What Makes Firms Stop Doing R&D in Switzerland? – Project Commissioned by SERI

ETH zürich

KOF Swiss Economic Institute

What Makes Firms Stop Doing R&D in Switzerland? – Project Commissioned by SERI

Michael D. König, Andrin Spescha, Martin Wörter and Sabien Dobbelaere

KOF Studies, No. 169, April 2022

KOF

Imprint

Editor

KOF Swiss Economic Institute, ETH Zurich

Authors

Michael D. König
Andrin Spescha
Martin Wörter
Sabien Dobbelaere

KOF

ETH Zurich
KOF Swiss Economic Institute
LEE G 116
Leonhardstrasse 21
8092 Zurich, Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

What Makes Firms Stop Doing R&D in Switzerland?
Project Commissioned by SERI

Sabien Dobbelaere[*], Michael D. König[†], Andrin Spescha[‡] and Martin Wörter[§]

Contents

[*]Vrije Universiteit Amsterdam. Email: sabien.dobbelaere@vu.nl

[†]KOF Swiss Economic Institute. Email: koenig@kof.ethz.ch.

[‡]KOF Swiss Economic Institute. Email: spescha@kof.ethz.ch.

[§]KOF Swiss Economic Institute. Email: woerter@kof.ethz.ch.

1 Executive Summary

The fraction of companies with R&D activities has changed significantly over time, not only in Switzerland but also in many other European countries. While we are observing a sharp decline in the share of R&D active firms in Switzerland and Germany, we are seeing an increase in this share in the Netherlands, Austria, and Finland, for example. Against the background that simultaneously the sales share of R&D expenditures has risen sharply in both Germany and Switzerland, we are observing a strong concentration of R&D activities in these two countries. This descriptive finding naturally raises questions about the economic consequences of such a structural development. This is a complex question that we want to address in this study.

On the basis of a "survival analysis", this study shows that the ability of a company to carry out R&D over a longer period of time is characterized by several factors. The extent of a company's **past innovation efforts** and its **competitiveness** are the most important features for the survival probability in the R&D markets. While competitive companies had a 70% probability of continuing R&D until 2017, this probability increases to 80% if the company is not only competitive but also one of the top innovators. Although sector affiliation (high-tech manufacturing, low-tech manufacturing, modern services, traditional services) impacts the probability to maintain R&D over a longer time span, competitiveness is clearly more important.

More specifically, the **company's size** is a decisive feature for the survival probability in the R&D markets. Larger companies tend to have, for example, so-called "complementary" factors, such as international sales channels or extensive marketing departments, which increase the profitability of R&D activities and thus reduce the probability of exiting R&D. The fraction of **well-educated employees** increases the probability to remain R&D active, since it augments the "absorptive capacity" of a company, which allows it to better understand and absorb external relevant knowledge for its innovation activities. **Access to international markets** (export activities) is also a key feature for R&D "survivors". R&D is expensive and involves high fixed costs. Access to large markets increases the commercialization possibilities of the innovative product and thus the growth prospects of the company. Fixed costs are spread over a larger output and reduce the risk of commercial success of the R&D efforts, which in turn significantly reduces the probability of exiting the R&D markets. A company's **technological potential** indicates the worldwide privately and publicly available technological knowledge available to the firm for bringing about marketable innovation. A high technological potential means that the company can draw on a large knowledge base (including basic research) in its R&D activities, which can reduce the technical risk of R&D projects. According to our analysis, this characteristic also significantly increases the probability of survival in the R&D markets. Finally, the majority of a company's R&D is financed from internal resources. The extent of internal funding is strongly related to the productivity of a company, especially since external investors are usually less inclined to finance risky R&D projects. The "survival analysis" shows accordingly that **labor productivity** is also an important element for the survival of a company in the R&D markets.

In contrast, companies show a significantly higher probability of **exiting** the R&D markets if they suffer from a **lack of equity** in their innovation activities and if their innovative products can be **copied too quickly and easily**. This deprives innovative companies of their "first-mover" advantage and makes innovation efforts less profitable.

Company characteristics that are important for the continuation of R&D activities are also the characteristics that increase the probability that a company will start conducting R&D. There is one exception to this rule, however. Productive, non-R&D performing firms have a significantly lower probability of entering R&D markets. One explanation for this finding is that they perform very well outside these specific innovation markets and therefore remain R&D inactive.

The findings on R&D entry and exit suggest that the share of competitive and "deep" innovators within the R&D markets has not diminished significantly and maybe has even increased. This raises the question of what this structural change means for productivity growth and the allocation of productivity gains.

To answer this question, we investigate - in a first step - if theoretically important factors (in-house R&D success probability, imitation success probability, R&D costs) are related to the decision of a company to conduct R&D. These factors indeed all have a significant effect, but with a different magnitude. The estimation results also show that the fraction of higher-educated employees, the access to international markets (export activities), the technological potential, and access to university knowledge are significantly and positively related to the R&D success probability and consequently to the R&D decision of a company, whereas the number of principal competitors worldwide shows a negative correlation. The latter indicates that intense competition makes it difficult for companies to conduct R&D. This confirms the findings in the relevant literature. In the course of the time, however, we see a decline in the R&D success probability which indicates that it has become more difficult to innovate. By contrast, the importance of innovation costs and the imitation success probability hardly changes across time.

In a second step, we simulate how the productivity growth rate and dispersion depend upon the in-house R&D success probability and the imitation success probability. The results show that if we increase the in-house R&D success probability from 0.8 to 1 (full success), productivity would grow by 14%, while setting the imitation success probability from 0.85 to 1 would increase productivity by 6%. Hence, policies targeting to increase the in-house R&D success probability are more effective than those targeting the diffusion/adoption of technologies (imitation success probability). To go for the more effective policy option, however, has some side effects. An increase in in-house R&D success probability also increases the dispersion of productivity growth leading to greater inequality in the Swiss business sectors. By contrast, an increase in the imitation success rate, decreases inequality.

In a third step, we simulate if public support to reduce the private R&D costs - for instance through R&D tax credits or subsidies - significantly increases the productivity growth rate (due to an increase in the in-house R&D success probability). Compared to a situation without such a policy, the productivity growth rate would increase only marginally by 0.05%. Since this growth effect results from companies entering the R&D markets, the number of R&D active companies would increase, which leads to a reduction in the productivity growth dispersion in the economy. The effect from such a policy measure is thus modest. Policies aimed at facilitating access to international markets, increasing the availability of highly qualified workers, or promoting cooperation between universities and industry to improve the in-house R&D success probability might be more effective for the time being.

Comparison with the Netherlands: Since we see country-specific developments in the share of R&D-active companies, the question naturally arises which country-specific factors could be driving such differences. In cooperation with the VU Amsterdam, we have analyzed the development in the Netherlands and compared it with Switzerland. The Netherlands is an interesting country for comparison, as it not only shows a completely different development in the share of R&D-active companies, but also pursues different innovation funding strategies than Switzerland. For instance, it introduced the patent/innovation box as early as 2007, while a patent box has been introduced in Switzerland only in 2020. The comparative analysis shows that although we observe a similar decline in the survivor function for both countries over the 20 years under investigation, the share of companies exiting R&D has declined significantly less in the Netherlands than in Switzerland over the past 10 years. Thus, the increase in R&D propensity in the Netherlands is partly based on a less pronounced exit rate. Moreover, we see that competitive and deep innovators in Switzerland are more resistant to exit R&D than in the Netherlands, while less competitive and less deep innovators in Switzerland are more inclined to exit R&D than in the Netherlands. One possible reason for this difference could be the more pronounced innovation support, for example in the context of the patent/innovation box, in the Netherlands. And indeed, based on a linear probability model, we see that the patent/innovation box has a positive and significant impact on the firm's R&D decision. This shows that R&D policy incentives can influence firms' decision to conduct R&D. However, it remains to be investigated how such policy instrument affects firm-level productivity.

2 Introduction and Research Question

According to Global Innovation Index or the European Innovation Scoreboard, Switzerland is on of the most innovative countries in the world. This top position is primarily due to various innovation-relevant infrastructure indicators: for example, general human capital, which can be measured by the density of doctoral degrees or as "life-long learning," as well as various patent indicators and the attractiveness of the higher education system. If we look at direct innovation indicators at the company level, such as the proportion of R&D active companies, we see a different development. The proportion of companies with R&D activities is falling significantly in Switzerland and also in Germany, while this indicator has developed very positively in the Netherlands, for example. Accordingly, a very well developed, innovation-relevant infrastructure is a necessary, but by no means a sufficient condition for sustainable innovation efforts by companies.

If companies decide to exit R&D, this is often a rational decision, probably made on the basis of cost and risk considerations. Such decisions are only noteworthy from an economic perspective if they have a negative impact on the country's competitiveness and productivity growth. In this case, economic policy measures are appropriate. Accordingly, the main objective of this study is, first, to identify which type of firms continue or stop performing R&D and, second, whether a declining R&D rate has consequences for productivity growth and thus a country's competitiveness. In particular, we aim to address the following research questions: How did innovation activities in Switzerland compare to selected countries such as the Netherlands and Germany? What are the main factors associated with the decline in R&D active firms in Switzerland? Are there significant differences between the Netherlands and Switzerland in these factors? Are there asymmetric effects? That is, do the factors responsible for the withdrawal from R&D activities differ from the factors motivating firms to start R&D activities? Do competitive pressures and cyclical fluctuations influence the (termination of) R&D activities? Does past innovation success increase the likelihood of continuing R&D activities?

In order to respond to these questions we need to apply a variety of methods. First, we present descriptive statistics based on firm-level information from the Community Innovation Survey (CIS) and the Swiss innovation survey (SIS). Second, we use "survival analyses" to identify the factors for R&D entry/exit. Third, we perform linear probability estimation to identify the main factors for the R&D decision of a company. Fourth, we apply an endogenous growth model (simulated methods of moments procedure) to estimate the significance of R&D activities for productivity growth and dispersion and simulate the effectiveness of policy designs to reduce the R&D cost (e.g. R&D tax credits or subsidies).

The analyses show that there is a great heterogeneity in the development of the R&D rate in European countries. While Switzerland and Germany show a decreasing trend, the Netherlands, Finland, and Austria show increasing R&D rates over time. A large number of employees (firm size), productivity, human capital, exports, and technological potential are the most important individual factors that reduce the probability of exiting R&D. In addition, we see that a strong focus on innovation activities and past economic success (competitiveness) increases a company's probability of remaining active in R&D over a long period of time. Lack of equity and the ease of copying innovations are important characteristics that increase the probability of exit. In general, we observe a high symmetry of the factors determining entry into or exit from R&D markets, with one exception: productivity reduces the probability of exit but does not increase the probability of entry. The main reason for this is that productive, non-R&D firms may see no reason to enter risky and costly R&D markets. Our results also show that the proportion of employees with higher education, access to international markets, and university-industry partnerships increase the likelihood of in-house R&D success and thus the likelihood of performing R&D.

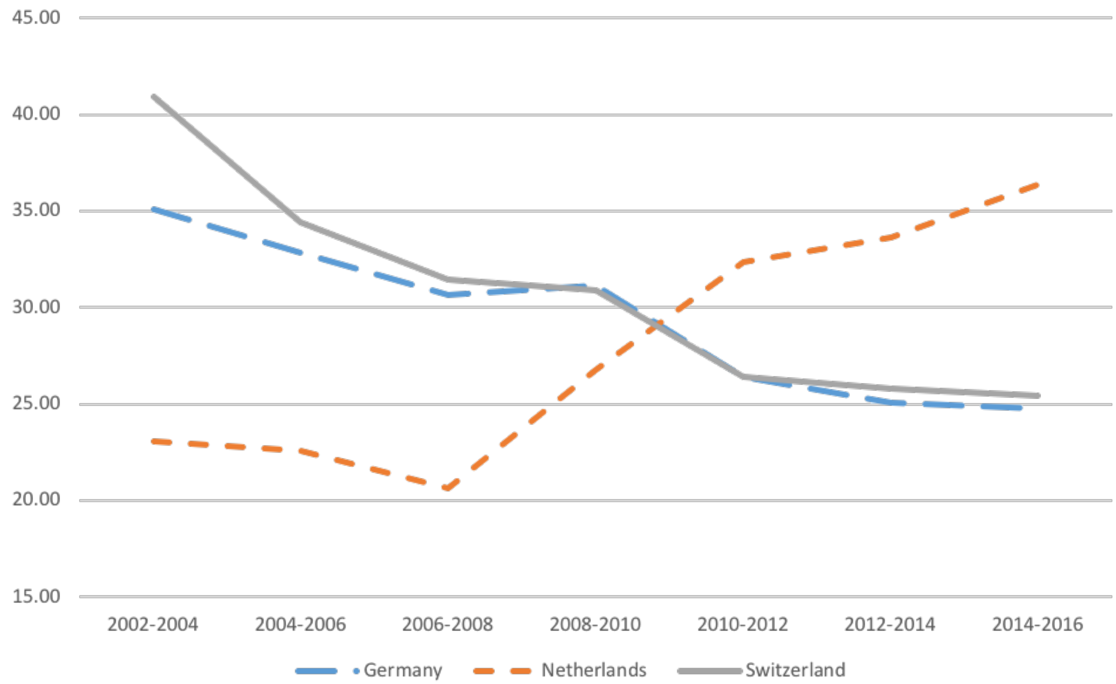
It is worrying that the internal R&D success probability is decreasing over time. This suggests that it has become more difficult to innovate. Furthermore, this development has a negative impact on productivity growth and its dispersion. Therefore, policies should be designed that increase the internal R&D success probability or the imitation probability. What kind of innovation policy could be effective? Simulations of the productivity effects of policy designs, such as subsidies or R&D tax credits, that reduce the private costs of R&D show very small positive effects. Therefore, policies that improve access to international markets, policies that reduce competitive pressures for a limited period of time, or policy designs to further improve university-industry partnerships or the availability of highly skilled workers might be more effective.

In Section 3 we provide a detailed discussion of the recent literature on R&D and productivity. Section 4 gives an overview of descriptive statistics R&D and productivity in Switzerland and abroad. In Section 5 we discuss the results of a survival analysis of why firms start or stop doing R&D. In Section 6 we analyze the determinants of firms' decision to conduct R&D. Finally, Section 7 analyzes how in-house R&D, technology diffusion, and R&D subsidies affect productivity growth and inequality. Section 8 concludes. Additional relevant material and proofs can be found in the Appendix.

3 Literature Review and Conceptual Notions

The concentration of R&D activities among ever fewer firms we observe in Switzerland is not an isolated economic phenomena. We observe similar tendencies of increasing concentration across different economic activities in most developed countries. For instance, Autor et al. (2020) find an increasing concentration in market shares, labor shares, and markups over the last decades not only

Figure 1: The fraction of R&D active companies - an international comparison. Source: EUROSTAT 2019.



in the US but internationally.

