 2

International Trade and Innovation Dynamics

2.1 Introduction

How does international trade affect productivity growth? Due to its obvious importance, this question has been widely analysed in the academic literature. However, while economic theory has identified a number of channels through which trade may affect productivity growth, empirical evidence lags behind: "*many regressions using cross-country data have been computed, [but] they teach us little about mechanisms at work*" (Grossman and Helpman (2015, p.104)).

These identification problems could potentially be overcome by exploiting micro-level heterogeneity, that is, the potentially different responses of different firms and industries to trade. Indeed, heterogeneous-firm models are omnipresent in the theory of international trade since the pioneering work of Melitz (2003). Melitz showed that the reallocation of resources between firms with different productivities is a key source of aggregate productivity gains after a fall in trade costs. In his framework, firms' productivity is exogenous. In reality, however, productivity is constantly improved by innovations created by firms' R&D investments. These innovations enter the economic system through creative destruction, as innovating firms expand at the expense of those that fail to innovate.

In this chapter, I analyse how international trade affects this firm-level creative destruction process, both in tradable and nontradable industries. To do so, I extend the most influential heterogeneous-firm endogenous growth model, due to Klette and Kortum (2004), to an international setting. Their model shows how firm-level R&D decisions determine creative destruction, aggregate productivity and the size distribution of innovating firms, matching many important facts on firm dynamics (Aghion et al. (2014)).[1] I modify their basic setup by considering two symmetric countries and allowing for a fraction of products to be exported subject to a trade cost. Therefore, every country has a tradable and a nontradable sector, linked through the national labour market. This model provides a new set of predictions at the firm and at the industry level, shedding more light on the way trade expansions affect productivity growth.

My model predicts that due to a market size effect, creative destruction and productivity growth in the tradable sector are always higher than in the nontradable one. Thus, the effect of an expansion in international trade is quite different depending on whether it is due to falling trade costs in the tradable sector, or to formerly nontradable products becoming tradable.
Falling trade costs in the tradable sector allow innovators to charge higher mark-ups abroad, and therefore encourage innovation and entry in this sector. Productivity growth increases, and greater creative destruction shifts the size distribution of innovating firms to the left. The nontradable sector is not affected, because the wage-to-GDP ratio does not change: higher labour demand for tradable sector R&D is exactly offset by lower labour demand for trade costs. Using a dataset on trade costs and productivity growth in US manufacturing industries due to Bernard et al. (2003) and Becker et al. (2013), I find supporting evidence for this prediction: a fall in the trade costs of an industry is positively correlated with an increase in its long-run productivity growth rate.
An increase in the fraction of tradable products increases innovation, entry and productivity growth for the newly tradable products. However, it also drives up labour demand and thus the wage-to-GDP ratio. This increases the relative cost of innovation and therefore lowers innovation incentives for the producers of all other products. The effect of these changes on aggregate productivity growth and on the size distribution of innovating firms is ambiguous. These predictions are consistent with industry-level evidence from the United States. In the beginning of the 1980s, productivity growth in service industries was markedly lower than in (tradable) goods-producing industries. As technological advances made a range of services tradable for the first time in history, productivity growth in services accelerated, while productivity growth in goods-producing industries stagnated or fell.

[1] According to Melitz and Redding, there is "*relatively little understanding of the processes through which large and successful firms emerge and the implications of these processes for the transitional dynamics of the economy's response to trade liberalization*." (Melitz and Redding (2014, p.49)). I argue in this chapter that Klette and Kortum's approach provides a good way to understand how large and successful firms emerge, and one can therefore learn a lot by using it to analyse an economy's reaction to international trade.

My analysis is related to a vast literature on trade and productivity growth. In a recent review, Grossman and Helpman (2015) note that this literature has identified four main channels through which trade affects growth: market size effects, knowledge spillovers, competition and technology diffusion. These channels have been analysed in one-sector, representative-firm endogenous growth models by, among many others, Grossman and Helpman (1991), Rivera-Batiz and Romer (1991) or Bloom et al. (2014).[2] In my model, I analyse how they affect firm-level creative destruction dynamics.[3]

Some authors have incorporated R&D and innovation into Melitz's heterogeneous-firm model. Baldwin and Robert-Nicoud (2008) introduce knowledge spillovers à la Romer (1990) for the creation of new product varieties, and find that trade has ambiguous effects on productivity growth. Their model maintains the assumption of an exogenously given productivity distribution, only has one (tradable) sector, and does not feature creative destruction.[4] Bustos (2011) does not assume spillovers, but introduces a technology allowing firms to upgrade their productivity when paying a fixed cost. This generates a complementarity between exports and innovation, and heterogeneous responses of firms to trade liberalization.[5]

Bustos finds evidence for these heterogeneous responses when analysing Argentinean firms' decisions to upgrade technology in response to the creation of the MERCOSUR customs union. Her findings are similar in spirit to Lileeva and Trefler (2010), Aw et al. (2011) and Steinwender (2015), who all point out a heterogeneous response of R&D or technology upgrading at the firm or plant level after trade liberalizations. Their evidence supports the view that accounting for firm and industry-level heterogeneities is important to understand how trade affects productivity growth.

The remainder of the chapter is structured as follows. Section 2.2 describes the assumptions and solves the theoretical model. Section 2.3 discusses its main predictions, and an extension to international knowledge spillovers. Section 2.4 presents the empirical evidence, and Section 2.5 concludes. Section 2.6 contains the Appendix.

[2] All these papers study the effect on trade on productivity growth in endogenous growth models. However, the same question has also been studied in Ricardian models with exogenous productivity improvements (Eaton and Kortum (2001)) and in models of technology diffusion (Perla et al. (2015)).

[3] Market size, competition and technology diffusion all play an important role in my analysis. However, in the baseline version of my model, I assume that there are no international knowledge spillovers. In Section 2.3.4, I extend the model to incorporate these spillovers.

[4] Gustafsson and Segerstrom (2010) provide a similar model which does not feature scale effects.

[5] Atkeson and Burstein (2010) analyze a different version of this model, to assess the effect of innovation composition on the gains from trade.

2.2 The model

2.2.1 Assumptions

Time is continuous ($t \in \mathbb{R}^{+}$), and the world is composed of two symmetric countries, Home and Foreign. Foreign variables are denoted with an asterisk.

Consumers Every country is populated by a mass L of identical consumers, who value consumption paths according to the intertemporal utility function

$$U(t) = \int_{t}^{+\infty} \exp\left(-(s-t)\rho\right) \ln C(s)\, ds, \tag{2.1}$$

where $C(s)$ stands for the utility derived from consumption at time s. This utility is a Cobb-Douglas aggregate of the consumption levels for a continuum of differentiated products:

$$C(t) = \exp\left(\int_{0}^{1} \ln c(t,j)\, dj\right), \tag{2.2}$$

where $c(t,j)$ is the quantity of product j consumed at time t. Consumers inelastically supply one unit of labour at every time t and receive an equal share of firms' profits.

Nontradable and tradable sectors The unit mass of products consumed in Home is composed of a fraction $1-\theta$ of nontradable products (which can neither be exported or imported), and a fraction θ of tradable products. Tradable products are subject to an iceberg trade cost τ: for one unit of a tradable product to arrive in Foreign, a Home producer needs to ship τ units (with $\tau > 1$).

Firms There is a continuum of firms, which can operate either in the tradable or in the nontradable sector, but not in both. A firm i is characterized by a function $a_i(t,\cdot)$, associating to every product of its sector a productivity with which the firm can produce. Precisely,

$$y_i(t,j) = a_i(t,j)\, l_i(t,j), \tag{2.3}$$

where $y_i(t,j)$ stands for output of product j by firm i, and $l_i(t,j)$ for the labour it employs. Firms compete à la Bertrand on the market for each differentiated product. This implies that in equilibrium, only one firm (the one with the highest productivity) will produce a given product. I assume that at $t = 0$, any given firm has the highest productivity worldwide for at most a finite number of products. The number of products a firm produces is denoted by n.

Radical innovation Innovations drive productivity growth. I refer to them as radical, because they entail displacement and creative destruction.

An incumbent firm which has the highest productivity for n different products in its sector can at any time generate a Poisson arrival rate Z of radical innovations by hiring $\zeta n \left(\frac{Z}{n}\right)^{\eta}$ (with $\zeta > 0$ and $\eta > 1$) units of labour. This cost function is decreasing in n, indicating that incumbent firms can build on the knowledge capital generated by past innovations (which gave it the highest productivity for n products) to generate new ones.[6]

Upon arrival, a radical innovation enables the firm to increase the highest domestic productivity for a randomly selected domestically produced product of its sector by a factor of γ. I denote the highest domestic productivity for product j at time t by $a(t, j)$, where $a(t, j) = \max_{i \in \mathscr{I}_t} a_i(t, j)$ (and $\mathscr{I}_t$ is the set of domestic firms active at time t). Then, a radical innovation enables the innovating firm to produce with productivity $\gamma a(t, j)$, and thereby to displace the previous producer and expand at its expense.