3.1 The Rise of the Superstar Firms

Dorn et al. (2017) and Autor et al. (2020) put forward the rise of superstar firms like Google, Amazon, or Facebook as an explanation for this increasing concentration. Industries are characterized more and more by a "winners-take-it-all" characteristic, where one or only a few firms gain very large market shares. Superstar firms capture higher market shares because they produce more efficiently; they are more productive and thus become larger than their competitors (Dorn et al., 2017). Autor et al. (2020) argue that globalization and technological change have pushed market shares toward these most productive firms in all industry. More specifically, factors like new technologies with positive network effects, diffusion of competitive platforms, or information intensive goods with high fixed costs (first-copy costs) have all played a key role for increasing the concentration of market shares (Dorn et al., 2017). Indeed, Autor et al. (2020) find that dynamic industries exhibiting the fastest technological progress are also those industries that experienced the most rapid concentration of market shares. Thus, one central driver behind the concentration of economic activity is the concentration in innovative activity. The increasing concentration of R&D activities that we are seeing in Switzerland can therefore have an impact on the entire economy, as superstar companies such as Novartis, UBS or Nestle gain increasing economic importance at the national level over time.

However, even though highly productive firms managed to gain ever larger market shares, aggregate productivity in the economy did not rise by more than in earlier decades. In fact, productivity growth has deteriorated over the last decade.[1] To explain this discrepancy, Autor et al. (2020) argue that knowledge diffusion between leaders and laggards in each industry may have slowed down and that therefore the productivity gap between leaders and laggards has accentuated. In contrast, an alternative explanation would be that innovations simply have become harder to develop than in earlier decades (Bloom et al., 2020). Today firms need much more resources to achieve technological progress, and only the superstar firms are able to do so. This latter mechanism would cause a similar widening in the productivity distribution between firms as a slowdown in knowledge diffusion would. In the following sections, we will discuss these two potential culprits for the increasing concentration in economic activities in general and innovation activities in particular: i) a decrease in knowledge spillovers and ii) an increase innovation costs. They are both candidates to explain the increasing (productivity) gaps between leaders and laggards, culminating in a concentration of economic activity.

3.2 Knowledge Spillovers

Akicigit and Ates (2019) go beyond just the concentration of economic activity and review the literature on the broader category of business dynamism. They highlight ten stylized facts describing the declining business dynamism in the United States. Many of these stylized facts are also observable in other economies. Important for our context are the following stylized facts: i) market concentration has risen, ii) average profits have increased, iii) the labor productivity gap between frontier and laggard firms has widened, iv) firm entry rate has declined, v) the dispersion of firm growth has decreased. Hence, Akicigit and Ates (2019) observe a concentration of economic activity that is even broader than the one discussed by Dorn et al. (2017) and Autor et al. (2020). They propose a unified theoretical framework to explain this declining business dynamism. Like Autor et al. (2020), they find that a decline in the intensity of knowledge diffusion between firms must be a significant driver of the stylized facts observed in the literature. However, the exact nature of knowledge diffusion remains open. Akicigit and Ates (2019) propose: a) the data-dependent nature of production, b) regulations favoring established firms, c) off-shoring of production abroad, and d) anti-competitive use of intellectual property as potential mechanisms that could have lowered the diffusion of knowledge. However, the authors do not present empirical support for their claim that knowledge spillovers have declined.

Before we proceed to empirical evidence regarding the strength of knowledge spillovers, we have to shortly discuss their nature. Knowledge spillovers have two central properties that affect how they spread from leaders to laggards (Eeckhout and Jovanovic, 2002). First, while it is commonly accepted that discoveries at the technological frontier provide an advantage for the inventors, firms that are further away from the technological frontier can profit from such discoveries as well. Second, the advantages of discoveries for the leaders are often hard to protect from the copying efforts of the laggards. This is the positive side of knowledge spillovers; they guarantee that discoveries flow from leaders to laggards. Knowledge spillovers thus contribute to a convergence of firms. Copying allows followers to access knowledge from the technological frontier without having to bear the full cost of development. Knowledge spillovers thus level the playing field between leaders and followers. However, without an adequate protection of discoveries, the incentives for leaders to push

[1] Measurement of productivity is an important indicator for the performance of an economy (Bartelsman and Doms, 2000), and the productivity across firms is often highly unequal (Bartelsman and Wolf, 2018).

the technological frontier are diminished. Firms cease to bring forward inventions when they expect to be copied immediately. The free-riding behavior of followers can therefore reduce investments in innovation below the optimal level (Eeckhout and Jovanovic, 2002). To quantify the extent of knowledge spillovers, we need empirical evidence. This gap is filled by the empirical studies of Andrews et al. (2015), Andrews et al. (2016), and Lucking et al. (2019).

The diffusion of technology from frontier to laggard firms is the central topic in Andrews et al. (2015). The authors find that, despite the slowdown in aggregate productivity, productivity at the global frontier remained robust. In contrast, a rising productivity gap between firms at this global frontier and laggard firms has emerged, opening up the question why seemingly non-rival technologies have not diffused to this latter group of firms. Andrews et al. (2015) document a highly uneven process of technology diffusion, where global frontier technologies only reach laggard firms through the respective national frontier firms. Technologies from the global frontier spread at an increasing speed across countries but at a decreasing speed within countries. Andrews et al. (2015) argue that this slowdown in technology diffusion from the global technology frontier to laggards in each country can explain the overall slowdown in aggregate productivity growth in OECD countries over the past decade.

In sharp contrast, Lucking et al. (2019) show that the diffusion of technological knowledge has been more or less constant over time. They differentiate between positive and negative knowledge spillovers. Positive knowledge spillovers can offset a decrease in R&D productivity. They would cause a more stable R&D productivity among firms. Negative knowledge spillovers in the form of product market rivalry (i.e., business stealing through increasing market shares), on the other hand, diminish R&D productivity. Lucking et al. (2019) investigate how positive and negative spillovers have developed from 1980 to 2015. They find a stable ratio of positive to negative knowledge spillovers over time.

3.3 Lack of Complementary Investments

However, while Lucking et al. (2019) find stable levels of both positive and negative knowledge spillovers for R&D over time, there remains the increasing productivity gap between global frontier and laggard firms identified by Andrews et al. (2015) and Andrews et al. (2016). If it is not the knowledge spillovers in the R&D process which have diminished, knowledge spillovers must have diminished on other dimensions. Andrews et al. (2016) argue that global frontier firms are not just able to introduce technologically innovative goods and services but also have the capacity to optimally combine intangibles in their production process, such as technological, organizational, and human capital. Laggard firms, in contrast, lack these important capacities. The sophistication required for combining complementary investments has strongly increased, as technologies have become increasingly complex, demanding a high degree of tacit knowledge. The adoption of ICT technologies requires high learning and high adjustments costs that need to go hand in hand with knowledge based capital. Andrews et al. (2016) argue that we are still in the transition phase toward the digital economy and firms are still learning. Consequently, the slower adoption of technologies by laggard firms is caused by increasing costs of the transition from an economy based on production to one based on ideas (Andrews et al., 2016).

3.4 Higher Innovation Costs

However, Andrews et al. (2016) find that, although more robust, even the global frontier firms have shown declining productivity growth over the last years. To explain the concentration of economic activity toward fewer firms, we must thus go beyond the gap between leaders and laggards and search for potential problems in the innovation process itself, hindering the development of innovations.

Cavenaile et al. (2019) examine how rising costs of innovation have been responsible for the macroeconomic trends we have discussed at the outset, namely the rise of the superstar firms, with industries becoming increasingly dominated by a small number of large firms. They develop a theoretical framework that can explain rising market concentration, markups, profits, and, importantly, rising R&D spending, as well as the decrease in firm entries, productivity growth, and labor shares. Moreover, their model is also consistent with the inverted U-shape relationship between innovation and competition described by for instance Aghion et al. (2005). Cavenaile et al. (2019) find that, while the increase in markups through superstar firms caused significant static welfare loss, this loss was outweighed by dynamic welfare gains. Due to the expected higher profits, firms have faced increased incentives to innovative. This suggests that increasing market concentration is not only bad for innovation and economic growth, and superstar firms are not necessarily a problem as such. However, the findings of Cavenaile et al. (2019) leave open the question why we then observe a decline in productivity growth. They argue that the culprit here is the increasing costs of innovation, that is, "ideas are getting harder to find" (Bloom et al., 2020). Their model predicts that if the costs of innovation for firms were set to the levels of some decades ago, the decline in productivity growth would have been more than offset. Thus, in our discussion of a concentration of R&D activities, we need to focus on the costs of innovation. The study of Bloom et al. (2020) offers empirical data in this respect.

Bloom et al. (2020) investigate whether, over time, the same amount of R&D input has led to the same amount of output. In other words, they ask whether the R&D productivity of firms has decreased or not. Their paper is thus also entitled "Are ideas getting harder to find?". Bloom et al. (2020) apply simple growth accounting to the production function for new ideas. Economic growth emerges from the product of i) the effective number of researchers and ii) their research productivity. Bloom et al. (2020) present a wide range of empirical evidence showing that research effort is rising, while research productivity is falling. The economic growth we observe is still positive because of the increase in research effort, which has so far always offset the decrease in research productivity. For instance, for the US economy as a whole, Bloom et al. (2020) find that research productivity has declined by a factor of 41 since 1930. In other words, it requires 41 times more researchers today than in 1930 to achieve the same level of economic growth. Thus, consistent economic growth require the employment of ever more (effective) researcher. This is clear evidence of a "draining out" of ideas. The low hanging fruits have been picked, and we now need more and more effort to also get at the high hanging fruits. The concentration of innovation activities among ever fewer firms thus makes sense economically. Only the superstar firms are able to still profitably unearth further innovations.

Jones (2009) describes the increasing difficulty to innovate as a "burden of knowledge". If knowledge is of a cumulative nature, later innovators will have it more difficult, as they need to learn more until they can reach the knowledge frontier. This implies that, today, innovators require much more and longer education than previous generations of innovators. The educational burden has increased. However, next to further education, innovators have two more ways to cope with the burden of knowledge: specialization and teamwork. They both take away pressure from individual innovators, as they distribute the efforts among additional minds. Importantly, both specialization and teamwork have profound implications for the organization of the work place. Using a rich patent data set, Jones (2009) finds empirical support for a burden of knowledge. Age at first invention, specialization, and teamwork all have strongly increased over time. Thus, Jones (2009) provides a potential micro founded explanation why productivity growth did not accelerate in the last decades despite strongly increasing research efforts. Innovations have simply become too costly for smaller firms to pursue.

3.5 High Fixed Costs

Notably, the two opposing trends of a falling share of innovative firms and of rising innovation expenditures we observe in Switzerland show up strongly also in Germany. Rammer and Schubert (2018) investigate the mechanisms behind this concentration on the few. They argue that small firms are disadvantaged in the innovation process because they have fewer resources available to cover the fixed costs of their innovation projects. According to Cohen and Klepper (1996), large firms have a larger product base upon which they can distribute their fixed costs. Consequently, if fix costs are rising, due to for example digital technologies or diminishing technological opportunities, small firms will be more likely to abandon their innovation activities, as they cannot afford the accruing fixed costs anymore. Only large firms will be able to pay the high fixed costs to still profitably exploit the ideas that are getting harder to find.

Rammer and Schubert (2018) argue that the observation that small firms are disproportionally stopping their innovation activities is problematic on two grounds. First, once firms have given up their innovation activities, it will be harder for them to reenter again. Those firms that discontinue innovation might lose important capabilities. While such decisions may save costs in the short run, they may erode the long-term competitive position of the respective firms. Second, the increasing concentration of innovation activities on large firms can lead to a concentration of risk. Dependence on a small number of very successful firms makes the entire economy more vulnerable to external shocks. If one or a few very large firms suffer adverse shocks, the economy will be hit harder than when the shock affects some smaller firms operating in different industries.

Aghion et al. (2019) provide yet another take to explain the declining business dynamism and rising concentration in developed economies. They formalize how superior process efficiency in firms leads to falling economic growth, increasing market concentration, and lower labor shares. Aghion et al. (2019) argue that firm such as Walmart, Microsoft, or Amazon can profit from organizational capacities that have evolved over time and are very hard to reverse engineer for competitors, such as their highly sophisticated logistics. Such highly process efficient firms have entered many new geographical and product markets over the last two decades. Aghion et al. (2019) argue that this was due to the IT wave, which allowed firms to extend to a wider array of product lines, as it increased the optimal boundary of firms. The costs of spanning multiple markets have fallen. Since firms with superior process efficiency have higher markups and lower labor shares, their expansion into new geographical and product markets will increase aggregate markups and decrease aggregate labor shares. Moreover, it deters innovation because of the increased levels of competition between firms. If competition becomes too strong, the existing firms tend to shy away from innovation, as they cannot appropriate rents anymore, especially because they do not have the means to become as process efficient. The vast majority of firms can never afford the high fixed costs that setting up the logistics of Amazon requires. Aghion et al. (2019) thus conclude that, counterintuitively, it may be welfare enhancing to constrain the expansion of the most efficient firms, as they restrict innovation and thus long-run growth. With their dominance, highly process efficient firms create an environment that inhibits steady levels of innovation.

3.6 Knowledge Spillovers vs. Innovation Costs

König et al. (2016) investigate a theoretical model where firms can improve productivity either by pursing in-house R&D or by imitating other firms' technologies. The choice of these two strategies is thereby based on profit maximization. Firms choose the strategy that offers higher profits. Whether in-house R&D or imitation provides higher profits depends on the distance of the firms to the technological frontier. Firms close to the technological frontier have fewer opportunities for imitation and thus choose in-house R&D simply because there are fewer better firms they can learn from. In contrast, firms further from the technological frontier have more opportunities to imitate and consequently also choose imitation; there are more firms they can learn from. In the model of König et al. (2016), the technology spillovers from the frontier to the laggard firms assure that the overall productivity distribution does not diverge. Repeatedly unsuccessful firms can switch to imitation to avoid falling behind too much, closing the ranks to the technological frontier. König et al. (2016) test their model on TFP distributions of French firms over the years 1995-2003, which together show the pattern of a traveling wave, with increasing mean and variance. The model is indeed able to successfully reproduce the development of the productivity distributions observed in the data.

König et al. (2020) build on the model of König et al. (2016). They similarly focus on the trade-off between innovation and imitation. Profit-maximizing firms seek to upgrade their technologies. They can either try to learn from other firms or discover new technologies. Each firm's productivity determines which of the two strategies is more profitable. Firms further behind from the technological frontier are more likely to gain from imitation, whereas firms closer to the technological frontier have limited opportunities to learn from others and thus need to break new grounds through innovation. In contrast to König et al. (2016), König et al. (2020) also incorporate R&D subsidies and output taxes in their model, which allows identifying the impact of misallocations on the evolution of productivity growth. König et al. (2020) apply their model to data from Chinese manufacturing firms. Over the last decade, the Chinese government has increasingly focused on supporting the innovation activities of domestic firms, which has led to boom in R&D expenditures in China. However, König et al. (2020) find that a substantial portion of these R&D expenditures corresponds to fake R&D. Chinese firms respond to fiscal incentives by the government through fudging R&D, that is, they relabel part of their operational expenditures as R&D in order to obtain public subsidies. In contrast, König et al. (2020) show that when the same model is applied to data from Taiwan, there is no sign of "fake" R&D. This is important for our report, as, in contrast to Switzerland, the Netherlands has strongly supported R&D activities as well. We will investigate whether this R&D in the Netherlands is fully effective or not.