Radical innovation can also be done by entrants. In both sectors, potential entrants may employ φ units of labour in order to generate radical innovations with a Poisson arrival rate 1. When a radical innovation arrives, the entrant becomes an incumbent firm with one product.[7]

Knowledge spillovers and technology diffusion My assumptions on the innovation technology imply that domestic innovators can only improve on the highest domestic productivity. This is in line with empirical evidence showing that most knowledge spillovers are local

[6]This radical innovation technology is substantially different from the one considered in Chapter 1. Instead of assuming that radical innovation opportunities are independent of firm size, I follow here the original Klette and Kortum model by assuming a radical R&D cost function which implies that in equilibrium, radical innovation choices will be exactly proportional to firm size. Thus, in the present chapter, the relative innovation capacity of a firm does not depend on its size.

I make this assumption for simplicity and to facilitate comparisons with Klette and Kortum. However, it is not crucial for my results, which would equally hold for alternative radical R&D cost functions generating a higher relative innovation capacity of small firms. This is due to the fact that trade expansions affect all firms symmetrically in my model (while in Chapter 1, innovation capacity differences mattered because small and large firms were also affected differently by aggregate shocks). In reality, of course, trade expansions may have asymmetric effects for firms of different sizes: access to the Chinese market probably does not mean the same thing to Google than to a small Silicon Valley start-up. Extending the model in this chapter to allow for such asymmetries is an important objective for future work.

[7]In principle, incumbents could also use the entrant technology to generate radical innovations. However, when $\varphi \geq \zeta$, which I impose from now on, they never want to do so, as their own technology is more efficient.

(Jaffe et al. (1993), Murata et al. (2014)). However, international trade (and economic openness more generally) certainly also generate knowledge spillovers between countries, that is, instances where Home firms improve on Foreign technologies. In Section 2.3.4, I analyse a version of my model incorporating these spillovers.
Finally, I allow for technology diffusion. I assume that when a product is sold on one country's market (whether by a Home or a Foreign producer), all domestic firms learn, through imitation or reverse engineering, to produce it with a productivity that is only by a factor γ lower than that of the current producer. In the presence of trade costs, technology diffusion limits exporters' mark-ups and thereby generates interesting economic effects.

When $\theta = 0$, my model reduces to the simplified version of Klette and Kortum's closed economy model exposed in Aghion et al. (2014). When $\theta > 0$, however, international trade generates interesting new results. These are derived and discussed in the next sections.

2.2.2 The equilibrium

I consider a competitive equilibrium in which, in both countries, consumers maximize their intertemporal utility subject to their budget constraints, and firms choose prices and R&D to maximize the expected net present value of their lifetime profits. Furthermore, there is free entry in both sectors, trade is balanced, and all product and labour markets clear.
For simplicity, I focus on a symmetric balanced growth path equilibrium, in which all aggregate variables are equal in Home and Foreign and grow at the same rate.[8]

2.2.2.1 Consumption

In every country, the consumers' problem yields the standard intertemporal Euler equation

$$\frac{\mathring{C}(t)}{C(t)} = r - \rho, \tag{2.4}$$

where r is the balanced growth path interest rate. Optimal choices within every time period give the demand functions

$$\forall j \in [0,1], \; c(t,j) = \frac{C(t)\,P(t)}{p(t,j)}, \text{ with } P(t) = \exp\left(\int_0^1 \ln p(t,j)\right). \tag{2.5}$$

I normalize the price index $P(t)$ to 1.

[8]This makes the balanced trade condition redundant, as trade is balanced by definition.

2.2.2.2 Pricing decisions

Pricing decisions are static. As firms are atomistic, they are taken product by product. Bertrand competition implies limit pricing: the most productive firm for a given product charges a price equal to the average cost of the second most productive one.
Consider first Home's nontradable sector. There, limit pricing yields for any product j

$$p(t,j) = \gamma \frac{w(t)}{a(t,j)}, \tag{2.6}$$

because the productivity of the most productive firm is always by a factor γ higher than the one of the second most productive firm. Accordingly, the flow profit of Home firms from producing one nontradable product is $\left(1-\frac{1}{\gamma}\right)C(t)$.
Now, consider a product j in the tradable sector for which a Home firm has the highest productivity in the world (that is, $\gamma a^*(t,j) \leq a(t,j)$). This Home firm then also serves the Foreign market if and only if $\tau \leq \gamma$, a restriction I impose from now onwards. Indeed, if $\tau > \gamma$, the technology diffusion generated by exporting would lower the production cost of Foreign firms below that of the Home one, making exporting unprofitable and eliminating all international trade in equilibrium. The Home firm charges a price

$$p(t,j) = \gamma \frac{w(t)}{a(t,j)} \text{ in Home, and } \quad p^*(t,j) = \frac{\gamma}{\tau}\frac{\tau w(t)}{a(t,j)} = \gamma \frac{w(t)}{a(t,j)} \text{ in Foreign.} \tag{2.7}$$

It therefore earns a flow profit $\left(1-\frac{1}{\gamma}\right)C(t)$ in the Home market and $\left(1-\frac{\tau}{\gamma}\right)C^*(t)$ in the Foreign market. Note that trade costs and technology diffusion jointly limit the mark-ups of exporters, and therefore the profits from exporting.[9]
Exactly analogous formulas apply when a Foreign producer has the highest productivity in the world for a tradable product. Finally, under a simple assumption on the productivity distribution at $t=0$, there are no products for which the highest Home and Foreign productivities coincide in equilibrium. This point is explained in greater detail in the next section.

2.2.2.3 Specialization in the tradable sector

I assume that at $t=0$, half of all tradable products have an absolute advantage for Home firms (i.e., $\gamma a^*(0,j) \leq a(0,j)$), while the other half have an absolute advantage for Foreign firms (i.e., $\gamma a(0,j) \leq a^*(0,j)$). As wages are equal, every country exactly produces the set

[9]In reality, a firm may circumvent trade costs through foreign direct investment (FDI). FDI is certainly important for productivity growth and technology diffusion, but outside of the scope of this chapter.

of products for which it has an absolute advantage at $t = 0$. Moreover, as domestic innovators can only improve on domestically produced products, the set of tradable products initially produced by Home firms will be produced by Home firms forever.

My assumption on the initial productivity distribution is not as arbitrary as it may seem. Indeed, when Home and Foreign start with the same productivity for a given set of products and generate innovations at the same Poisson rate, the fraction of products for which Home has a higher productivity tends to $\frac{1}{2}$ as time goes on, while the fraction of products for which Home and Foreign have the same productivity tends to 0. Accordingly, one may interpret $t = 0$ as the time when a mass θ of products becomes tradable for the first time, after a long period of complete autarky.[10]

I now analyse firms' R&D and innovation decisions, starting with the nontradable sector.

2.2.2.4 Innovation and entry in the nontradable sector

The only endogenous state variable for an incumbent firm in the nontradable sector is n, the number of products which it currently produces. This determines both the firm's current cash flow and its R&D costs. The Hamilton-Jacobi-Bellman equation for its value function V_{NT} is

$$rV_{NT,t}(n) - \mathring{V}_{NT,t}(n) = \max_{z_{NT}} \begin{bmatrix} n\left(1-\frac{1}{\gamma}\right)C(t) - \zeta n z_{NT}^{\eta} w(t) \\ +n z_{NT}\left(V_{NT,t}(n+1) - V_{NT,t}(n)\right) \\ -n\frac{x_{NT}}{1-\theta}\left(V_{NT,t}(n) - V_{NT,t}(n-1)\right) \end{bmatrix}, \tag{2.8}$$

where $z_{NT} = \frac{Z_{NT}}{n}$. Rewriting the problem in terms of z_{NT} considerably simplifies things by making the incumbent's R&D cost function linear in n. x_{NT} is the Poisson arrival rate of radical innovations in the Home nontradable sector. To obtain the arrival rate of innovations for any given product, it needs to be multiplied with $\frac{1}{1-\theta}$, the density of a uniform distribution on the set of nontradable products, as every product is equally likely to be innovated upon.

I conjecture that the value function solving (2.8) is linear in n, holding $V_{NT,t}(n) = n v_{NT} C(t)$, where v_{NT} is a positive constant. Then, the first-order condition gives

$$z_{NT} = \left(\frac{v_{NT}\frac{C(t)}{w(t)}}{\zeta\eta}\right)^{\frac{1}{\eta-1}}, \tag{2.9}$$

[10]Therefore, at $t = 0$, both countries experience static gains from trade because of productivity differences, as in a classic Ricardian trade model. Section 2.3.3 discusses these aspects in greater detail.

and solving for v_{NT} using the method of undetermined coefficients yields

$$v_{NT} = \frac{\left(1-\frac{1}{\gamma}\right) - \zeta \frac{w(t)}{C(t)} z_{NT}^{\eta}}{\rho + \frac{x_{NT}}{1-\theta} - z_{NT}}. \tag{2.10}$$

In an equilibrium with positive entry, the flow cost of entry $\varphi w(t)$ must be equalized to the expected flow profits from entry $v_{NT} C(t)$.[11] Replacing this equality into Equation (2.9) gives

$$z_{NT} = \left(\frac{\varphi}{\zeta \eta}\right)^{\frac{1}{\eta-1}}. \tag{2.11}$$

This simple expression for the R&D intensity of incumbent firms directly follows from the linearity of the value function V_{NT}. Indeed, linearity implies that the value of adding a product is independent of n. It must therefore equal the value of getting the very first product, pinned down by free entry.

Finally, substituting the free entry condition into Equation (2.10) gives

$$\frac{x_{NT}}{1-\theta} = \frac{\left(1-\frac{1}{\gamma}\right)\frac{C(t)}{w(t)} - \zeta z_{NT}^{\eta}}{\varphi} - (\rho - z_{NT}). \tag{2.12}$$

2.2.2.5 Innovation and entry in the tradable sector

The problem of a firm in the tradable sector is almost identical to the one of a firm in the nontradable sector. The only difference is that a tradable sector firm can sell its products on two markets. Therefore, the Hamilton-Jacobi-Bellman equation for its value function is

$$\begin{aligned} rV_{T,t}(n) - \mathring{V}_{T,t}(n) = \max_{z_T} \quad & \left[n\left(1-\frac{1}{\gamma}\right)C(t) + n\left(1-\frac{\tau}{\gamma}\right)C^*(t) - \zeta n z_T^{\eta} w(t) \right. \\ & + n z_T \left(V_{T,t}(n+1) - V_{T,t}(n)\right) \\ & \left. - n\frac{2x_T}{\theta}\left(V_{T,t}(n) - V_{T,t}(n-1)\right)\right] \end{aligned}, \tag{2.13}$$

where x_T denotes the Poisson arrival rate of radical innovations by Home firms in the tradable sector. The arrival rate of innovations for a particular product is obtained by multiplying this rate with the density of a uniform distribution on the set of tradable products produced by Home firms, $\frac{1}{\frac{\theta}{2}}$. Through the same steps as in the previous section, I get

[11] It can be shown that such an equilibrium exists if and only if $\left(1-\frac{1}{\gamma}\right)\left(\frac{L}{\varphi} - \frac{\theta}{2}\rho\right) \geq \frac{\rho}{\gamma} + \frac{1}{\eta}\left(\frac{\varphi}{\eta\zeta}\right)^{\frac{1}{\eta-1}}$, and I assume from now on that this parameter condition holds.

$$z_T = \left(\frac{\varphi}{\zeta\eta}\right)^{\frac{1}{\eta-1}}, \tag{2.14}$$

$$\frac{2x_T}{\theta} = \frac{\left(2 - \frac{1+\tau}{\gamma}\right)\frac{C(t)}{w(t)} - \zeta z_T^{\eta}}{\varphi} - (\rho - z_T), \tag{2.15}$$

where the last equation makes uses the fact that by symmetry, $C(t) = C^*(t)$. Note that as long as entry costs and R&D cost function parameters are the same in both sectors, $z_T = z_{NT} = z$.

2.2.2.6 Labour market clearing

Equations (2.12) and (2.15) determine the arrival rates of innovations in each sector as a function of parameters and of the wage-to-GDP ratio, $\frac{w(t)}{C(t)}$. The latter is pinned down by the labour market clearing condition.[12] Table 2.1 summarizes labour demands for all different uses and sectors in Home.

Table 2.1: Labour demands in Home

Use/Sector	Nontradable	Tradable
Production (including trade costs)	$(1-\theta)\frac{C(t)}{w(t)}\frac{1}{\gamma}$	$\theta\frac{C(t)}{w(t)}\frac{1+\tau}{2}\frac{1}{\gamma}$
Incumbent radical R&D	$(1-\theta)\zeta z^{\eta}$	$\frac{\theta}{2}\zeta z^{\eta}$
Entrant radical R&D	$(x_{NT} - (1-\theta)z)\varphi$	$\left(x_T - \frac{\theta}{2}z\right)\varphi$

Summing everything up, I get

$$\frac{C(t)}{w(t)} = L + \left(1 - \frac{\theta}{2}\right)\rho\varphi. \tag{2.16}$$

2.2.3 Aggregates and growth rates

Combining Equations (2.5), (2.6), (2.7) and (2.16), I can determine the level of aggregate consumption.

$$C(t) = A_{NT}(t)^{1-\theta} A_T(t)^{\frac{\theta}{2}} A_T^*(t)^{\frac{\theta}{2}} \frac{L + \left(1 - \frac{\theta}{2}\right)\rho\varphi}{\gamma}, \tag{2.17}$$

[12] By symmetry, I only consider the labour market of one country. Walras' law implies that if labour markets clear, aggregate income in every country equals aggregate consumption, so product markets clear as well.

where $A_{NT}(t) = \exp\left(\frac{1}{1-\theta}\int_{j\in NT} \ln a(t,j)\,dj\right)$, $A_T(t) = \exp\left(\frac{2}{\theta}\int_{j\in T} \ln a(t,j)\,dj\right)$ and $A_T^*(t) = \exp\left(\frac{2}{\theta}\int_{j\in T^*} \ln a^*(t,j)\,dj\right)$ are measures of aggregate productivity in the nontradable sector (NT) and in the parts of the tradable sector produced by Home and Foreign firms (T and T^*). The aggregate productivity growth rate is a weighted average of the growth rates of the different sectors:

$$g = \frac{\mathring{C}(t)}{C(t)} = (1-\theta)\frac{\mathring{A}_{NT}(t)}{A_{NT}(t)} + \frac{\theta}{2}\frac{\mathring{A}_T(t)}{A_T(t)} + \frac{\theta}{2}\frac{\mathring{A}_T^*(t)}{A_T^*(t)}. \tag{2.18}$$

Productivity growth rates in both sectors depend on innovation arrival rates:

$$g_{NT} = \frac{\mathring{A}_{NT}(t)}{A_{NT}(t)} = \frac{x_{NT}}{1-\theta}\ln\gamma, \quad g_T = \frac{\mathring{A}_T(t)}{A_T(t)} = \frac{\mathring{A}_T^*(t)}{A_T^*(t)} = \frac{2x_T}{\theta}\ln\gamma. \tag{2.19}$$

Therefore, using Equation (2.16), growth rates can finally be expressed as

$$\frac{g_{NT}}{\ln\gamma} = \left(1-\frac{1}{\gamma}\right)\left(\frac{L}{\varphi}-\frac{\theta}{2}\rho\right) - \frac{\rho}{\gamma} + \frac{\eta-1}{\eta}z, \tag{2.20}$$

$$\frac{g_T}{\ln\gamma} = \left(2-\frac{1+\tau}{\gamma}\right)\left(\frac{L}{\varphi}-\frac{\theta}{2}\rho\right) + \left(1-\frac{1+\tau}{\gamma}\right)\rho + \frac{\eta-1}{\eta}z, \tag{2.21}$$

$$g = (1-\theta)g_{NT} + \theta g_T. \tag{2.22}$$

2.2.4 The size distribution

On the balanced growth path, there is a unique invariant size distribution of firms by number of products, given by

$$\forall n \in \mathbb{N}^*, \; m_{NT,n} = (1-\theta)\frac{\frac{e_{NT}}{z}}{n\left(1+\frac{e_{NT}}{z}\right)^n} \text{ and } m_{T,n} = \frac{\theta}{2}\frac{\frac{e_T}{z}}{n\left(1+\frac{e_T}{z}\right)^n}. \tag{2.23}$$

$m_{NT,n}$ stands for the mass of firms producing n products in the nontradable sector, and e_{NT} for the mass of potential entrants investing in R&D in the nontradable sector (holding $e_{NT} = x_{NT} - (1-\theta)z_{NT}$).[13]
There is a one-to-one mapping between this distribution and the employment distribution at

[13] The expressions in Equation (2.23) are derived exactly as in the original Klette and Kortum model (see Aghion et al. (2014, p. 537)). I therefore do not provide a proof in this chapter.

the sectoral level, as every product of the same sector has the same employment level. This is not true in the aggregate, as tradable products command more employment than nontradable ones. The respective employment levels needed to produce a product are

$$l_{NT} = \frac{L + \left(1 - \frac{\theta}{2}\right)\rho\varphi}{\gamma} \text{ and } l_T = (1+\tau)\frac{L + \left(1 - \frac{\theta}{2}\right)\rho\varphi}{\gamma}, \tag{2.24}$$

where employment for tradable products includes the labour needed to pay the trade cost. Knowing these employment levels and the size distribution of firms by number of products, the employment size distribution can easily be deduced.