3.7 R&D Subsidies

Arqué-Castells and Mohnen (2015) investigate whether there is room for R&D subsidies to influence the continuation of or entry in R&D. They estimate a model of the firms' optimal R&D decision, where firms decide not only about whether they pursue R&D but also how much they invest in R&D. The impact the R&D subsidies can have depends on the relation between both fixed costs and sunk costs of R&D. If the fixed costs are positive but the sunk costs are negative (i.e., there is a sunk profit), R&D subsidies can play the role of a one-shot trigger and induce permanent R&D activity. Arqué-Castells and Mohnen (2015) find that R&D subsidies do affect both the decision for R&D and the level of R&D expenditures. Moreover, there is a state dependence: firms with R&D are much more likely to be R&D active in the next period than firms without R&D. Hence, because the negative sunk costs, which compensate some of the positive fixed costs, there is the possibility of a permanent inducement effect. Firms that entered R&D because of the subsidies may remain R&D active afterwards, even if the subsidies are withdrawn. However, the effect of such a trigger policy fades away after seven years. In addition, if the fixed costs are high in relation to the sunk costs, more structural policies aimed at reducing fixed costs will be necessary.

Using a sample of German manufacturing firms, Peters et al. (2017) estimate a dynamic structural model of R&D investments to quantify their long-run benefits. They use the structural parameters of their model to simulate how a change in innovation costs, such as a tax break or an R&D subsidy, affects the firms R&D choice and its future productivity. They find that among R&D active high-tech firms 20% lower maintenance costs of R&D lead after ten years to a 9%-points higher probability for firms to pursue R&D and a 1.4% higher productivity. The respective numbers for the low-tech sector are smaller: companies in this sector are only 7% points more likely to continue R&D, while the productivity effects are very small. In contrast, for high-tech firms just beginning with R&D, 20% lower maintenance costs has very little impact on the probability for the firm to pursue R&D and thus also on its productivity. For low-tech firms these numbers are again different. Low-tech firms just beginning with R&D have a 7%-points higher probability to pursue R&D and a 2.1% higher productivity. Thus, reducing innovation costs has an effect on both continuation of and entry into R&D.

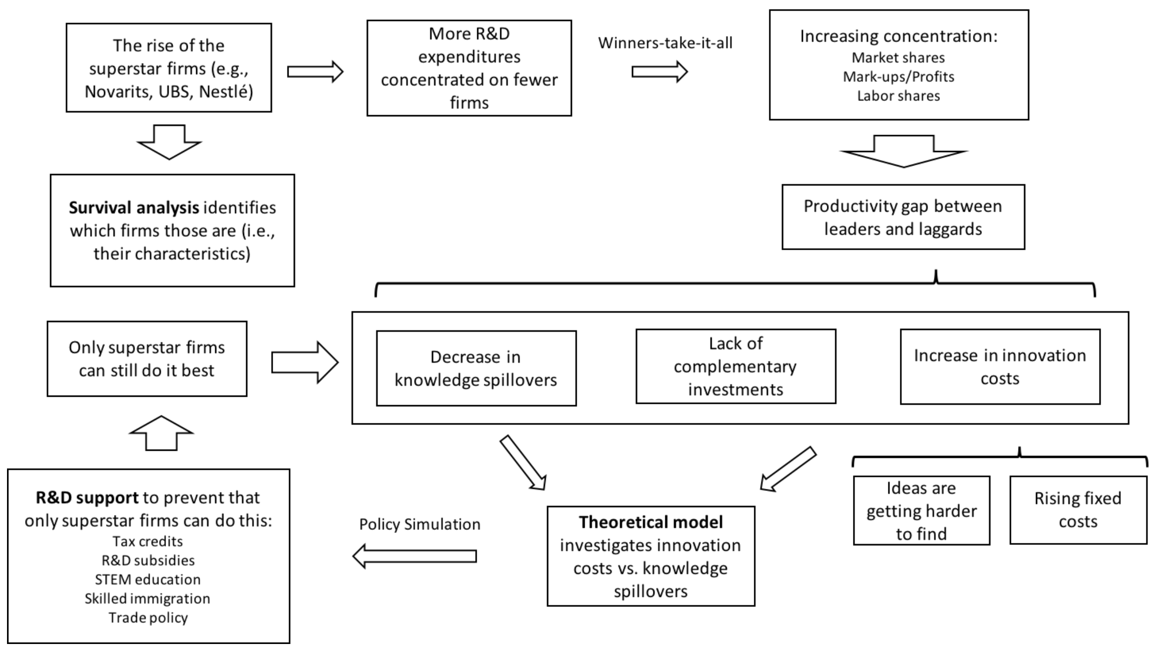
3.8 Comparison of the effectiveness of policy measures

There is a broad range of possible policy instruments to foster both entry into R&D and R&D expenditures. Bloom et al. (2019) summarize these different instruments and provide a toolkit for policymakers. They rank tax credits for R&D as the most effective instrument in the short run. They quality and conclusiveness of the evidence that they provide a net benefit for R&D activities are high. Despite a danger of relocation of R&D activity, the effect of tax credits on R&D and productivity is substantial. Also positive already in the short-run are R&D subsidies. They profit from direct targeting to promising industries and can also directly fund basic research at for instance universities. However, the quality and conclusiveness of the literature is somewhat less clear than for tax credits. The problem is whether this targeting indeed proves effective. The wrong projects may be targeted and there may be a crowing out of industry R&D. Bloom et al. (2019) argue that increasing human capital through for instance higher university enrolment in STEM fields is most effective in the long-run. More R&D workers increase the volume and lower the price of R&D. The literature provides medium quality and conclusiveness in the respect. Already yielding a positive net benefit in the short run is immigration of skilled R&D workers. The literature is very unambiguous in that such an increase in human capital is very effective. Importantly, as the only available policy instruments, an increase in human capital will also lower inequality in the economy, as the competition for this scare resource decreases. Another possible policy lever is product market competition and international trade. Bloom et al. (2019) argue that the empirical literature shows that greater competition and trade openness typically increase innovation. The quality and conclusiveness of the literature is high with regard to such policies. Moreover, the financial costs are low, but these policies will clearly further increase inequality in the economy. Bloom et al. (2019) rank patent boxes as least effective instrument and even ascribe it a negative net benefit. They only encourage firms to relocate their intellectual property royalties into different tax jurisdictions. Patent boxes are instruments that allow multinational firms to minimize their tax burden. Bloom et al. (2019) argue that they are an example of a harmful form of tax competition that disguises as pro-innovation policy. Finally, there is evidence that co-locating many small high-tech firms together proves valuable to produce agglomeration effects. However, the literature is not as unambiguous as local policy makers might think. Bloom et al. (2019) argue that what is important is fostering the R&D activities of young firms, such as angel finance or venture capital initiatives. In our report, we will focus on instruments such as tax credits and R&D subsidies. They are broad instruments with the capacity to stabilize the share of R&D active firms already in the short run.

Akcigit and Stantcheva (2020) focus more narrowly on tax policy, which, if applied appropriately, can provide incentives for innovation. If applied inappropriately, it can create disincentives and inhibit innovation. It is thus important to provide the correct tax policy. Akcigit and Stantcheva (2020) divide tax policies into general tax policies, such as personal or corporate income tax, and targeted tax policies, such as R&D tax credits and subsidies for specific types of research. Innovation reacts to tax policy because it is an economic activity as well, which depends on intentional efforts and forward-looking investments; firms form expectation about the net present values of the returns to R&D. Akcigit et al. (2018) show that in the US both corporate and personal income taxes negatively affected the quantity and quality of innovation throughout the 20th century, with similar responses as for instance labor supply. Akcigit and Stantcheva (2020) argue that the response of firms to tax policy depends on where they are in their life cycle, that is, where they are located in their development from new entrants to established incumbents. R&D tax credits support big and profitable firms most. However, it is the small firms that contribute the most radical innovations. Akcigit and Stantcheva (2020) thus argue that R&D tax breaks for small firms would foster breakthrough innovations most. Moreover, the tax regime should also preferentially treat venture capital firms supporting starts up.

Akcigit et al. (2019) address the fact that not all firms are equally productive in their R&D process. R&D subsidies can thus have differential effects on firms. However, the productivity of the R&D process is seldom directly observable. Akcigit et al. (2019) thus recommend a policy mix of taxing incumbents and subsidizing R&D. The subsidies in R&D solve the problem that knowledge spillovers are not internalized, whereas the taxes on incumbents force less productive firms to exit and free up resources. Akcigit et al. (2019) describe the optimal tax as a lower marginal tax for more profitable firms and lower marginal subsidies for firms with high R&D investment levels. The latter is beneficial because it incentivizes firms that are not as productive in their R&D process to imitate those firms that are more productive in their R&D process. The former is beneficial because firms that are productive in their R&D process can produce more profits with the same R&D investments. Taxing profits to a lesser extent is therefore attractive for firms that are more productive in R&D. However, Akcigit et al. (2019) argue that a linear corporate profit tax would also be feasible, as the benefits for less profitable firms are not sizeable; these firms anyway pay relatively low taxes. In contrast, a non-linear tax policy for R&D subsidies is very crucial, as it would encourage a more efficient R&D process among firms.

Figure 2: Literature overview.



3.9 Contribution to the Literature

Overall, the above literature discussion shows that those firms that are most productive also remain innovative. These superstar firms continue to push the technological frontier in a robust way, even though they also have to fight the increasing costs of innovation. Superstar firms cause through their expansion a concentration not only in market shares but on many other dimensions as well. Most important, the rise of superstar firms goes hand in hand with an increasing concentration of R&D activities. They are the group of firms who can still best afford the ever increasing costs of innovation. In the report, we will analyze whether the developments we observe for the US and other developed countries also hold for Switzerland. We pursue a survival analysis, where we analyze the decisions to exit from and enter into R&D. The target of this survival analysis is to identify which group of firms still has the highest likelihood to remain R&D active and which group is most likely to exit from R&D. In case we should observe that those firms who are still R&D active are primarily the type of superstar firms, we can except to observe a similar concentration of economic activity in the hands of fewer firms in the future.

To identify the causes for the increasing market concentration, we rely on the theoretical model of König et al. (2016) and its extension in König et al. (2020). It allows to investigate to which extent rising innovation costs or decreasing knowledge spillovers are responsible for the decrease in the share of R&D active firms in Switzerland. Moreover, it makes it possible to investigate the impact of potential R&D subsidies on these developments.

4 Descriptive Information and International Comparison

In Section 4.1 we provide an international comparison, while Section 4.2 provides descriptive statistics for the Swiss economy.

4.1 International Comparison

The results for an international comparison of R&D activities have already been published elsewhere (Spescha and Woerter, 2020). Here we will only reconsider the developments most relevant to this report. Figure 3 shows the share of R&D active firms for eight selected European countries. Switzerland exhibits a very similar downward trend as Germany; both have experienced a sharp decrease in the share of R&D active firms. All other countries show a much more stable development over time. France, Finland, and especially the Netherlands show an increasing share of R&D active firms over the entire time span.

Figure 4 shows how the share of R&D expenditures in sales has developed for these same eight countries. Interestingly, here we see an increase for in this indicator for both Switzerland and Germany. Hence, in these two countries we observe an increasing concentration of more R&D expenditures on an ever smaller number of R&D active firms. Austria and Sweden show an increase in the share of R&D expenditures in sales, too. Yet, because they have not experienced a similar decrease in the share of R&D active firms, their R&D activities do not exhibit increasing concentration. The only countries that show a decrease in the share of R&D expenditures in sales are France and, in the last period, Finland. Interestingly, the Netherlands also show an increase in the sales share of R&D expenditures. However, we do not observe the Netherlands after 2012 anymore, so we cannot make as strong a statement. Nonetheless, the data suggest that firms in the Netherlands are not only more frequently pursuing R&D but also increasing their respective R&D expenditures.

This report does not only tackle why we observe a decreasing share of R&D active firms in Switzerland, but also why we observe an increasing share of R&D active firms in the Netherlands. These two contrasting developments might have a lasting effect on the respective competitiveness of the two countries.

Figure 3: The fraction of R&D active companies - an international comparison. Source: EUROSTAT 2019.

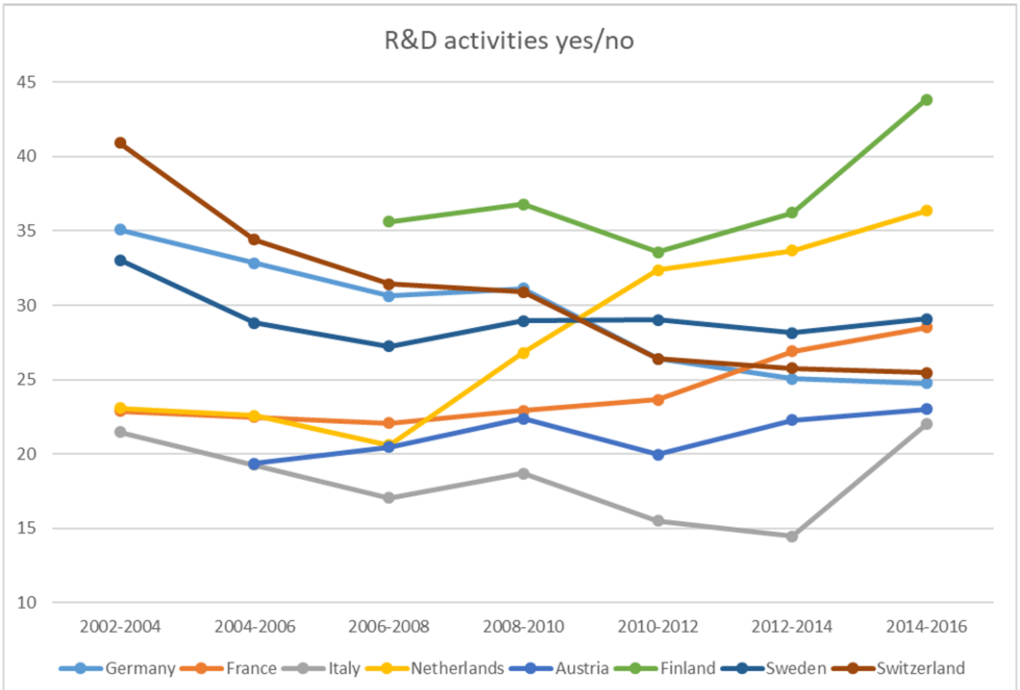
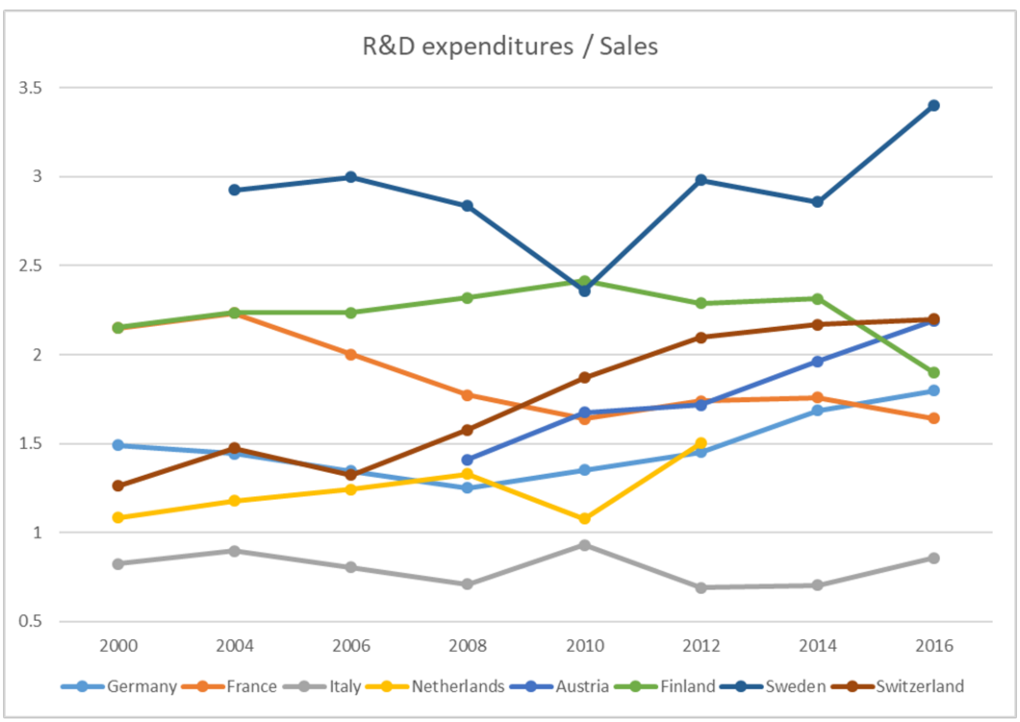


Figure 4: The fraction of R&D expenditures over sales - an international comparison. Source: EUROSTAT 2019.