2.3 Predictions

This section discusses my model's main predictions. I first consider permanent differences between the two sectors. Then, I turn to the effect of expansions in international trade, due either to falling trade costs within the tradable sector or to nontradable products becoming tradable.

2.3.1 Permanent differences between sectors

Proposition 1. *$g_T \geq g_{NT}$ and $e_T \geq e_{NT}$, with strict inequality if and only if $\tau < \gamma$.*

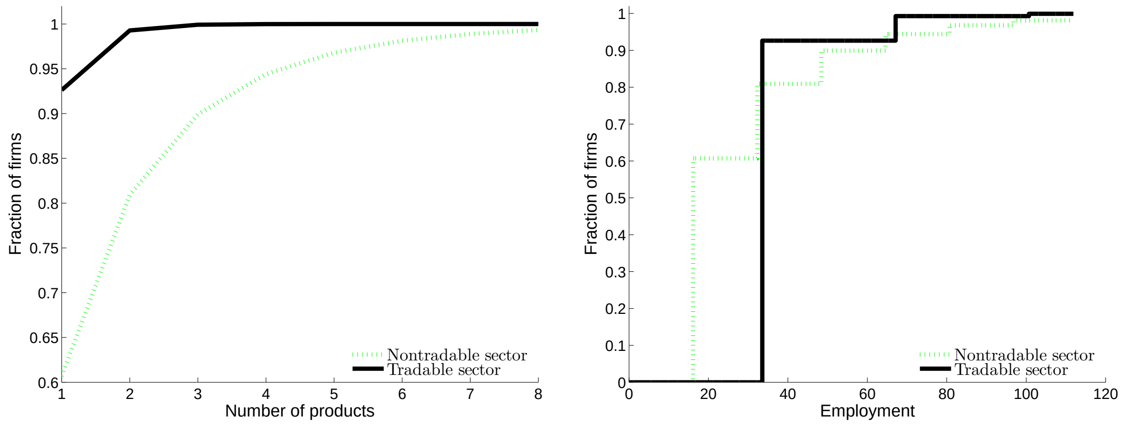
Proof. Follows directly from (2.20) and (2.21). □

Productivity growth in the tradable sector is larger than in the nontradable sector because of a market size effect. A successful innovation in the tradable sector generates more profits for the innovating firm, as it can sell its new product to two markets rather than one. As the entry costs are equal in both sectors, entry (and therefore radical innovation) in the tradable sector must be higher, to balance the potentially higher profits with a higher destruction rate.[14]

As entry is higher in the tradable sector, its size distribution by products is always more left-skewed (see the left panel of Figure 2.1). This is not true for the employment distribution, as a tradable product commands greater employment than a nontradable one. In particular, the smallest employers in any country are always nontradable firms (see the right panel of Figure 2.1).

[14]The destruction rate is the rate at which firms lose their products ($\frac{x_{NT}}{1-\theta}$ in the nontradable sector). Obviously, an increase in the destruction rate lowers the value of being an incumbent firm.

Figure 2.1: Size distributions in the tradable and nontradable sectors



Notes: The graphs show cumulative distribution functions (i.e., the fraction of firms which have less than a given number of products or employment level). Parameter values used are given in Appendix Section 2.6.1.2.

2.3.2 Falling trade costs in the tradable sector

Proposition 2. *Suppose that at time t_0, τ unexpectedly and permanently falls to τ_{new}, holding $\tau_{new} < \tau$.*
1. This does not affect consumption and wages at time t_0.
2. This increases growth in the tradable sector ($g_{T,new} > g_T$), leaves growth in the nontradable sector unchanged ($g_{NT,new} = g_{NT}$), and increases aggregate growth ($g_{new} > g$).

Proof. The first part of the proposition follows from Equation (2.17), while the second part follows from Equations (2.20) to (2.22). □

Falling trade costs increase growth in the tradable sector because they allow exporters to charge higher mark-ups and therefore to earn higher flow profits. These higher flow profits need to be balanced by higher radical innovation and entry, so that the free entry condition continues to hold.

These changes do not affect the nontradable sector, as they do not change the wage-to-GDP ratio (given by Equation (2.16)). Indeed, higher labour demand by tradable sector entrants is exactly offset by lower labour needs for trade costs.[15] Consumption also does not change on impact, as exporters' higher mark-ups are offset by their lower marginal costs.

Higher entry and lower labour needs for trade costs gradually shift the tradable sector's size distributions of firms (both by products and by employment) to the left. While changes in

[15]These two effects cancel out exactly with a Cobb-Douglas utility function. However, with different preferences, a reduction in trade costs may affect the nontradable sector by changing the wage-to-GDP ratio.

the size distribution take some time, growth rates immediately jump to their new balanced growth path values, without transition dynamics.

2.3.3 Increases in the fraction of tradable products

Proposition 3. *Suppose that at time t_0, a mass θ_{new} of products which were previously nontradable unexpectedly and permanently become tradable.*
1. This immediately increases consumption and wages.
2. This decreases productivity growth in the tradable and nontradable sectors ($g_{T,new} < g_T$ and $g_{NT,new} < g_{NT}$). The reaction of aggregate growth depends on the level of trade costs. If $\tau < \gamma - (\gamma - 1)\frac{\rho}{2}\frac{\varphi}{L}$, the aggregate growth rate increases ($g_{new} > g$). If $\frac{2\gamma L + \rho\varphi}{2L + \rho\varphi} < \tau < \gamma$, the aggregate growth rate decreases ($g_{new} < g$). In all other cases, the reaction depends on the exact values of θ and θ_{new}.

Proof. For the first part, see below. The second part of the proposition follows from Equations (2.20) to (2.22). □

An increase in the fraction of tradable products has static and dynamic effects. To determine the static ones, I first focus on the productivity distribution for the newly tradable products. As these were previously nontradable, the numbers of innovations on one particular product in Home and in Foreign during this time (denoted $B(t_0)$ and $B^*(t_0)$) follow independent Poisson distributions with parameter $\frac{t_0 x_{NT}}{1-\theta}$. It can be shown (see Appendix Section 2.6.1.1) that $\lim_{t_0 \to +\infty} \mathbb{P}(B(t_0) = B^*(t_0)) = 0$ and $\lim_{t_0 \to +\infty} \mathbb{P}(B(t_0) > B^*(t_0)) = \frac{1}{2}$. Thus, if t_0 is large (and if both countries had the same productivity levels for nontradable products at $t = 0$), approximately half of the newly tradable products have an absolute advantage for Home. As wages are equal in both countries and knowledge spillovers are local, they will be produced by Home firms forever onwards. This specialization driven by productivity differences increases consumption and wages just as in the classic Ricardian model of Dornbusch, Fischer and Samuelson (Dornbusch et al. (1977)).[16] My model embeds their framework in a more general setting, where the productivity distribution is endogenous and determined by countries' innovation history.

I now turn to the dynamic effects. Productivity growth rates in the tradable and nontradable sector decrease as more products become tradable. This somewhat surprising result is due to an increase in labour demand. Indeed, the firms producing the newly tradable products now

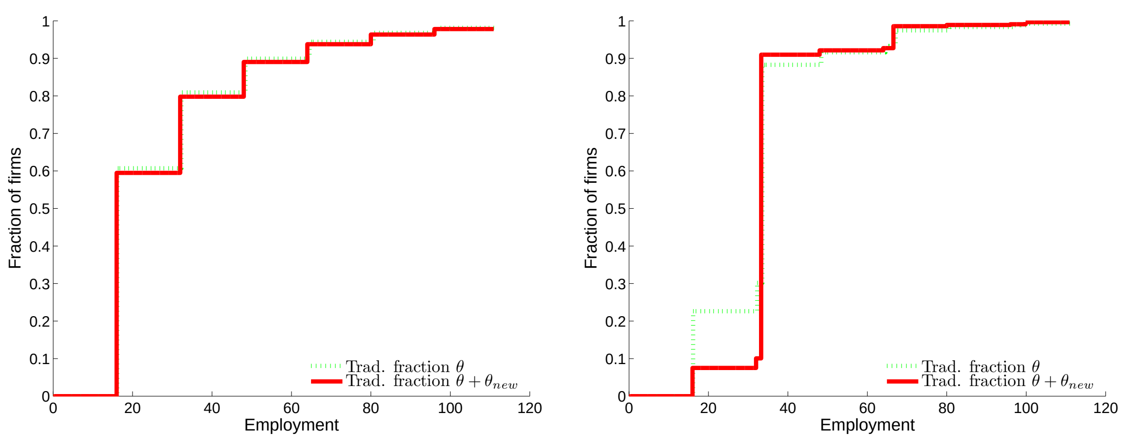
[16] The level of consumption is still given by Equation (2.17), but specialization for the newly tradable products increases $A_T(t_0)$ and $A_T^*(t_0)$.

need to hire labour to pay trade costs, and entrants demand more labour for innovation on these products (because, as shown in Proposition 1, entry and innovation rates in the tradable sector are higher than in the nontradable one). Higher labour demand increases the wage-to-GDP ratio, and this discourages innovation, because its costs (payable in wages) increase with respect to its payoffs (proportional to GDP). Thus, within every sector, entry rates fall and the size distribution of firms gradually shifts to the right.

However, even though the pace of innovation slows down in the tradable and nontradable sectors taken as a whole, it increases for the newly tradable products, as they switch from the nontradable to the higher tradable balanced growth rate. The aggregate effect of these changes is ambiguous and depends on parameter values. In particular, when the level of trade costs τ is low, exporters' profits and innovation incentives in the tradable sector are high, and not much labour is needed to pay for the newly arising trade costs. Then, the innovation-enhancing effect dominates. However, when τ is high, exporters' profits are low and a lot of labour needs to be spent on trade costs, so the negative effect dominates.[17]

The aggregate size distribution is affected by several forces. As the nontradable sector shrinks, the share of the very smallest firms in each country falls (see the right panel of Figure 2.2). However, various other mechanisms act on the distribution. Most importantly, it is affected by the change in aggregate entry, which is one-to-one related to the change in aggregate productivity growth. Moreover, the fall in the employment level for each product (due to the increasing wage-to-GDP ratio) tends to shift the distribution to the left.

Figure 2.2: Changes in the employment size distribution



Nontradable sector | Aggregate economy

Notes: See Figure 2.1.

[17] Note also that when L is sufficiently large, aggregate productivity growth always increases. Indeed, the wage-to-GDP ratio is essentially equivalent to L when L is large, and therefore does not change much when the tradable sector expands.

2.3.4 International knowledge spillovers

In the baseline version of my model, I assumed that innovators could only improve on domestically produced products. In this section, I analyse an alternative model with "unconditional" international knowledge spillovers. In this alternative model, innovations in both sectors apply to a randomly chosen product (irrespective of whether that product is currently produced in Home or in Foreign) and improve on the highest productivity level worldwide.

International knowledge spillovers do not change the problem of a firm in the nontradable sector. Indeed, neither the profits from innovation (which are independent of the productivity level of the improved product) nor the probability to be displaced are affected by spillovers. Therefore, the firm's problem is still defined by (2.8), and the innovation arrival rate in the Home nontradable sector is still given by Equation (2.12).
The problem of a tradable firm also does not change. Of course, the firm can now be displaced by Foreign firms, so the arrival rate of innovations in its sector becomes $x_T + x_T^*$ ($2x_T$ in the symmetric equilibrium). However, these innovations also apply to a larger mass of products θ, so the destruction rate for any particular product is unchanged at $\frac{2x_T}{\theta}$. Intuitively, the greater likelihood that Foreign innovations affect a Home firm is exactly offset by the greater likelihood that Home innovations do not affect it. Thus, the worldwide arrival rate of innovations in the tradable sector, $2x_T$, is still given by Equation (2.15).
Private decisions being unchanged, the labour market clearing condition is also unaffected, and the equilibrium wage-to-GDP ratio is still given by Equation (2.16). International knowledge spillovers are therefore neutral for the tradable sector's growth rate, which is still given by Equation (2.21). However, they increase growth in the nontradable sector: the arrival rate for improvements in the productivity of a nontradable product is now $2x_{NT}$, so that

$$\frac{g_{NT}^{IKS}}{\ln \gamma} = 2\left(\left(1-\frac{1}{\gamma}\right)\left(\frac{L}{\varphi}-\frac{\theta}{2}\rho\right)-\frac{\rho}{\gamma}+\frac{\eta-1}{\eta}z\right). \tag{2.25}$$

To which extent the predictions of the previous sections still hold in this alternative model? Proposition 2 clearly remains true: a fall in trade costs in the tradable sector still increases innovation in that sector. However, it is not guaranteed any more that productivity growth in the tradable sector is higher than in the nontradable sector. While the tradable sector still benefits from greater market size, international knowledge spillovers allow the nontradable sector to increase productivity growth without increasing the destruction rate and thereby lowering R&D incentives.[18] The market size effect dominates, and g_T remains higher than

[18] Indeed, in the nontradable sector, a Home firm can improve on a Foreign firm's productivity without dis-

g_{NT} if and only if

$$\rho > \frac{\tau-1}{\gamma}\left(\frac{L}{\varphi}+\left(1-\frac{\theta}{2}\right)\rho\right)+\frac{\eta-1}{\eta}\left(\frac{\varphi}{\zeta\eta}\right)^{\frac{1}{\eta-1}}. \tag{2.26}$$

Under this condition, Proposition 3 remains valid (with different thresholds for trade costs). Otherwise, increases in the fraction of tradable products unambiguously reduce aggregate growth, as they raise the wage-to-GDP ratio and redistribute activity to a low-growth sector.

Thus, in the presence of international knowledge spillovers, it is in general less likely that an expansion in international trade stimulates productivity growth. Note, however, that I have assumed a very strong version of knowledge spillovers, and that the continuous time structure of the model implies that there is no overlap between nontradable innovations in both countries. Assuming spillovers are limited in the absence of trade[19], or that Home and Foreign nontradable innovations may overlap, would lower g_{NT}^{IKS} and bring this alternative model closer to the baseline version.

2.4 Empirical evidence

2.4.1 Increases in the fraction of tradable products: the case of services

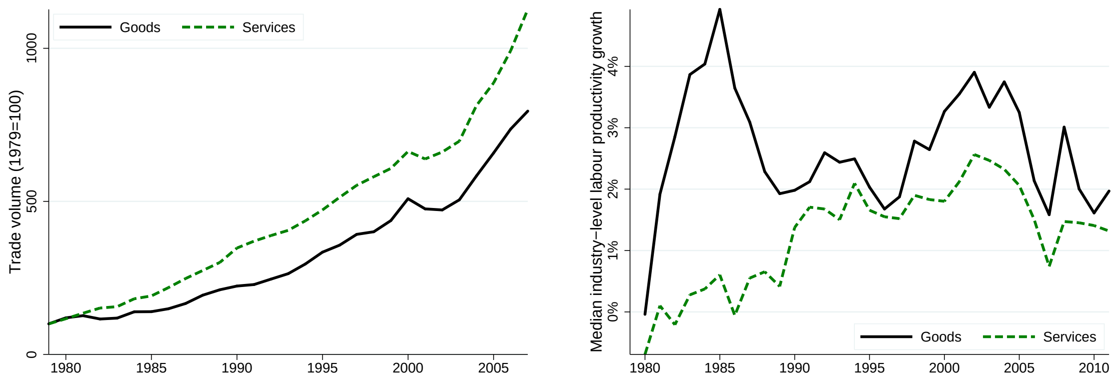
To compare my model's predictions from Propositions 1 and 3 to the data, I need to identify a set of products which at a certain point in time have switched from being nontradable to becoming tradable.

The products of a line of service industries are natural candidates for this exercise. Indeed, in the 1980s and 1990s, advances in information technology and targeted trade liberalizations made a wide range of services tradable for the first time in history (Francois and Hoekman (2010)). Accordingly, the left panel of Figure 2.3 shows that in the United States, trade in services has outpaced trade in goods, especially in the 1980s and early 1990s.

placing it: even if it can produce a product more efficiently than its foreign competitors, it cannot export.

[19]Instead of the "unconditional" spillovers assumed in this section, one could consider a model of "conditional" spillovers, in which Home innovators can improve on the productivity of a Foreign firm if and only if that firm sells on the Home market. Then, by definition, spillovers do not apply to the nontradable sector. In the tradable sector, firms now face a trade-off when deciding whether to export: selling in Foreign increases their profits, but also provides Foreign firms with an opportunity to displace them. It can be shown that if trade costs are lower than some threshold $\overline{\tau}$, all firms still export. Then, innovation incentives in the tradable sector are identical to the baseline model, and the equilibrium of the conditional spillover model coincides with the baseline. When trade costs are very high, however, profits from exporting are so low that the threat of displacement makes all products effectively nontraded and locks up countries on a low growth path.

Figure 2.3: Trade and labour productivity growth in goods and service industries



Sources: BEA and own calculations. Trade in goods or services is defined as the sum of nominal exports and imports in a given year, taken from the BEA's International Transaction Database. See Appendix Section 2.6.2.1 for details on the labour productivity measure.