4.2 Descriptive Information for Switzerland

In the following we present descriptive evidence about the distribution of productivity and the innovation (R&D) decisions of firms in Switzerland using the Swiss innovation survey data over the years 1999 to 2017.

Figure 5 shows the log-productivity distribution in Switzerland over the years 1999 to 2017. Productivity is measured as value added per employee. The figure illustrates a decline in the productivity growth and an increasing dispersion in the productivity distribution. Next, Figure 6 shows the average productivity over the years ranging from 1996 to 2017, as well as the average productivity for different productivity percentiles (that is, the average is computed over all firms with productivity below the 10th, 20th, ..., 100th percentile). While the average is increasing in the highest percentiles, it is declining in the lowest percentiles. This suggests that the trend in the lowest percentile plays an important role in the overall increase in productivity disparity. Figure 7 shows the productivity growth rate and the standard deviation over the years from 1996 to 2017. The growth rate shows a declining trend while the standard deviation shows an increasing trend over time. The later indicates an increasing dispersion in the productivity distribution. Moreover, Figure 8 shows the innovation decision (whether firms conduct R&D or not) over their productivity levels. We see that firms with higher productivity tend to innovate more. Further, Figure 9 shows the correlation of the innovation decision with productivity. Except for the year 2005, the correlation shows a declining trend. We also see that the innovation threshold (i.e. the lowest productivity level at which the innovation decision is more likely than 50%) is increasing over the years of observation.

To summarize, a descriptive statistical analysis using survey data for Switzerland allows us to make the following observations:

(1) We observe a declining productivity growth rate (see Figures 5, 6 and the left panel in Figure 7).

(2) We find an increasing dispersion in the productivity distribution (see Figure 5 and the right panel in Figure 7).

(3) We find a decline in the fraction of R&D active firms (concentration of R&D).[2] This comes along with an increasing productivity threshold in the decision of firms to innovate (see Figures 8 and 9).

In the following sections we will analyze the potential causes for the above listed stylized facts.

[2]See Rammer and Schubert (2016) for a related study of German firms showing a similar declining trend.

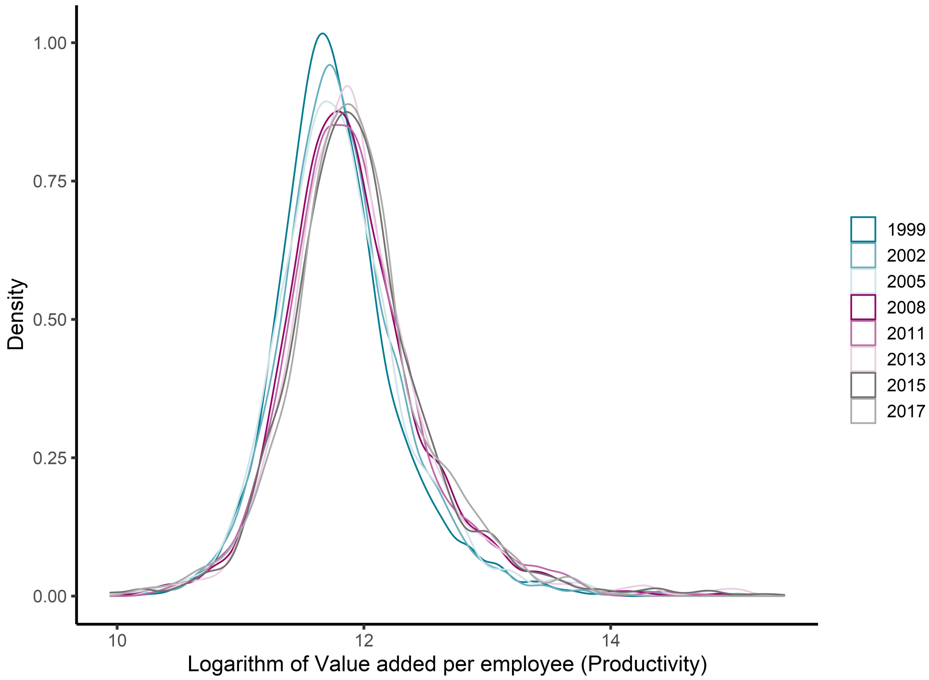


Figure 5: The log-productivity distribution in Switzerland over the years 1999 to 2017.

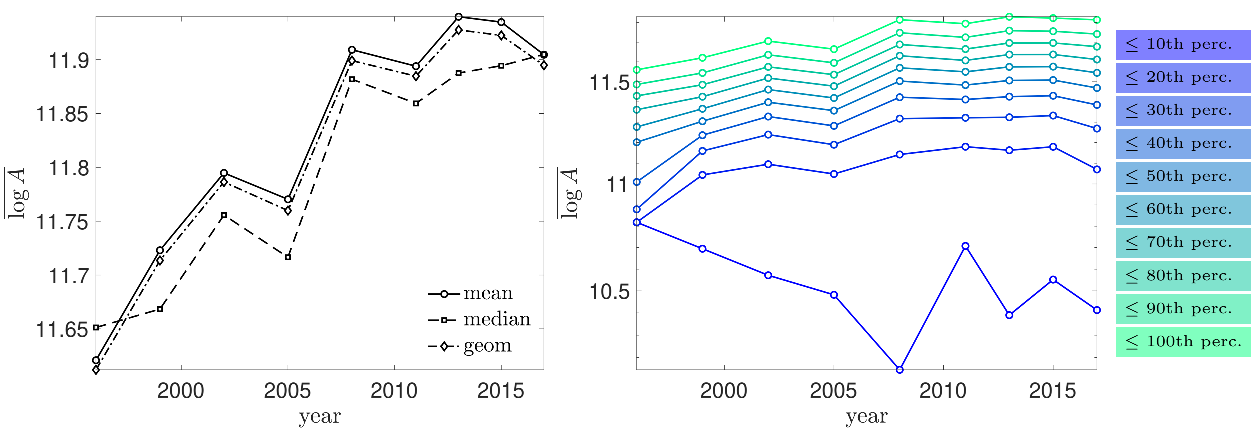


Figure 6: (Left panel) The average productivity $\overline{A}$ over the years ranging from 1996 to 2017. Productivity is measured as value added per employee. The arithmetic, geometric mean and the median are shown and all statistics follow a similar trend. (Right panel) The average productivity (arithmetic mean) for different productivity percentiles (that is, the average is computed over all firms with productivity below the 10th, 20th, ..., 100th percentile).

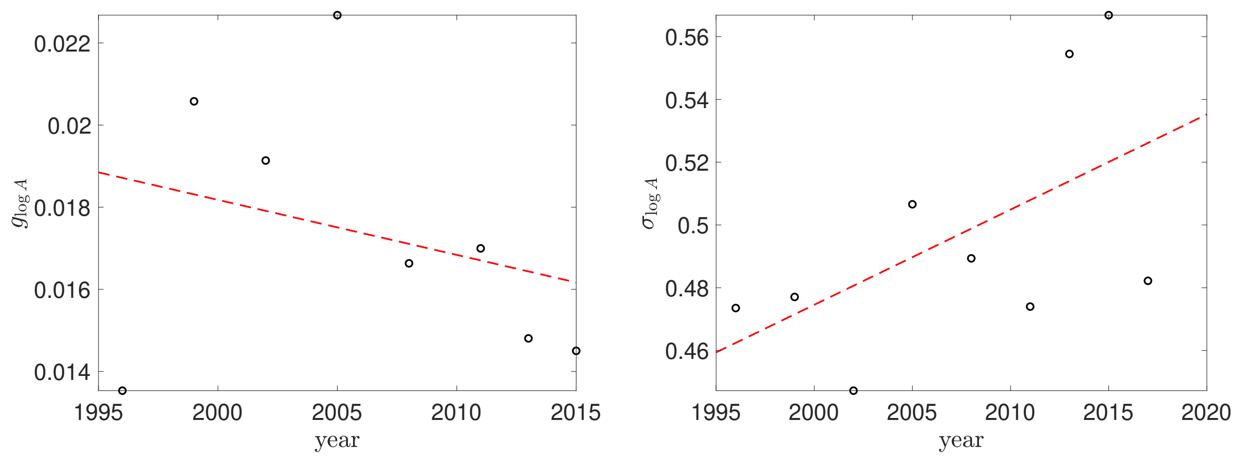


Figure 7: (Left panel) The productivity growth rate, g_A over the years from 1996 to 2017. The growth rate shows a declining trend over time. (Right panel) The productivity standard deviation, σ_A, over the years from 1996 to 2017. The standard deviation shows an increasing trend over time. Productivity is measured as value added per employee. Dashed lines indicate a linear regression fit.

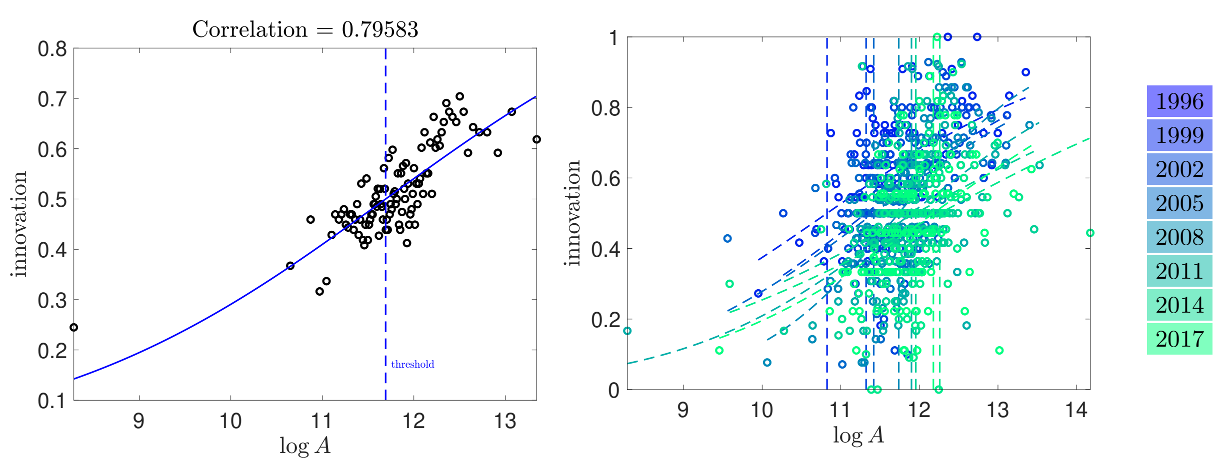


Figure 8: (Left panel) The innovation decision over productivity pooled across all years of observation. Firms with higher productivity tend to innovate more. The solid line shows the fit of a logistic function. The threshold (dashed line) indicates the lowest productivity level at which the innovation decision is more likely than 50%. (Right panel) The innovation decision over productivity across different years. Productivity is measured as value added per employee.

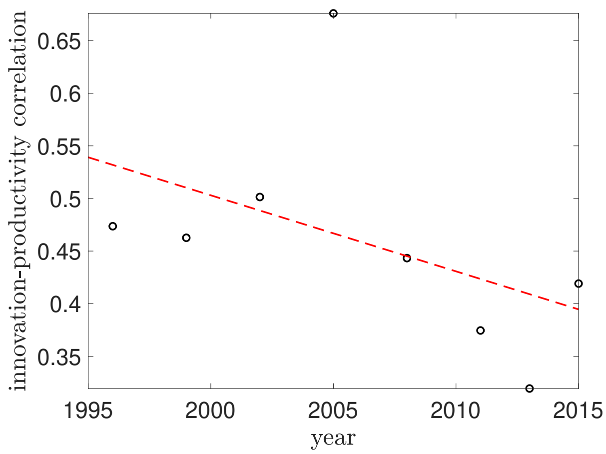


Figure 9: (Left panel) The correlation of the innovation decision with productivity. Except for the year 2005, the correlation shows a declining trend. (Right panel) The innovation threshold is increasing over the years of observation. Productivity is measured as value added per employee. Dashed lines indicate a linear regression fit.

5 Survival Analysis

5.1 The Cox proportional hazard model

Survival analysis concerns the analysis of the time until the occurrence of an event. In our context, the years until firms exit from R&D, or, vice versa, the years until firms enter R&D. OLS is not appropriate in the context of survival analysis because the time until failure is often distributed in a non-symmetrical way, and OLS is not robust to this type of violations (Cleves et al., 2008). Survival analysis substitutes the normality assumption of the error terms required by OLS with something more appropriate.

More specifically, we use the Cox proportional hazard model or Cox (1972) model. It is a semi-parametric model, which means that it does not require any assumptions about the distribution of failure times. Instead, the Cox model relies on an ordering of failure times. It does not matter when exactly the failures took place, but only whether a specific failure occurred before or after another failure.

$$h(t) = h_0(t)exp(\beta_1 x_1 + ... + \beta_k x_k)$$

The core of every survival analysis is the hazard rate h(t). Cleves et al. (2008) define the hazard rate as the (limiting) probability that the failure event occurs in a given time interval, conditional upon the subject having survived to the beginning of that interval, divided by the width of the interval. The hazard rate is an intuitive measure of the risk of failure a subject faces at a given point in time. For example, the risk of dying for humans follows a hazard rate that has the shape of a bathtub. After birth, the hazard of dying is relatively high, with the hazard falling for a time until it reaches a low point. It then remains flat for a long time and starts to increase constantly only in higher age, until with an age of more than 100 it starts to reach values near infinity. In our context, the hazard rate measures the risk a firm, which is still R&D-active until time t, faces to exit R&D from t to t+1. The hazard rate has units 1/t, that is, it measures an increase or decrease in the risk of failure in one unit of analysis time. The interpretation of the hazard rate is therefore the following: for a hazard rate of 0.05/t, if we were to continue for one unit of time, we would expect 0.05 exits for a subject. For instance, if the hazard rate of exit from R&D is 0.05/year, and we were to continue for another year, the average firm, which has survived as R&D-active until right now, would face a probability of 0.05 to exit R&D.

In survival analysis, the survivor function S(t) is usually of interest, too. It measures the probability of the subjects surviving beyond time t. In our context, the survivor function is the probability that a given firm does not exit R&D before the year t. The survivor function S(t) is equal to 1 at the year where the survival analysis starts and monotonically decreases to 0 with increasing time t.