The right panel of Figure 2.3 shows the five-year moving average of the median annual growth rate of labour productivity in goods and service industries. Goods-producing industries are made up by agriculture, mining and manufacturing, while services contain the remainder of the economy, excluding health care, education and government.[20] Labour productivity is measured as real value added by employee, and I report five-year moving averages in order to smooth out business cycle fluctuations.[21]

In the beginning of the 1980s, when most services were still nontradable, labour productivity growth rates were in line with Proposition 1: the median growth rate for (tradable) goods-producing industries considerably exceeded the one for service industries. However, as services became more tradable, their productivity growth rate greatly increased, while the one of goods-producing industries had a slightly decreasing trend since the mid-1980s, just as predicted by Proposition 3. As a result, productivity growth rates in services and goods-producing industries have become almost identical in the 2000s.[22]

Figure 2.4 provides further disaggregation. Indeed, while the products of some service in-

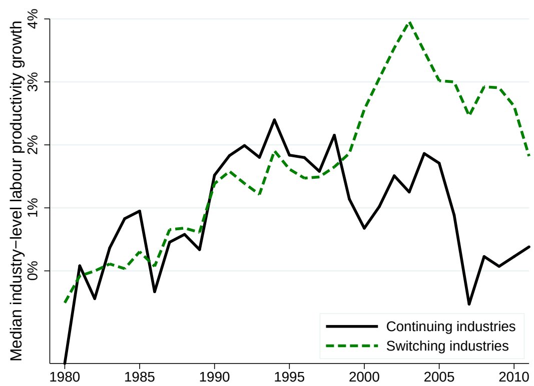
[20]Industries are defined by the BEA and roughly correspond to the NAICS 3-digit level. There are 24 goods-producing and 31 service industries (a full list is provided in Appendix Section 2.6.2.2).

[21]I consider the median of industry productivity growth rates because averages or aggregates are driven by outliers. Houseman et al. (2014) have noted that almost all manufacturing productivity growth between 1997 and 2007 is due to only one industry, Computer and Electronic Product Manufacturing, and warn against generalizing from its exceptional performance to the remainder of the manufacturing sector.

[22]This pattern does not change if I consider total factor productivity as estimated by Jorgenson et al. (2012, available at http://www.worldklems.net/data.htm) instead of labour productivity. The acceleration of productivity growth in services has been pointed out before (for instance, by Triplett and Bosworth (2003)). My model offers a potential explanation for this phenomenon.

dustries switched from nontradable to tradable in the 1980s and 1990s, other services (such as transports) had been tradable for a long time, while still others (such as barber shops) remained nontradable. I therefore compare the set of "ICT-enabled tradable services" (defined by the United Nations' Conference on Trade and Development (2008)), which can be considered to have switched from nontradable to tradable, to the remainder of the service sector.[23] Figure 2.4 suggests that only the switching industries settled on a permanently higher balanced growth path for productivity after the 1980s.[24]

Figure 2.4: The labour productivity growth experience of different service industries



Sources: BEA and own calculations. See Appendix Section 2.6.2.2 for a list of continuing and switching industries.

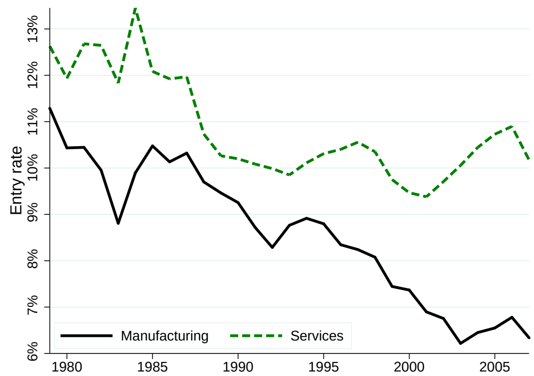
In my model, productivity growth is one-to-one related to entry. Figure 2.5 plots the evolution of entry rates since the 1980s for manufacturing and services.[25] As has been highlighted by Decker et al. (2014), entry rates have been falling in all industries. This trend cannot be explained by my model. However, Proposition 3 does provide a potential explanation for why the fall of the entry rate in services has been much smaller than the fall of the entry rate in manufacturing.

[23]These services include Communication, Finance and insurance, Computer and information services, royalties and licence fees, other business services, and personal, cultural and recreational services. Unfortunately, it is impossible to map this classification one-to-one to the industries defined by the BEA. I report my necessarily subjective mapping in Appendix Section 2.6.2.2.

[24]The evidence in this section is also in line with industry-level R&D investment data provided by the BEA (available at http://www.bea.gov/national/rd.htm.) Software publishing, Finance and insurance, and R&D services, which are switching industries, have all had an annual growth rate of real R&D investment of at least 11% between 1987 and 2007. The corresponding number for the remaining service industries was below 7%, and for manufacturing, just 2%.

[25]This figure uses data from the US Census Bureau, which has a much narrower definition of services than the one used in Figures 2.3 and 2.4. In particular, the Census Bureau excludes retail, transport and finance from the service sector. However, the productivity growth pattern for the Census service industries is the same than the one shown in Figure 2.3.

Figure 2.5: Entry rates in manufacturing and services



Source: US Census Bureau, Center for Economic Studies, Business Dynamics Statistics. The entry rate for a sector is defined as the ratio between the number of firms of age 0 and the total number of firms in a year.

2.4.2 Falling trade costs in the tradable sector

Finally, in this section, I test Proposition 2, which predicts that a fall in trade costs leads to an acceleration of productivity growth in tradable industries. I build on the work of Bernard et al. (2003), who calculated trade cost measures for every year between 1974 and 1999 at the four-digit SIC level for manufacturing industries, and combined this information with the NBER-CES Manufacturing Database. They show that a fall in trade costs between years $t-5$ and t is positively correlated with productivity growth between years t and $t+5$, and interpret this as supporting evidence for the Melitz (2003) model.[26]

I modify their estimation strategy slightly and estimate[27]

$$\forall t \in \{1979, 1980, ..., 1989\}, \quad \overline{g}_k^{t+5,t+10} - \overline{g}_k^{t-5,t} = \beta \left(\tau_{k,t-5} - \tau_{k,t} \right) + \delta_k + \delta_t + \varepsilon_{k,t}, \quad (2.27)$$

where $\overline{g}_k^{t_1,t_2}$ stands for the average productivity growth rate between t_1 and t_2 in industry k, and $\tau_{k,t}$ for the level of trade costs in year t. δ_k is an industry fixed effect and δ_t a year fixed effect. Thus, instead of analysing like Bernard et al. whether a fall in trade costs raises productivity in the short run (which could be explained by the static effect of greater specialization or selection), I analyse whether it is correlated with an increase in the balanced growth path of

[26]Their measure of trade costs mainly captures transport costs and tariffs, and is expressed as a fraction of industry sales. Productivity growth is calculated with respect to the total factor productivity measure provided in the NBER-CES Manufacturing Database (Becker et al. (2013)).

[27]I implicitly assume that all manufacturing industries were tradable in 1974. However, my results do not change if I drop the 10% of industries which had the lowest combined value of imports and exports in 1974.

the industry, that is, with a durable increase in average productivity growth.
Results are shown in Table 2.2. They show a strong positive correlation between trade cost reductions and increases in the productivity growth rate years ten years later.[28] This suggests, in line with my model, that falling trade costs increased the balanced growth rates and not only had a static effect.

Table 2.2: Trade cost reductions and productivity growth

	Change in av. TFP growth $\left(\overline{g}_k^{t+5,t+10} - \overline{g}_k^{t-5,t}\right)$
Trade cost reduction $(\tau_{k,t-5} - \tau_{k,t})$	0.076^{***}
R^2	0.39
Observations	4226

Notes: *** significant at 1%, ** significant at 5%, * significant at 10%. Coefficients for time and industry dummies are not shown. Standard errors are clustered by industry.

2.5 Conclusions

In this chapter, I extended the canonical Klette and Kortum (2004) model to an international setting, in order to analyse how expansions in international trade affect creative destruction and productivity growth. My model predicts that if an expansion in trade is due to falling trade costs in the tradable sector, it permanently increases productivity growth rates in that sector. An increase in the fraction of tradable products has more ambivalent effects. These predictions are consistent with industry-level evidence from the United States.
In the future, the empirical evidence presented in this chapter could be completed using firm-level data, both to test the model's firm-level predictions and to calibrate it in order to generate quantitative results. From a theoretical standpoint, it would also be interesting to explore the effects of trade between asymmetric countries.

[28]The results in Table 2.2 are robust to excluding Computer industries (SIC Codes 3570 to 3579).

2.6 Appendix to Chapter 2

2.6.1 Theoretical Appendix

2.6.1.1 Poisson innovation processes and absolute advantage

To simplify notation, I denote $\phi \equiv \frac{t_0 x_{NT}}{1-\theta}$. As $B(t_0)$ and $B^*(t_0)$ follow independent Poisson distributions with parameter ϕ,

$$\begin{aligned}\mathbb{P}(B(t_0) = B^*(t_0)) &= \sum_{n=0}^{+\infty} \mathbb{P}((B(t_0) = n) \cap (B^*(t_0) = n)) \\ &= \exp(-2\phi) \sum_{n=0}^{+\infty} \frac{\phi^{2n}}{n!n!}.\end{aligned}$$

As $\lim_{n\to+\infty} \frac{\phi^n}{n!} = 0$, there exists an integer n^* such that for every n, $\frac{\phi^n}{n!} \leq \frac{\phi^{n^*}}{n^*!}$. Thus,

$$\begin{aligned}\mathbb{P}(B(t_0) = B^*(t_0)) &\leq \exp(-2\phi) \frac{\phi^{n^*}}{n^*!} \left(\sum_{n=0}^{+\infty} \frac{\phi^n}{n!}\right). \\ &= \exp(-\phi) \frac{\phi^{n^*}}{n^*!}.\end{aligned}$$

However, as $\lim_{\phi\to+\infty} \exp(-\phi)\phi^{n^*} = 0$ for every integer n^*, this implies

$$\lim_{\phi\to+\infty} \mathbb{P}(B(t_0) = B^*(t_0)) = 0.$$

Finally, by symmetry, $\mathbb{P}(B(t_0) > B^*(t_0)) = \mathbb{P}(B(t_0) < B^*(t_0))$. This directly implies

$$\lim_{\phi\to+\infty} \mathbb{P}(B(t_0) > B^*(t_0)) = \frac{1}{2}.$$

2.6.1.2 Parameter values

Table 2.3: Parameter values

γ	1.1	η	1.1
τ	1.08	ζ	1
θ	0.5	L	17
θ_{new}	0.3	ρ	1
φ	1		

2.6.2 Data Appendix

2.6.2.1 Labour productivity data

The labour productivity of an industry is defined as the ratio of its real value added and its total employment. Data for real value added is taken from the BEA Industry Accounts (available at http://www.bea.gov/industry/gdpbyind_data.htm), and available for the period 1977-2013 (with a structural break in 1998). Employment is defined as the total number of full and part-time employees and is taken from NIPA Table 6.4 for the years 1998-2013, and from historical time series available at http://www.bea.gov/industry/NAICSemployment_datarelease.htm for the years 1977-1997.

To construct the time series shown in Figures 2.3 and 2.4, I calculate labour productivity growth for every year of the sample (except for 1998, where there is a structural break for all data series). I then determine the median growth rate for goods-producing and service industries (listed in the next section). Finally, I smooth the time series by taking five-year moving averages. That is, the number reported for year t is the average of the median growth rates between year $t-2$ and year $t+2$.

2.6.2.2 The BEA industry classification

The following tables list the goods-producing and service industries used in Figures 2.3 and 2.4. Service industries which are considered to have switched from nontradable to tradable are marked with an S.

NAICS code	Goods-producing industries (24 industries)
111, 112	Farms
113, 114, 115	Forestry, fishing, and related activities
211	Oil and gas extraction
212	Mining, except oil and gas
213	Support activities for mining
321	Wood products
327	Nonmetallic mineral products
331	Primary metals
332	Fabricated metal products
333	Machinery
334	Computer and electronic products
335	Electrical equipment, appliances, and components

NAICS code	Goods-producing industries (continued)
3361-3363	Motor vehicles, bodies and trailers, and parts
3364-3369	Other transportation equipment
337	Furniture and related products
339	Miscellaneous manufacturing
311-312	Food and beverage and tobacco products
313-314	Textile mills and textile product mills
315-316	Apparel and leather and allied products
322	Paper products
323	Printing and related support activities
324	Petroleum and coal products
325	Chemical products
326	Plastics and rubber products

NAICS code	Service producing industries (31 industries)
42	Wholesale trade (S)
44-45	Retail trade (S)
481	Air transportation
482	Rail transportation
483	Water transportation
484	Truck transportation
485	Transit and ground passenger transportation
486	Pipeline transportation
487-492	Other transportation and support activities
493	Warehousing and storage
511	Publishing industries (includes software) (S)
512	Motion picture and sound recording industries (S)
513	Broadcasting and telecommunications (S)
514	Information and data processing services (S)
521, 522	Federal Reserve banks, credit intermediation, and related activities (S)
523	Securities, commodity contracts, and investments (S)
524	Insurance carriers and related activities (S)
525	Funds, trusts, and other financial vehicles (S)
531	Real estate

NAICS code	Service producing industries (31 industries)
532, 533	Rental and leasing services and lessors of intangible assets (S)
5411	Legal services (S)
5415	Computer systems design and related services (S)
5412-5414, 5416-5419	Miscellaneous professional, scientific, and technical services (S)
551	Management of companies and enterprises (S)
561	Administrative and support services (S)
562	Waste management and remediation services
711, 712	Performing arts, spectator sports, museums, and related activities
713	Amusements, gambling, and recreation industries (S)
721	Accommodation
722	Food services and drinking places
81	Other services, except government

The only industries which are excluded from this listing are Utilities (NAICS Code 22), Construction (23), Educational services, health care, and social assistance (6) and Public administration (92).

Bibliography

[1] **Abraham, Katharine G. and Lawrence F. Katz**, "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?," *Journal of Political Economy*, 1986, *94* (3), pp. 507–522.

[2] **Acemoğlu, Daron, Ufuk Akçiğit, Nicholas Bloom, and William R. Kerr**, "Innovation, Reallocation and Growth," NBER Working Paper 18993, April 2013.

[3] **Acs, Zoltan J. and David B. Audretsch**, *Innovation and Small Firms*, MIT Press, December 1990.

[4] **Aghion, Philippe and Gilles Saint-Paul**, "Uncovering Some Causal Relationships between Productivity Growth and the Structure of Economic Fluctuations: A Tentative Survey," NBER Working Paper 4603, December 1993.

[5] ___ **and Peter Howitt**, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–351.

[6] ___ **, George-Marios Angeletos, Abhijit Banerjee, and Kalina Manova**, "Volatility and Growth: Credit Constraints and the Composition of Investment," *Journal of Monetary Economics*, 2010, *57* (3), 246 – 265.

[7] ___ **, Philippe Askenazy, Nicolas Berman, Gilbert Cette, and Laurent Eymard**, "Credit Constraints and the Cyclicality of R&D Investment: Evidence from France," *Journal of the European Economic Association*, 2012, *10* (5), 1001–1024.

[8] ___ **, Ufuk Akçiğit, and Peter Howitt**, "What Do We Learn From Schumpeterian Growth Theory?," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Elsevier, June 2014, pp. 515–563.

[9] **Akçiğit, Ufuk and William R. Kerr**, "Growth Through Heterogeneous Innovations," NBER Working Paper 16443, November 2010.

[10] **Ateş, Sînâ T. and Felipe E. Saffie**, "Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection," Working Paper 14-043, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania, November 2014.

[11] **Atkeson, Andrew and Ariel T. Burstein**, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, June 2010, *118* (3), 433–484.

[12] **Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu**, "R&D Investment, Exporting, and Productivity Dynamics," *American Economic Review*, June 2011, *101* (4), 1312–1344.

[13] **Baldwin, Richard E. and Frédéric Robert-Nicoud**, "Trade and growth with heterogeneous firms," *Journal of International Economics*, January 2008, *74* (1), 21–34.

[14] **Ball, Laurence M.**, "Long-Term Damage from the Great Recession in OECD Countries," NBER Working Paper 20185, May 2014.

[15] **Barlevy, Gadi**, "On the Cyclicality of Research and Development," *American Economic Review*, September 2007, *97* (4), 1131–1164.

[16] **Baumol, William J.**, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *The American Economic Review*, 1967, *57* (3), pp. 415–426.

[17] **Becker, Randy, Wayne Gray, and Jordan Marvakov**, *NBER-CES Manufacturing Industry Database*, http://www.nber.org/nberces/, June 2013.

[18] **Bernanke, Ben and Mark Gertler**, "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, March 1989, *79* (1), 14–31.

[19] **Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott**, "Falling Trade Costs, Heterogeneous Firms, and Industry Dynamics," NBER Working Paper 9639, April 2003.

[20] **Bloom, Nicholas, Paul M. Romer, Stephen J. Terry, and John Van Reenen**, "Trapped Factors and China's Impact on Global Growth," NBER Working Paper 19951, March 2014.

[21] **Bustos, Paula**, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, February 2011, *101* (1), 304–40.

[22] **Caballero, Ricardo J. and Mohamad L. Hammour**, "The Cost of Recessions Revisited: A Reverse-Liquidationist View," *The Review of Economic Studies*, 2005, *72* (2), pp. 313–341.

[23] **Cohen, Wesley M. and Steven Klepper**, "A Reprise of Size and R&D," *The Economic Journal*, 1996, *106* (437), pp. 925–951.

[24] **Comin, Diego and Mark Gertler**, "Medium-Term Business Cycles," *American Economic Review*, June 2006, *96* (3), 523–551.

[25] **Davis, Steven, John Haltiwanger, and Scott Schuh**, *Job Creation and Destruction*, MIT Press, 1998.