There will be two types of analyses. The first analysis investigates exit from R&D. The analysis starts out in 1996 with all firms that have R&D activities. Firms with R&D activities that appear later on in the panel join the analysis, too, and we explicitly designate them as such. The failure event is when firms exit from R&D. Firms are not allowed to enter into R&D again after they have exited. In other words, we do not allow for multiple failure events: firms cannot exit R&D, enter R&D again, and then exit R&D from anew. Firms staying R&D active until the analysis ends (the year 2017) or firms that drop from the analysis because of non-response are coded as censored observations. The second analysis investigates entry into R&D. The analysis starts out in 1996 as well, with all firms that have no R&D-activities. Those firms with no R&D activities that appear later on in the panel also join the analysis. The failure event is when firms enter R&D. Firms are not allowed to enter into R&D, exit R&D, and then enter R&D again. That is, also the second analysis does not allow for multiple failures. Those firms that have no R&D activities until the analysis ends (the year 2017) or drop from the analysis because of non-response are coded as censored.

5.2 Descriptive Statistics

Table 1 and 2 show the descriptive statistics for the first survival analysis, where firms initially pursue R&D activities and the failure event is exit from R&D. In total, we have 3,948 observations, consisting of 1,963 firms, which result in all together 999 exits from R&D. These 1,963 firms have been at risk for existing from R&D for an average of about 8 years, whereby the distribution of failure times is somewhat right-skewed with a median of 6 years. Table 2 shows the estimates from the Kaplan-Meier estimator, which in the absence of covariates is equivalent to the Cox model (Cleves et al., 2008). The fifth column shows the survivor function, which gives the probability that a firm survives past the respective years listed in the first column. For example, the probability that a firm, which pursues R&D in 1996, still pursues R&D in 2005 is still 55.7%, whereas in 2017 it reduces to 21.0

Table 3 and 4 show the descriptive statistics for the second survival analysis. They mirror Table 1 and 2 in their structure, only that here we are looking at firms that start out with no R&D activities and the failure event is entry into R&D. In total, we have 5,614 observations, consisting of 2,581 firms (there are more firms without than with R&D), which result in 654 entries into R&D. The 2,581 firms have been at risk for entering R&D for an average of about 8 years, whereby the distribution is also somewhat right-skewed, with a median of 6 years. The fifth column of 4 shows the survivor function from the Kaplan-Meier estimator again. For example, the probability that a firm, which had no R&D in 1996, still has no R&D in 2005 is 69.7%, whereas until the year 2017 this number

Table 1: Analysis of exit from R&D

Observations	3,948				
Firms	1,963				
Exits from R&D	999				
		Mean	Min.	Median	Max.
Time at risk:		8.06 years	2 years	6 years	21 years

Table 2: Kaplan Meier estimator: Exit from R&D

Year	Firms at risk	Exits	Ent.-Cens.	Survivor function	Std. Error
1999	762	140	-339	0.8163	0.014
2002	961	169	-178	0.6727	0.0153
2005	970	167	-89	0.5569	0.0151
2008	892	142	39	0.4682	0.0144
2011	711	123	-78	0.3872	0.0136
2013	666	107	60	0.325	0.0127
2015	499	84	114	0.2703	0.0119
2017	301	67	234	0.2101	0.0113

reduces to 49.2%. The probability of surviving without R&D in Table 2 is considerably higher than the probability of surviving with R&D in Table 4 because there are more exits from R&D than entries into R&D in the sample.

5.3 Exit from R&D

5.3.1 Baseline model

The analysis starts out with all R&D-active firms, which can either continue or exit R&D. Given these R&D-active firms, we investigate whether they exit R&D and if so how long they lasted until they exit. Non-R&D active firms and firms, which enter into R&D, are not part of the analysis.

Table 3: Analysis of entry into R&D

Observations	5,614				
Firms	2,581				
Exits from R&D	654				
		Mean	Min.	Median	Max.
Time at risk:		8.09 years	2 years	6 years	21 years

Table 4: Kaplan Meier estimator: Entry into R&D

Year	Firms at risk	Entries	Ent.-Cens.	Survivor function	Std. Error
1999	639	82	-511	0.8717	0.0132
2002	1068	130	-227	0.7656	0.0145
2005	1165	104	-207	0.6972	0.0147
2008	1268	103	91	0.6406	0.0145
2011	1074	87	-105	0.5887	0.0144
2013	1092	63	117	0.5547	0.0142
2015	912	40	249	0.5304	0.0141
2017	623	45	578	0.4921	0.0142

The analysis time starts in 1996; however, firms can join the analysis at a later point in time, given that they are R&D active. As laid out before, to fix the start of the analysis to 1996 allows firms to experience the same risk of exiting R&D in each year. The pressure on firms to exit from R&D could vary widely between each calendar year. For instance, pressure might have been higher in the period 2008-2011, where the financial crisis happened, than in other periods. Accounting for the time line in this way allows calculating a hazard rate for exiting R&D for all available years of our panel. Note again that firms entering the analysis later only contribute to the risk profile of the years that they have been part of.

If we have no explanatory information about a firm (covariates), the reasons why such a firm exits from R&D remain unknown. In each period, firms are exposed to a certain risk that they exit from R&D. For example, it may be too expensive to conduct R&D. By using covariates, we can model the risk that firms exit R&D, as certain firm characteristics could affect the risk of firms for their decisions to exit from R&D. The setup of survival analysis is very common in medical studies. We have a set of patients (R&D firms) and we look at how long they survive, whereas characteristics of the patients (R&D firms) can increase or reduce the risk of death (exit from R&D). For instance, large firms could have a higher probability to remain R&D active.

Table 5 shows the coefficients for the explanatory variables of the estimated baseline Cox model. The coefficients in the table are exponentiated. They show by how much the explanatory variables shift the hazard of exiting from R&D. For instance, the hazard ratio of employment of 0.781 in the first column means that a one percent increase in employment reduces the hazard rate of exiting from R&D by 0.219%. The five explanatory variables shown in Table 5 are all statistically significant and negatively correlated with the hazard rate. Ceteris paribus, higher values of employment, productivity, human capital, exports, and technological potential are all associated with a lower hazard rate of exiting from R&D. Vice versa, a decrease in these five explanatory variables is associated with a higher hazard rate of exiting R&D. Importantly, the exponentiated coefficients in Table 5 change only marginally when they are part of the same model. Note that all specifications control for 2-digit industry dummies.

Figure 10 shows the survivor function of Model V from Table 5 over the period 1999 to 2017 with all explanatory variables held at their means. The displayed survivor function is thus relevant for an average firm. The Y-axis displays the values of the survivor function, that is, the probability of the

Table 5: Baseline R&D exit

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
Ln(Employment)	0.781***	0.783***	0.773***	0.793***	0.799***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)
Ln(Value added/employee)		0.812***	0.841***	0.874**	0.877**
		(0.048)	(0.049)	(0.051)	(0.051)
Share academics			0.984***	0.987***	0.988***
			(0.003)	(0.003)	(0.003)
Share higher education			0.992***	0.994***	0.994**
			(0.002)	(0.002)	(0.002)
Export share				0.992***	0.993***
				(0.001)	(0.001)
Technological potential					0.890***
					(0.023)
Observations	3,946	3,938	3,938	3,910	3,910
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

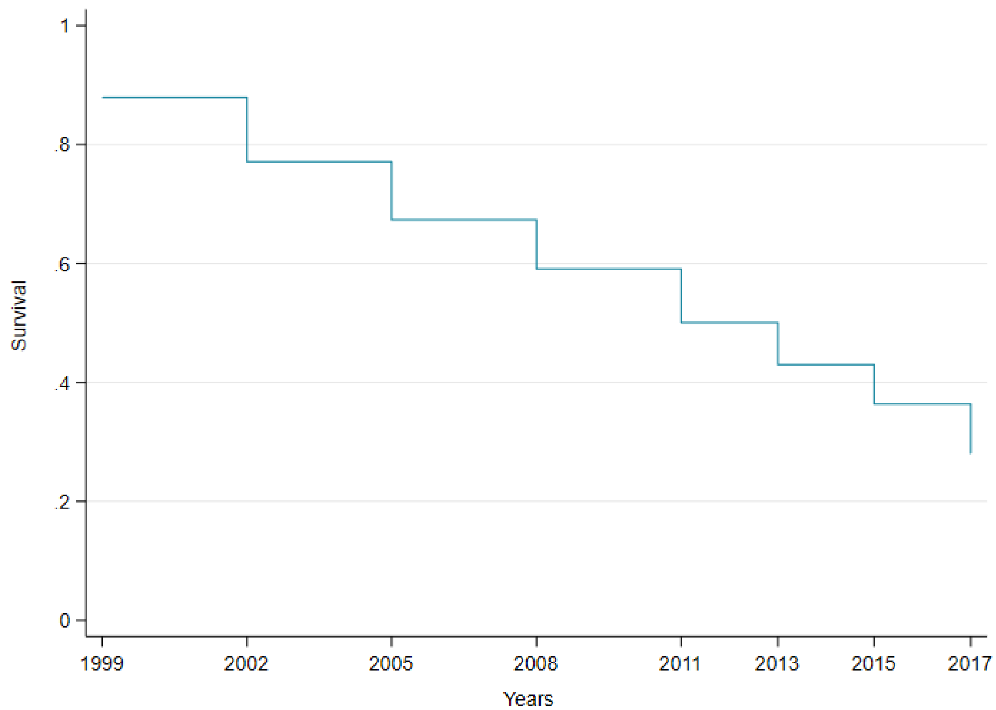
firms to make it until the years displayed on the X-axis[3]. The probability that firms maintain R&D activities decreases with time. For instance, an average firm that has R&D in 1996 experiences a probability to still have R&D in 2008 of about 0.6. In 2017, this probability further drops to less than 0.3.

Importantly, the survivor function falls much faster than the propensity of firms to conduct R&D (i.e., the share of firms in the economy that pursue R&D), because we only allow for exits from R&D and not for entries into R&D. The propensity to conduct R&D sums up firms that exit and firms that enter R&D, which makes its share more stable over time. The survivor function, in contrast, sums up only exits from R&D. Note that because there are more firms exiting than entering R&D, the propensity to conduct R&D falls over time as well, albeit, since it is more balanced by entries into R&D, at a lower rate.

Figure 11 shows the evolution of the hazard rate of Model V from Table 5. To calculate the hazard rate, the covariates shown in Table 5 are again held at their means. The hazard rate thus takes on the values of an average firm. Figure 11 shows that the average firm initially faces a hazard of 0.04 of exiting R&D per year. This means that, for instance in 2005, among those firms still R&D-active, the average firm faces a probability of exiting R&D of 0.04. Interestingly, while the hazard rate remains constant from 2002 to 2008, it almost doubles until 2015. Hence, given that they are still R&D-active, the probability of firms exiting R&D strongly increases after the financial crisis

[3] The survivor function is a step function because the Cox model estimates it non-parametrically; the survivor function is estimated directly from the data and no distribution of failure times is imposed.

Figure 10: Survivor functions exit from R&D

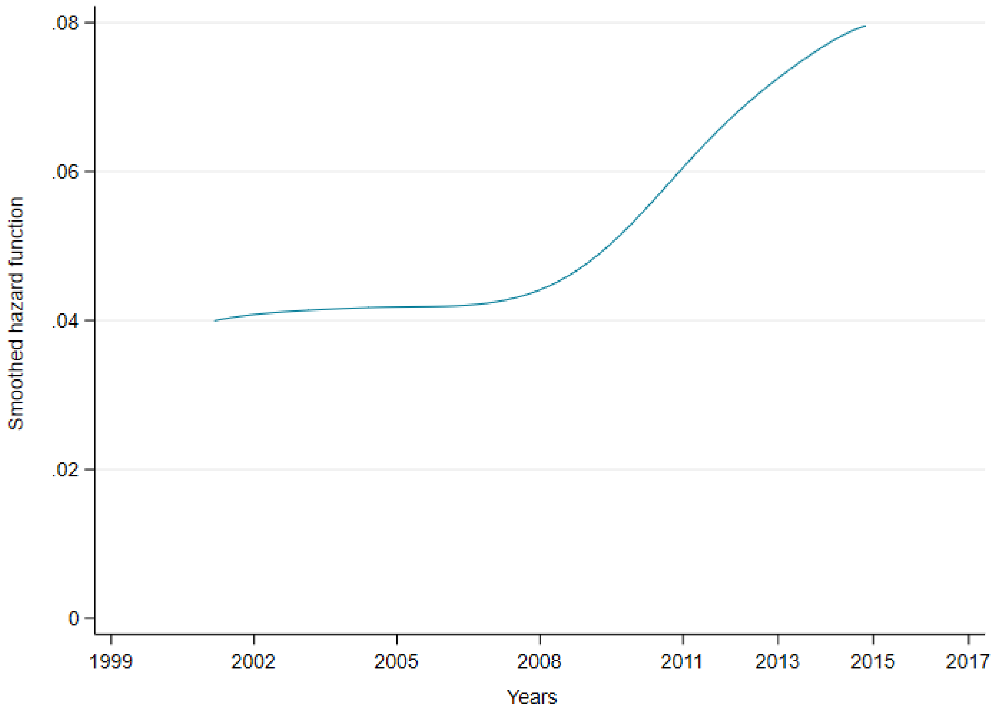


in 2008/09. Today firms apparently face an even harder environment when it comes to maintaining R&D than they have experienced in the past.

5.3.2 Competitive vs. less competitive firms

In this part we will investigate how the survival of different types of firms evolves over time. We will build three types of firms: those with low values for the firm characteristics displayed in Table 6, those with high values, and those firms whose values of the firm characteristics in Table 6 lie at the median. Table 6 shows the exact values these three types of firms take on. Except for employment, they are the values falling in the 10% quantile, the median, and the 90% quantile of the distribution of the variables. Together, the five variables of Table 6 correspond to how competitive firms are; firms with high values of the variables are generally more competitive than firms with low values. Figure 12 plots the survivor functions of the three types of firms. It shows that those firms scoring high on the firm characteristics in Table 6 are much less likely to exit R&D than firms scoring low on them. While at the end of the analysis time in 2017 firms scoring high on the firm characteristics in Table 6 have a probability of about 70% to be still R&D-active, firms scoring low on these characteristics have a probability of less than 10% to still be R&D active. Firms at the median of the firm characteristics in Table 6 have a probability of about 30% to still be R&D active. Hence, firms that are large, productive, human capital intensive, export oriented, and operate in fields with a high technological potential are much less likely to exit from R&D than firms from the opposite spectrum scoring low

Figure 11: Hazard functions exit from R&D



on all the variables in Table 6, especially when compared with those firms taking on the 10% smallest values. Among the firms scoring high on the firm characteristics, the exits from R&D are minimal, particularly when considering that they are also more prone to enter R&D (see Chapter 4). In sum, we can conclude competitive firms thus show a much lower probability to exit from R&D than less competitive firms.

Figure 13 displays the hazard rate for the three types of firms shown in Table 6. While firms scoring high on the firm characteristics in Table 6 face a very low hazard rate during the entire time span, firms scoring low on the firm characteristics face a much higher hazard, and correspondingly experience a much steeper increase in the last 10 years of the analysis time. Firms whose values lie at the median of the variables fall right in between the two other survivor functions; they are closer to the firms with the higher values though. While in 2015 firms scoring low on the firm characteristics in Table 6 face a hazard rate of over 0.15, firms scoring high on the characteristics face a hazard of only 0.02. Firms laying at the median of the variables face a hazard of about 0.08. In 2015, the probability of exiting from R&D was almost 8 times higher for less competitive firms than for competitive firms.

5.3.3 Innovation input

In Table 16, the explanatory variables from Table 5 are retained, although they are not displayed again, while we introduce five new explanatory variables that measure the inputs firms invest in

Table 6: Competitive vs. less competitive firms

Variable	10% smallest	50%: median	90% highest
Employment	50	250	500
Value added/employee	90'000	145'000	275'000
Share academics	0	3	20
Share higher eduction	3	12	33
Export share	0	33	95
Technological potential	2	3	5

Figure 12: Survivor functions competitive vs. less competitive firms

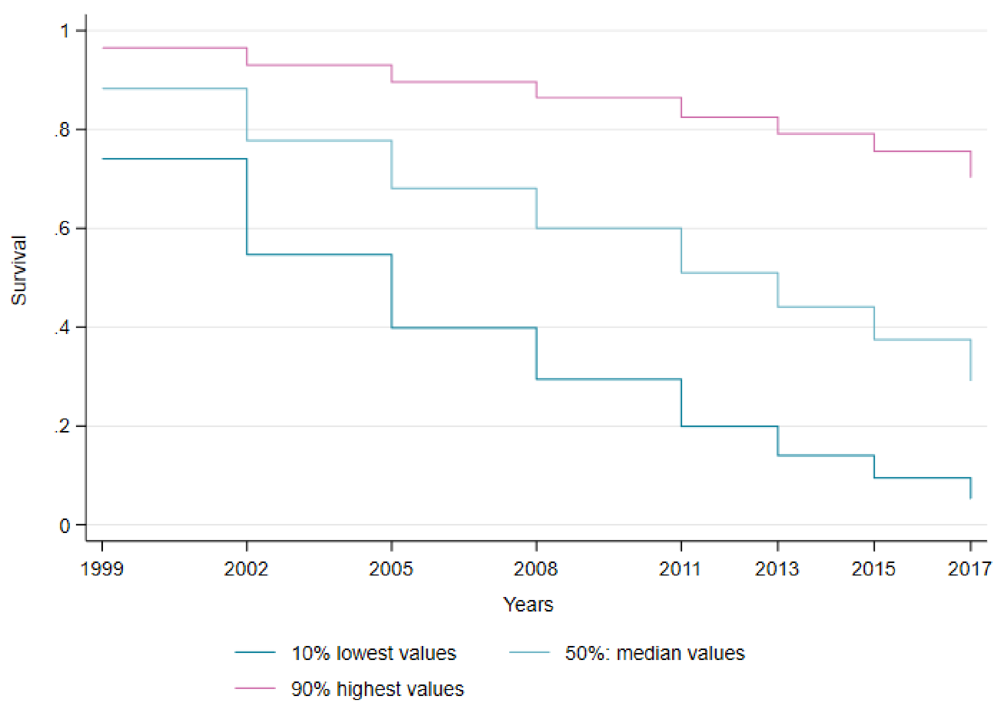
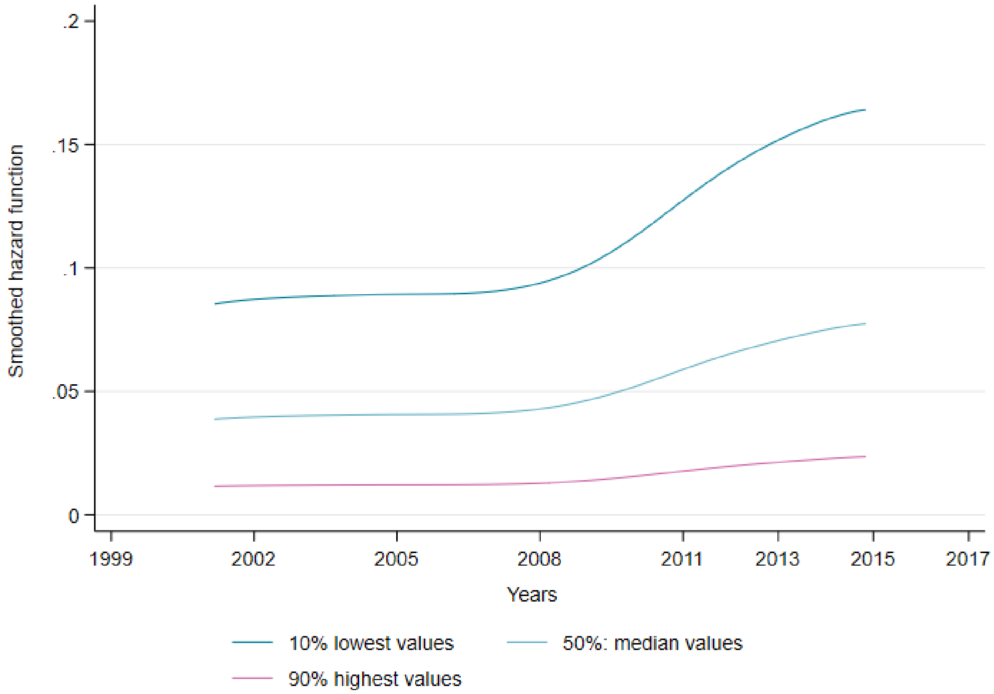


Figure 13: Hazard functions competitive vs. less competitive firms



their R&D activities: R&D expenditures, R&D cooperation with universities, R&D cooperation with other research institutes, and domestic and international innovation support. These variables are all negatively associated with the hazard rate of exiting R&D. The exponentiated coefficients in Table 7 are statistically significant and below the value of one. For instance, the exponentiated coefficient of R&D expenditures is 0.888, which means that a one percent increase in R&D expenditures reduces the hazard of exiting R&D by 0.112%. The only variable that is not statistically significant is international innovation support. In contrast, domestic innovation support shows an exponentiated coefficient of 0.504, which implies that the hazard of exiting R&D is only half as large for firms enjoying domestic innovation support than for firms without such support. Note that all specifications in Table 7 also control for 2-digit industry dummies. In sum, Table 7 shows that firms intensively engaged in their R&D activities also show a much lower hazard of exiting these R&D activities; those firms that are fully committed to R&D also stay in R&D.

As an example, Figure 14 shows the survivor functions of Table 7 for firms with R&D cooperation with universities. All variables are held at their means except for the variable that designates R&D cooperation with universities. The two survivor functions indicate what we have already observed in Table 7: firms that pursue R&D cooperation with universities show a slower decrease in the probability to exit R&D. The survivor function for these firms falls much slower than the survivor function for firms that do not pursue R&D cooperation with universities.

Table 7: Innovation input R&D exit

VARIABLES	(1) Model I	(2) Model II	(3) Model III
Ln(R&D expenditures)	0.888***		
	(0.017)		
R&D Cooperation universities		0.713***	
		(0.075)	
R&D Cooperation other research		0.786**	
		(0.094)	
Innovation support domestic			0.504***
			(0.114)
Innovation support international			0.633
			(0.220)
Observations	3,763	3,879	1,789
Control variables	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Figure 14: Survivor functions R&D cooperation with universities

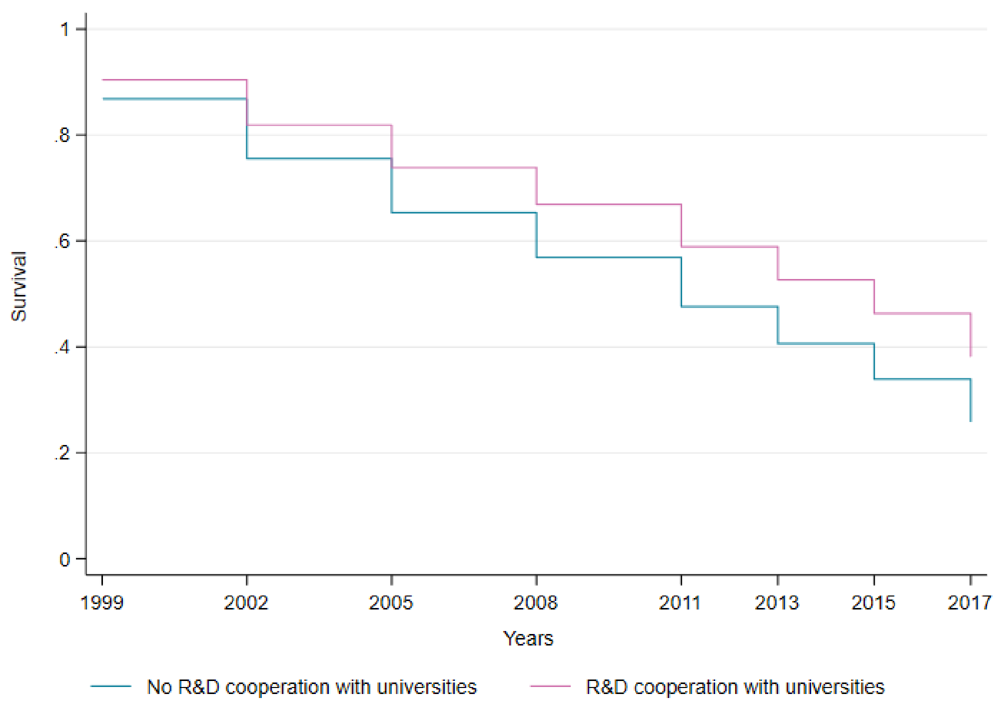


Table 8: Innovation output R&D exit

VARIABLES	(1) Model I	(2) Model II	(3) Model III
Sales share new products/services	0.993**		
	(0.003)		
Sales share improved products/services	0.998		
	(0.002)		
Sales share new to market		0.992*	
		(0.004)	
Sales share new to firm		0.997	
		(0.003)	
World first product/service			0.829**
			(0.069)
Observations	3,052	1,284	2,225
Control variables	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

5.3.4 Innovation output

In Table 8, we extend the baseline Cox model from Table 5 with variables measuring the innovation output of firms. The results show that firms with a higher sales share of new products, but not with a higher sales share of improved products, have a lower hazard of exiting from R&D. Similarly, firms with a higher sales share of products new to the market, but not with a higher sales share of products new to the firm, have lower hazard rates. However, this latter difference is less pronounced. Finally, firms that indicate they have introduced a world first product show a 17.1% lower hazard rate than firms without such a product. In analogy to the innovation input in Table 7, Table 8 shows that firms very active in producing highly innovative output are considerably less likely to exit from R&D.

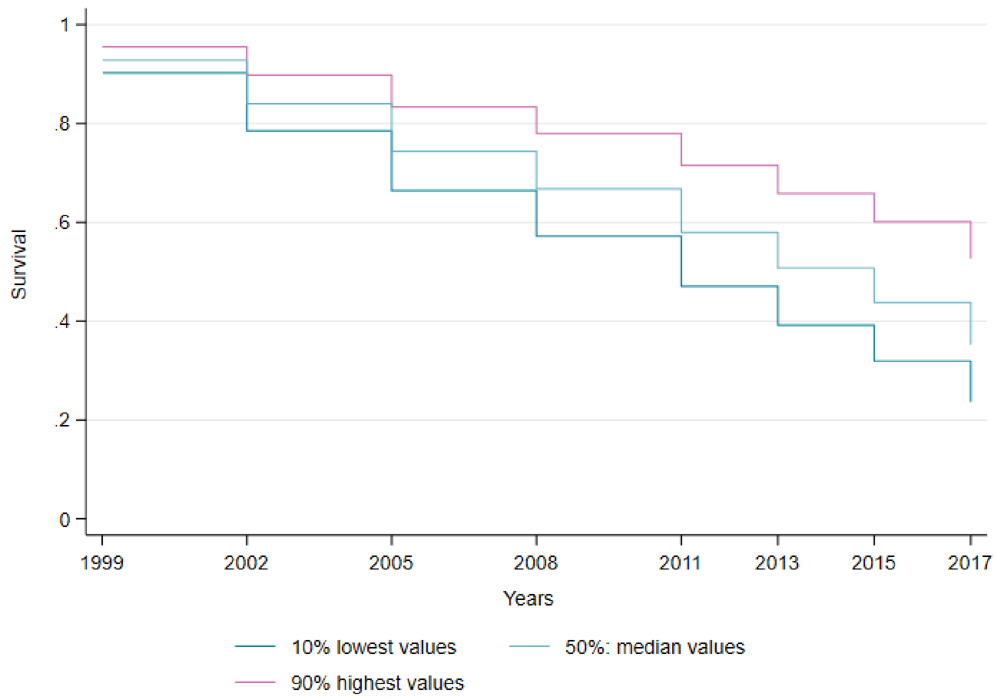
5.3.5 Deep innovators vs. less deep innovators

Figure 15 shows the survivor function for the so-called deep innovators. These are firms that are intensively engaged in their R&D activities, and which at the same also produce highly innovative output. Table 9 shows the characterization of deep innovators in comparison with less deep innovators. The less deep innovators have R&D expenditures that correspond to 1% of their sales, whereas deep innovators have R&D expenditures that corresponds to 10% of their sales. The average firms whose values are held at the median have share of R&D expenditures in sales of 2%. Moreover, deep innovators have a sales share of new products of 40% and a sales share of improved products of 50%. Deep innovators thus replace their product portfolio very rapidly, every 3-4 years. Less deep

Table 9: Deep innovators vs. less deep innovators

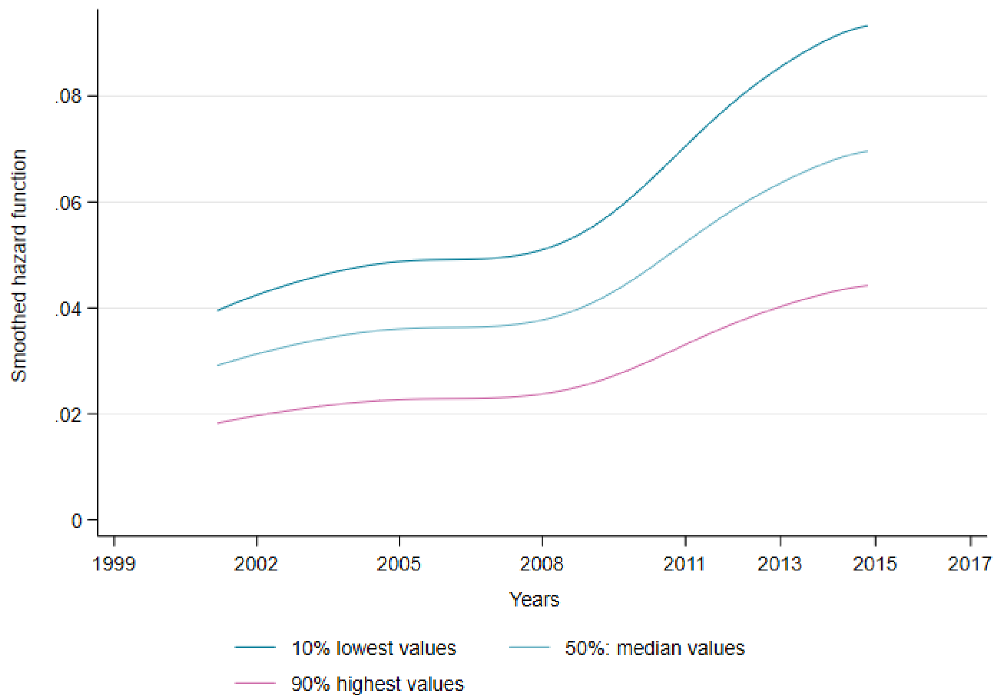
Variable	10% smallest	50%: median	90% highest
R&D expenditures/sales	0.001	0.020	0.010
Sales share new products	0	10	40
Sales share improved products	0	15	50

Figure 15: Survivor functions deep vs. less deep innovators



innovators, in contrast, have a sales share of new products and a sales share of improved products of 0%. The average firms held at the median have values of 10% and 15% for these two variables, respectively. The analysis incorporates the baseline Cox model of Table 5 and only these tree variables, because the other innovation input and output variables are only available in some of the cross-sections, which would reduce the power of the analysis too much. Figure 15 shows that deep innovators experience a much less pronounced decrease in the probability to exit from R&D than less deep innovators. The same pattern emerges in Figure 16 deep innovators show much lower hazard rates than less deep innovators. We can thus conclude that firms that are intensively engaged in their innovation activities and produce a highly innovative output are much less likely to exit from R&D than firms falling short of these important characteristics.