[26] **Decker, Ryan, John Haltiwanger, Ron S. Jarmin, and Javier Miranda**, "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 2014, *28* (3), 3–24.

[27] **Dornbusch, Rüdiger, Stanley Fischer, and Paul A. Samuelson**, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, December 1977, *67* (5), 823–839.

[28] **Eaton, Jonathan and Samuel Kortum**, "Technology, trade, and growth: A unified framework," *European Economic Review*, May 2001, *45* (4-6), 742–755.

[29] **Epifani, Paolo and Gino Gancia**, "Trade, markup heterogeneity and misallocations," *Journal of International Economics*, January 2011, *83* (1), 1–13.

[30] **Ewens, Michael and Christian Fons-Rosen**, "The Consequences of Entrepreneurial Firm Founding on Innovation," Working Paper, 2013.

[31] **Farre-Mensa, Joan and Alexander Ljungqvist**, "Do Measures of Financial Constraints Measure Financial Constraints?," NBER Working Paper 19551, October 2013.

[32] **Fernald, John**, "Productivity and Potential Output Before, During, and After the Great Recession," in "NBER Macroeconomics Annual 2014, Volume 29" October 2014.

[33] **Fort, Teresa C., John Haltiwanger, Ron S. Jarmin, and Javier Miranda**, "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," *IMF Economic Review*, August 2013, *61* (3), 520–559.

[34] **Francois, Joseph and Bernard Hoekman**, "Services Trade and Policy," *Journal of Economic Literature*, September 2010, *48* (3), 642–692.

[35] **Francois, Patrick and Huw Lloyd-Ellis**, "Animal Spirits Through Creative Destruction," *American Economic Review*, June 2003, *93* (3), 530–550.

[36] **Gabaix, Xavier**, "The Granular Origins of Aggregate Fluctuations," *Econometrica*, May 2011, *79* (3), 733–772.

[37] **Garcia-Macià, Daniel**, "The Financing of Ideas and The Great Deviation," Working Paper, 2014.

[38] **Gertler, Mark and Simon Gilchrist**, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *The Quarterly Journal of Economics*, May 1994, *109* (2), 309–340.

[39] **Grossman, Gene M. and Elhanan Helpman**, *Innovation and Growth in the Global Economy*, MIT Press, 1991.

[40] ____ **and** ____ , "Globalization and Growth," *American Economic Review, Papers & Proceedings*, May 2015, *105* (5), 100–104.

[41] **Gustafsson, Peter and Paul Segerstrom**, "Trade Liberalization and Productivity Growth," *Review of International Economics*, May 2010, *18* (2), 207–228.

[42] **Hall, Bronwyn H.**, "R&D and Innovation Expenditures in the Crisis," Presentation to the OECD workshop on R&D and innovation in the current macroeconomic climate, Paris 2011.

[43] **Hall, Robert E.**, "Quantifying the Lasting Harm to the U.S. Economy from the Financial Crisis," NBER Working Paper 20183, May 2014.

[44] **Hicks, John R.**, "Mr. Keynes and the "Classics"; A Suggested Interpretation," *Econometrica*, April 1937, *5* (2), 147–159.

[45] **Houseman, Susan N., Timothy J. Bartik, and Timothy J. Sturgeon**, "Measuring Manufacturing: How the Computer and Semiconductor Industries Affect the Numbers and Perceptions," Upjohn Working Papers and Journal Articles 14-209, W.E. Upjohn Institute for Employment Research, February 2014.

[46] **Hsieh, Chang-Tai and Peter J. Klenow**, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

[47] **Imbs, Jean**, "Growth and volatility," *Journal of Monetary Economics*, October 2007, *54* (7), 1848–1862.

[48] **Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson**, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, August 1993, *108* (3), 577–598.

[49] **Jorgenson, Dale W., Mun S. Ho, and Jon D. Samuels**, "Information Technology and US Productivity Growth: Evidence from a Prototype Industry Production Account," in "Industrial Productivity in Europe," Edward Elgar, 2012.

[50] **Keynes, John Maynard**, *The General Theory of Employment, Interest and Money*, Palgrave Macmillan, 1936.

[51] **Khan, Aubhik and Julia K. Thomas**, "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 2013, *121* (6), 1055 – 1107.

[52] **Klapper, Leora, Inessa Love, and Douglas Randall**, "New firm registration and the business cycle," Policy Research Working Paper Series 6775, The World Bank 2014.

[53] **Klette, Tor Jakob and Samuel Kortum**, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, October 2004, *112* (5), 986–1018.

[54] **Kortum, Samuel and Josh Lerner**, "Assessing the Contribution of Venture Capital to Innovation," *RAND Journal of Economics*, Winter 2000, *31* (4), 674–692.

[55] **Krueger, Alan B. and Sarah Charnes**, "JOLTS as a timely source of data by establishment size," *Monthly Labor Review*, May 2011.

[56] **Krusell, Per and Anthony A. Smith**, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

[57] **Lileeva, Alla and Daniel Trefler**, "Improved Access to Foreign Markets Raises Plant-Level Productivity... for Some Plants," *The Quarterly Journal of Economics*, August 2010, *125* (3), 1051–1099.

[58] **Matsuyama, Kiminori**, "Growing Through Cycles," *Econometrica*, March 1999, *67* (2), 335–348.

[59] **Melitz, Marc J.**, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

[60] **Moscarini, Giuseppe and Fabien Postel-Vinay**, "The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment," *American Economic Review*, October 2012, *102* (6), 2509–39.

[61] **Murata, Yasusada, Ryo Nakajima, Ryosuke Okamoto, and Ryuichi Tamura**, "Localized Knowledge Spillovers and Patent Citations: A Distance-Based Approach," *The Review of Economics and Statistics*, December 2014, *96* (5), 967–985.

[62] **OECD**, "Innovation in the crisis and beyond," *OECD Science, Technology and Industry Outlook 2012*, Chapter 1, 2012, pp. 21–57.

[63] **Ouyang, Min**, "On the Cyclicality of R&D," *The Review of Economics and Statistics*, May 2011, *93* (2), 542–553.

[64] **Paunov, Caroline**, "The global crisis and firms' investments in innovation," *Research Policy*, 2012, *41* (1), 24–35.

[65] **Pérez, Ander**, "Credit Constraints, Firms' Precautionary Investment, and the Business Cycle," Working Paper 1237, Universitat Pompeu Fabra, 2012.

[66] **Perla, Jesse, Christopher Tonetti, and Michael E. Waugh**, "Equilibrium Technology Diffusion, Trade, and Growth," NBER Working Paper 20881, January 2015.

[67] **Peters, Bettina**, "Persistence of innovation: stylised facts and panel data evidence," *The Journal of Technology Transfer*, April 2009, *34* (2), 226–243.

[68] **Peters, Michael**, "Heterogeneous Mark-Ups and Endogenous Misallocation," Working Paper, 2011.

[69] **Rammer, Christian, Birgit Aschhoff, Dirk Crass, Thorsten Doherr, Martin Hud, Christian Köhler, Bettina Peters, Torben Schubert, and Franz Schwiebacher**, "Innovationsverhalten der deutschen Wirtschaft," , ZEW Mannheim, 2010, 2011 and 2012.

[70] **Rivera-Batiz, Luis A. and Paul M. Romer**, "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, May 1991, *106* (2), 531–55.

[71] **Romer, Paul M.**, “Endogenous Technological Change,” *Journal of Political Economy*, October 1990, *98* (5), S71–102.

[72] **Schumpeter, Joseph A.**, *The Theory of Economic Development*, Harvard Economics Studies, 1934.

[73] ——— , *Capitalism, Socialism, and Democracy*, Harper, New York, 1942.

[74] **Sedláček, Petr and Vincent Sterk**, “Recession Scars and the Growth of Newborn Firms in General Equilibrium,” Working Paper, 2013.

[75] **Shleifer, Andrei**, “Implementation Cycles,” *Journal of Political Economy*, December 1986, *94* (6), 1163–90.

[76] **Siemer, Michael**, “Firm Entry and Employment Dynamics in the Great Recession,” Finance and Economics Discussion Paper 2014-56, Board of Governors of the Federal Reserve System (U.S.) July 2014.

[77] **Steinwender, Claudia**, “The Roles of Import Competition and Export Opportunities for Technical Change,” CEP Discussion Paper 1334, 2015.

[78] **Stiglitz, Joseph E.**, “Endogenous Growth and Cycles,” NBER Working Paper 4286, March 1993.

[79] **Syverson, Chad**, “What Determines Productivity?,” *Journal of Economic Literature*, 2011, *49* (2), 326–365.

[80] **Triplett, Jack E. and Barry P. Bosworth**, “Productivity measurement issues in services industries: "Baumol’s disease" has been cured,” *Economic Policy Review*, 2003, (Sep), 23–33.

[81] **UNCTAD**, “Information Economy Report 2007-2008. Science and technology for development: the new paradigm of ICT,” 2008.