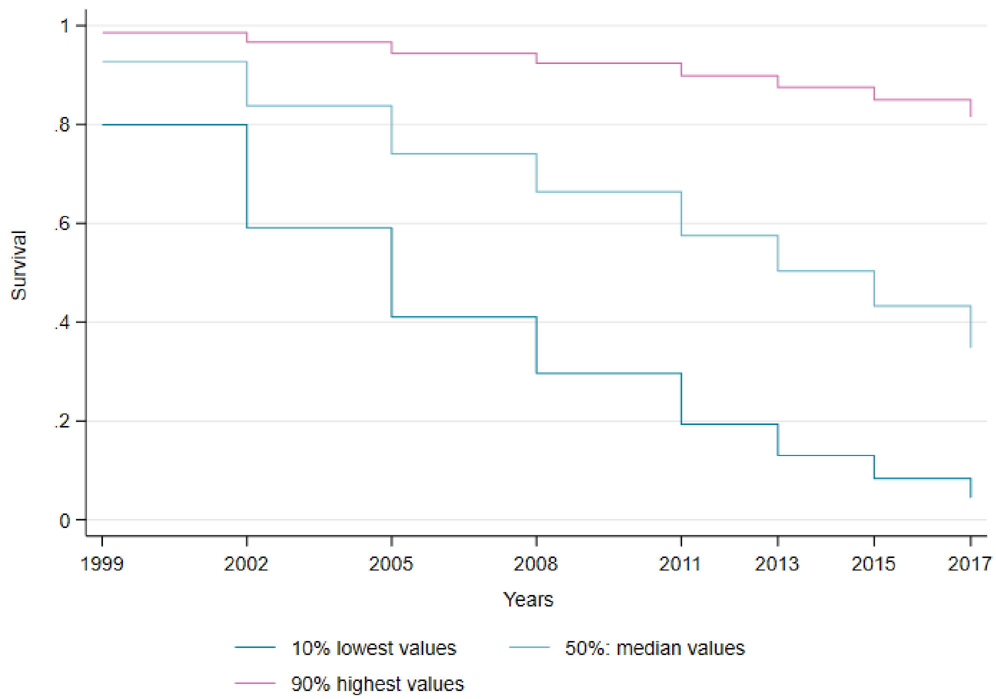
Figure 16: Hazard functions deep vs. less deep innovators



5.3.6 Competitive and innovative: a diverging pattern

Figure 17 shows the survivor functions from the baseline Cox model in Table 5 together with the three variables characterizing deep innovators. Thus, in calculating the survivor functions, it combines the dimensions: the pink line displays the survivor function for firms that are competitive, deep innovators, while the blue line displays the survivor function for firms that are less competitive, less deep innovators. The light blue line displays the survivor curve for the firms whose values for all variables are held at the median. Figure 17 thus combines the high and low values shown in Table 6 and 9, respectively. It shows that competitive, deep innovators are very unlikely to exit from R&D, even over a time span of 21 years. From 1996 to 2017, they face an over 80% probability to still be R&D active. The next chapter will show that competitive, deep innovators also experienced a pronounced shift toward entering R&D. Hence, the proportion of competitive, deep innovators with R&D activities in the economy is even likely to have increased. The development of the hazard rate in Figure 18 shows the same pattern as Figure 17. Competitive firms that are at the same time also deep innovators show a very low hazard of exiting R&D, although this rate has also slightly increased since the year 2011. The hazard is still very low though. In sharp contrast, less competitive, less innovative firms are practically certain to exit from R&D over a time span of 21 years. The hazard rate for this type of firms is very high right from the beginning and strongly increases over time. Thus, we have a divergence between competitive, deep innovators and less competitive firms, less deep innovators. Only few firms from the former category exit R&D, whereas the probability to exit

Figure 17: Survivor functions: a diverging pattern

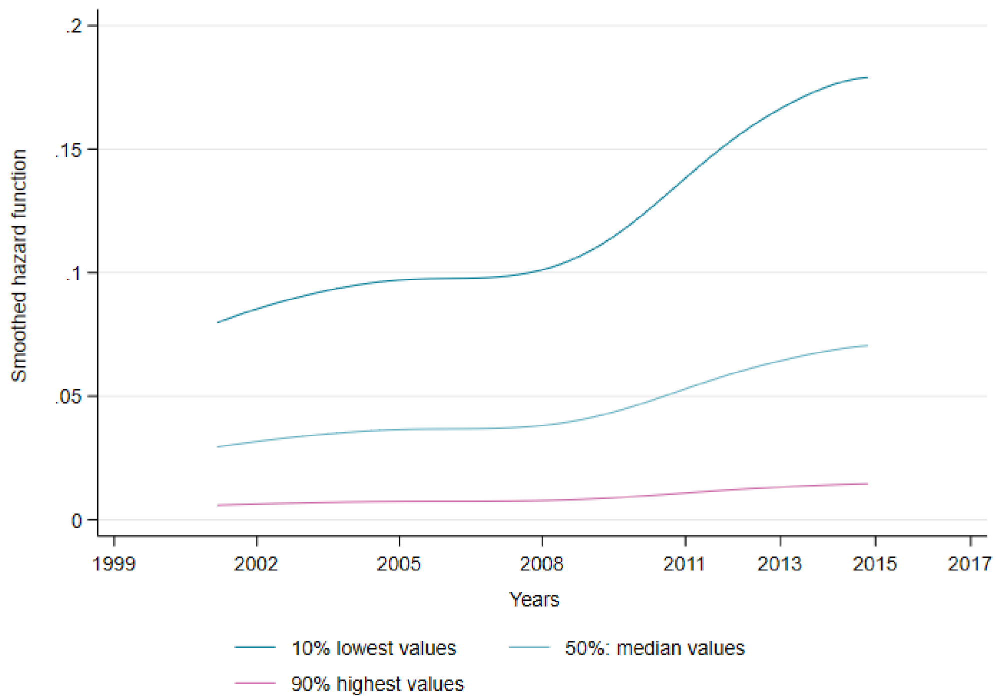


for a firm in the latter category reaches near certainty over the time span from 1996 to 2017. The survivor function of the group of firms whose values are held at the mean falls in-between the other two survivor functions, while being somewhat closer to the survivor function of the competitive, deep innovators. The hazard of the average firm to exit from R&D is thus a bit closer to the hazard of exiting for the group of competitive, deep innovators.

5.3.7 Hampering factors R&D exit

Table 10 shows different groups of hampering factors and how they influence the time until firms exit from R&D. The Cox model again contains all variables from the baseline model shown in Table 7 as control variables, including the 2-digit industry dummies. Interestingly, hampering factors one would consider as very relevant, such as high costs and payback time are not statistically significant in this model. In contrast, firms indicating equity constraints and copiability as important hampering factors show a significantly higher hazard rate of exiting R&D. Counterintuitively, firms with the hampering factors technological risk and lack of R&D employees show a significantly lower hazard of exiting R&D. A potential explanation of this pattern is that the deep innovators score higher values on these hampering factors, which is the reason why we observe such an inverse correlation with the hazard rate. The only hampering factor that is associated with an increase in the probability to exit from R&D are thus constraints in equity and ease of copiability. The pursuit of R&D activities obviously needs a stable financial basis, while the products and services should not lend themselves

Figure 18: Hazard functions: a diverging pattern



to easy imitation.

5.4 Entry into R&D

5.4.1 Baseline model

This second analysis estimates a Cox model that starts out with all non R&D-active firms, which can all either continue without R&D or enter R&D. It is the mirror image of the analysis in the previous chapter on exiting from R&D and analyzes whether firms enter R&D and if so how long they last until they enter. In this chapter, R&D active firms and firms that do not exit R&D are not part of the analysis. The analysis starts again in 1996, and firms can join the analysis given that they are without R&D activities. Table 11 shows the explanatory variables of the Cox model. It contains the same explanatory variables as Table 6. The coefficients in Table 11 are again exponentiated. They indicate by how much they shift the hazard of entering into R&D. For instance, the hazard ratio of employment of 1.266 means that a one percent increase in employment increases the hazard rate of entering R&D by 0.266%. Interestingly, except for value added per employee, all variables are again statistically significantly associated with the hazard rate. Higher values of employment, human capital, exports, and technological potential are all positively associated with the hazard of entering R&D. This means that the same firm characteristics causing firms to maintain R&D are also the firm characteristics making them more likely to enter R&D. Note that all specifications again control for 2-digit industry dummies.

Table 10: Hampering factors R&D exit

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
High costs	1.014				1.035
	(0.034)				(0.038)
Copiability	1.027				1.060*
	(0.028)				(0.032)
Payback time	0.975				0.985
	(0.033)				(0.037)
Technological risk		0.930**			0.924**
		(0.030)			(0.032)
Market risk		1.047			1.032
		(0.032)			(0.035)
Lack of equity			1.074**		1.069*
			(0.037)		(0.040)
Lack of credit			0.953		0.969
			(0.032)		(0.036)
Lack of R&D personnel				0.939**	0.936**
				(0.030)	(0.030)
Lack of skilled personnel				0.984	0.974
				(0.032)	(0.033)
Observations	3,910	3,898	3,907	3,688	3,688
Control variables	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Table 11: Baseline R&D entry

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
Ln(Employment)	1.266***	1.267***	1.269***	1.254***	1.242***
	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)
Ln(Value added/employee)		1.083	1.055	1.035	1.024
		(0.078)	(0.076)	(0.076)	(0.076)
Share academics			1.010***	1.008**	1.008**
			(0.003)	(0.003)	(0.003)
Share higher education			1.007***	1.007***	1.006**
			(0.003)	(0.003)	(0.003)
Export share				1.004***	1.004***
				(0.001)	(0.001)
Technological potential					1.123***
					(0.039)
Observations	5,613	5,604	5,604	5,541	5,540
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

5.4.2 Competitive vs. less competitive firms

As in Chapter 3.2, Figure 19 displays the survivor functions for more competitive, less competitive firms, and firms with medium competitiveness. Because the sample of non R&D-active firms is different from the sample of R&D active firms, the values for the characterization of competitiveness are somewhat different. The respective 10% smallest values, median values, and 90% highest values are shown in Table 12. In a mirror image to the part on exit from R&D, Figure 19 shows that firms scoring high on the firm characteristics in Table 12 (i.e., the competitive firms) are much more likely to enter R&D than firms scoring low on these firm characteristics. Note that the survivor function is again decreasing; it can by definition only decrease, even though we are here concerned with entries into R&D and not with exits from R&D. It measures how long non-R&D active firm can survive without entering R&D.

At the end of the analysis time in 2017, firms with high values for the variables in Table 12 have a probability of about 80% to enter R&D (you have to subtract the value of the survivor function from one). In contrast, firms with low values for the variables have a low probability of less than 40% to enter R&D. Firms with values held at the median still face a more than 50% to enter R&D. The hazard rates in Figure 20 show a somewhat different pattern than in the previous chapter. While competitive firms are more likely to enter R&D than less competitive firms, the hazard rate for all three types of firms are slightly decreasing. This means that the probability for all types of firms to enter R&D has slightly decreased over time, given that they made it this far. This observation stands in sharp contrast with the development for the hazard rate to exit from R&D, which has

Table 12: Competitive vs. less competitive firms

Variable	10% smallest	50%: median	90% highest
Employment	50	250	500
Value added/employee	75'000	130'000	265'000
Share academics	0	0	13
Share higher eduction	0	10	30
Export share	0	0	40
Technological potential	1	2	4

strongly increased over time.

In sum, firms that are large, human capital intensive, export oriented, and operate in technologically active fields are much more likely to enter R&D than firms scoring low on all these variables. Because the propensity to pursue R&D (i.e., the share of firms with R&D activities) has decreased over time, we know that the entries into R&D were not sufficient to compensate the exits from R&D we have observed. Moreover, in chapter 5, we will run a robustness check of how entries and exits compare when we exclude the switchers, those firms that enter and exit repeatedly.

Note that we cannot pursue an investigation of innovation input and innovation output for the entry into R&D, because non R&D-active firms do not have positive R&D expenditures and also show much less variation in how innovative they are. In short, we have too few deep innovators that do not purse R&D activities. The data therefore does not allow for an identification of these effects. These would be counterfactual possibilities we do not observe, neither in the data nor in the real world.

5.4.3 Hampering factors R&D entry

Table 13 shows the same hampering factors as in the chapter on exit from R&D and investigates how they influence the time until entry into R&D. The Cox model contains all variables from the model shown in Table 11 as control variables, including the 2-digit industry dummies. Table 13 shows the same patterns as the analysis of the hampering factors on exit from R&D, just, as expected, in the opposite direction. Notably, most of the hampering factors are only moderately correlated with entry into R&D. A higher payback time is associated with a significantly lower hazard of entering R&D. In addition, the lack of R&D employees is here positively associated with the hazard rate. As in the previous chapter, this result suggests that this hampering factor proxies for deep innovators. All other explanatory variables are not statistically significantly associated with the hazard of entering R&D. Most important, equity constraint seem therefore less important for entering R&D than they are for maintaining R&D. While firms with equity constraints are more likely to exit from R&D, the presence of equity itself does not spur entry into R&D. This means that the availability of sufficient equity is asymmetric. A lack of equity can move firms to give up R&D, but the availability of equity does not incline them to enter it. Equity is therefore a necessary but not a sufficient condition for the pursuit of R&D activities.

Figure 19: Survivor functions: competitive vs. less competitive firms

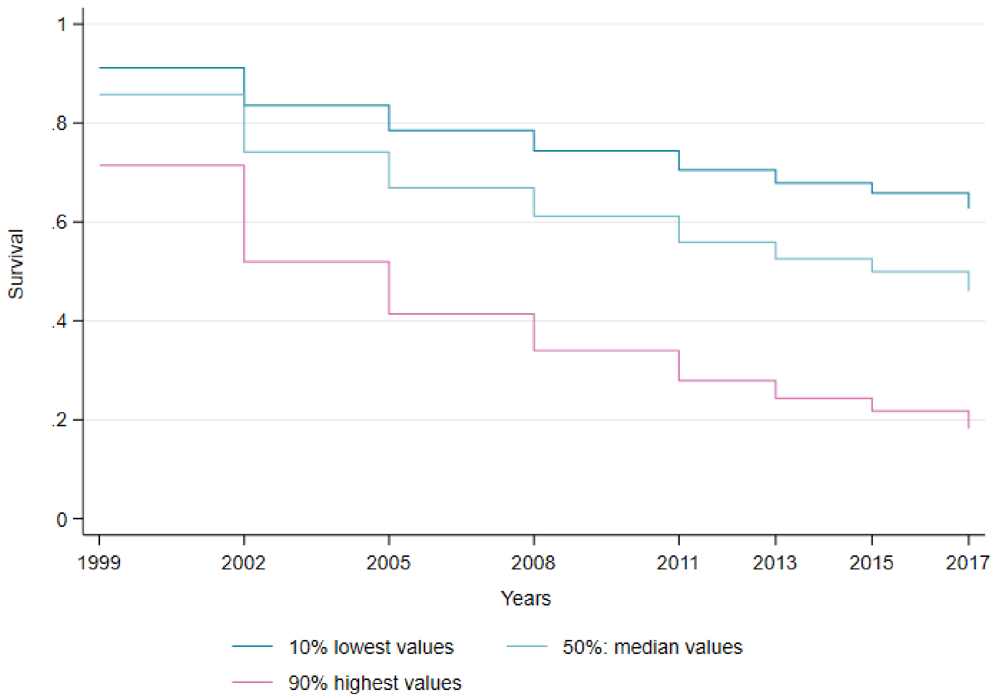


Figure 20: Hazard functions: competitive vs. less competitive firms

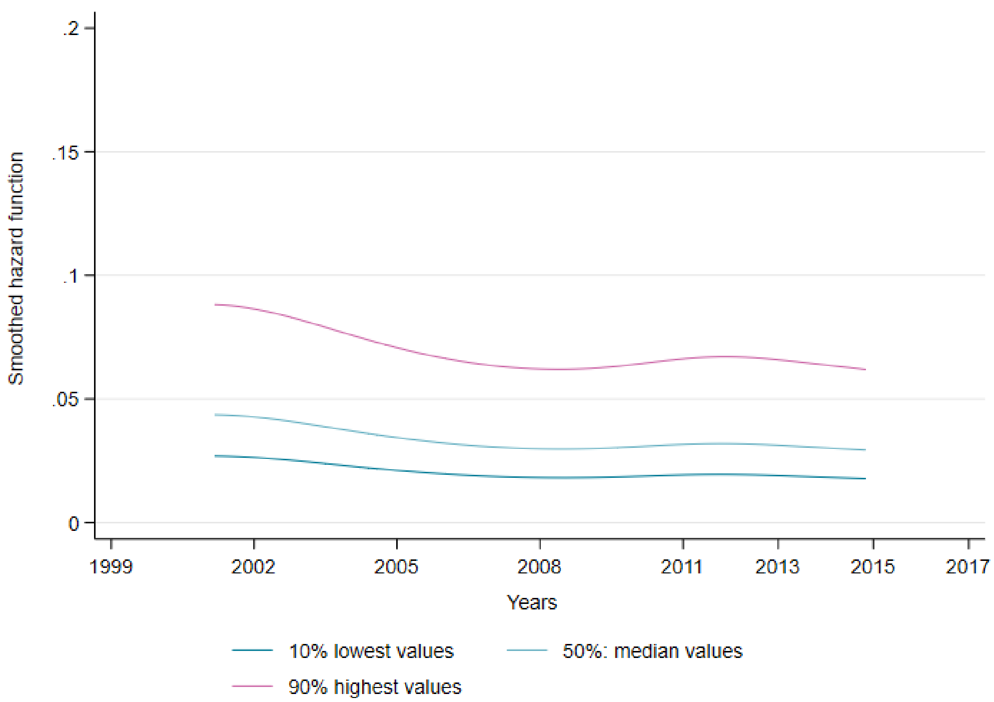


Table 13: Hampering factors R&D entry

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
High costs	1.067				1.030
	(0.053)				(0.057)
Copiability	1.077*				1.016
	(0.043)				(0.048)
Payback time	0.930				0.901*
	(0.047)				(0.052)
Technological risk		1.085*			1.078
		(0.052)			(0.058)
Market risk		1.069			1.084
		(0.050)			(0.059)
Lack of equity			1.034		1.013
			(0.060)		(0.066)
Lack of credit			0.991		0.976
			(0.060)		(0.064)
Lack of R&D personnel				1.143***	1.118**
				(0.053)	(0.053)
Lack of skilled personnel				0.939	0.925
				(0.045)	(0.046)
Observations	5,539	5,457	5,536	5,177	5,176
Control variables	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Table 14: Exit from and entry into R&D for switchers

Analysis of exit from R&D	
Observations	3,494
Firms	1,676
Exits from R&D	712

Analysis of entry into R&D	
Observations	5,122
Firms	2,258
Entry into R&D	331

5.5 Robustness checks

5.5.1 Without R&D switchers

In the previous two chapters, the switchers, that is, those firms that exit R&D and then enter it again were included to the analysis, too. The switchers constitute relevant information about the time until failure as well, as they are true exits and entries, even though they are more short-lived than permanent exits and entries since they do revert again. However, because the exits and entries of switchers are more short-lived and the initial decision is reverted over time, one might argue that the switchers are less important, as their decisions to enter or exit R&D seem not to be final. In the first chapter, we argued that firms that do not appear as switchers could potentially be so, too, and we just did not observe them for a long enough time span. Hence, the boundary between the two is fuzzy and we kept them in the analysis. However, to estimate the influence of these switchers, this chapter pursues the very same analysis as the in the previous chapters but excludes the switchers. The switchers make up 610 firms, which either exit R&D and enter it again or enter R&D and exit it again; they are thus distributed over both the analysis of exit from R&D and entry in R&D. Table 14 shows the descriptive statistics without the switchers. Excluding them reduces the observations for the analysis of exit and entry by 454 and 492, respectively, as many of the switchers appear in both analysis.

Table 15 shows that the results for exit from R&D are practically identical to the results of the baseline Cox model in Table 5. The changes in the coefficients are minor and the subtraction of the switchers from the analysis sample did not affect the results in any way. In the same vein, Table 16 shows that the results for entry into R&D are also practically identical to the results of the baseline Cox model in Table 11, even though the reduction in observations is much larger in this latter case of entry into R&D. From these two tables, we can conclude that the exclusion of the switchers does not alter the results in any substantial way. The coefficients stay more or less the same in both types of analysis.

As a further robustness check, Figure 17 from Section 3.6 is reproduced, where competitive and deep innovators are combined and compared with less competitive and less deep innovators, but this time without the switchers. We can only do this for exit from R&D, as we do not have sufficient data on the innovativeness of non R&D-active firms, and we thus cannot pursue this analysis for entry into R&D. The patterns in Figure 21 are practically identical to the patterns in Figure 17.

Table 15: Baseline model exit from R&D for switchers

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
Ln(Employment)	0.748***	0.751***	0.737***	0.758***	0.763***
	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)
Ln(Value added/employee)		0.815***	0.833***	0.867**	0.871**
		(0.053)	(0.056)	(0.058)	(0.059)
Share academics			0.983***	0.987***	0.988***
			(0.003)	(0.003)	(0.003)
Share higher education			0.990***	0.992***	0.993***
			(0.003)	(0.003)	(0.003)
Export share				0.991***	0.991***
				(0.001)	(0.001)
Technological potential					0.897***
					(0.028)
Observations	3,492	3,488	3,488	3,462	3,462
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

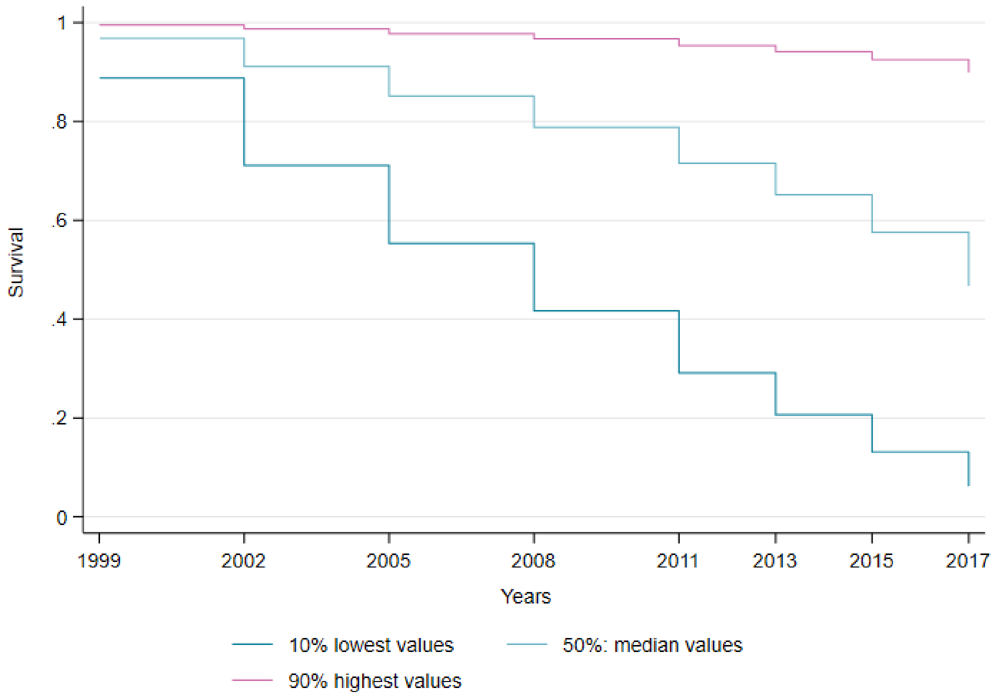
*** p<0.01, ** p<0.05, * p<0.1

Table 16: Baseline model entry into R&D for switchers

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V
Ln(Employment)	1.324***	1.325***	1.326***	1.304***	1.285***
	(0.051)	(0.052)	(0.051)	(0.053)	(0.053)
Ln(Value added/employee)		1.143	1.092	1.069	1.042
		(0.124)	(0.117)	(0.115)	(0.113)
Share academics			1.014***	1.012***	1.011**
			(0.005)	(0.005)	(0.005)
Share higher education			1.009***	1.009***	1.008**
			(0.003)	(0.003)	(0.003)
Export share				1.005***	1.005**
				(0.002)	(0.002)
Technological potential					1.211***
					(0.060)
Observations	5,121	5,114	5,114	5,060	5,060
Industry fixed effects	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

Figure 21: Survivor functions for switchers

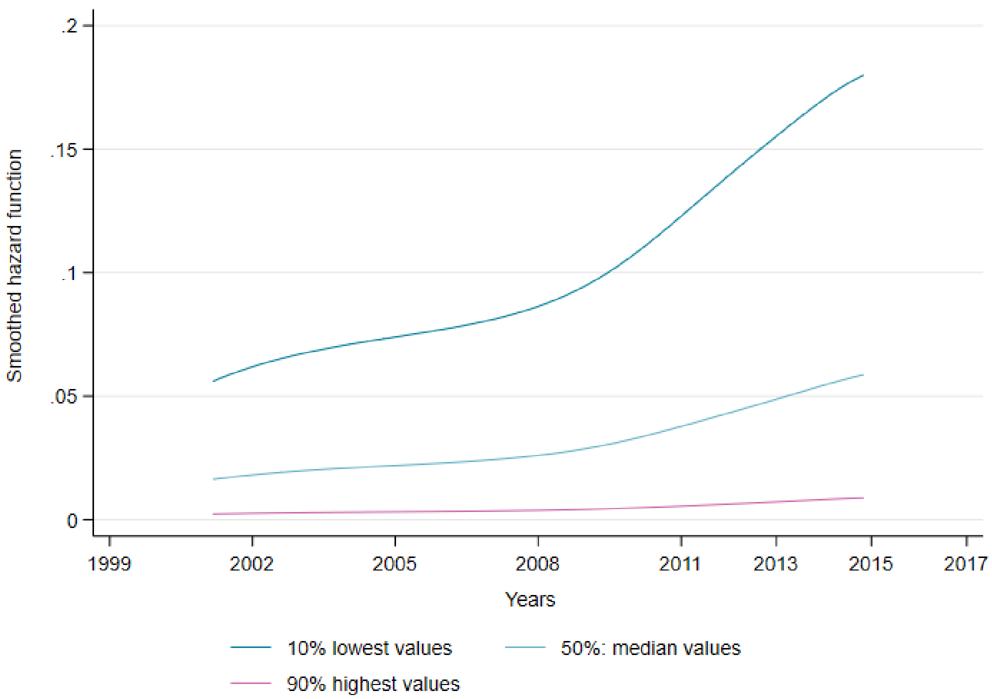


The only difference is that the survivor functions are somewhat flatter, as there are less exits from R&D for both groups. We see that the survivor function for the competitive firms that are also deep innovators is practically flat over time. Over a period of more than 21 years, the hazard of exiting for competitive, highly innovative firms does not even amount to 10%. In sharp contrast, less competitive, less innovative firms face an almost certainty to exit R&D over the entire period of 21 years, just as in Figure 17. The hazard rate in Figure 22 shows the same divergence; the differences in the hazard rates between competitive, deep innovators is even somewhat more pronounced than the differences in hazard rates in Figure 18.

5.5.2 Variance due to education

One might object that a high share of employees with a tertiary education implies that the firm must have successful innovation activities and that therefore the analysis of competitive vs. less competitive firms in Chapter 3.2 reduces to a tautology. However, there are many firms, especially in services, which employ high shares of skilled personnel and nevertheless do not pursue R&D activities, such as in the financial sector. Highly qualified personnel may be a necessary but it is not a sufficient condition for the pursuit of R&D activities. Nonetheless, to test whether our results may to a large degree be determined by this potential problem, we conduct a robustness check where we hold constant the share of employees with a tertiary education between competitive and less competitive firms. The values for the firms who lie at the median of the firm characteristics in Table

Figure 22: Hazard functions for switchers



6 are retained.

Figure 23 displays the survivor curves for competitive vs. less competitive firms using the same covariates as in Table 6 in Chapter 3.2. In contrast to the survivor functions in Figure 13, however, Figure 23 uses the same high levels of tertiary education for the less competitive firms as for the competitive firms. That is, the less competitive firms are assumed to have the same high share of academics of 20% and the same share of employees with higher education of 33% as the competitive firms. Figure 23 demonstrates that the probability for the less competitive firms to not exit from R&D until 2017 increases from about 5% to about 15%, while the probability to survive for the other two groups of firms stays, of course, the same. In the same manner, we can display the three survivor functions by inserting the low levels of tertiary education from the less competitive firms into the competitive firms. That is, for both the competitive and the less competitive firms, we assume a share of academics of 0% and a share of employees with higher education of 3%. Figure 24 shows that the probability for the competitive firm to not exit from R&D until 2017 now reduces from about 70% to somewhat less than 60%. The survivor function for the less competitive firms and the firms whose values are held at the median stay the same again. Hence, while tertiary education is certainly important for maintaining R&D activities in firms, it is not the sole factor. Its influence is on the hazard of exiting R&D is not more pronounced than the influence of the other four, equally relevant covariates proxying for the firms competitiveness: employment, value added per employee, export share, and technological potential.

Figure 23: Survivor functions: variance due to education

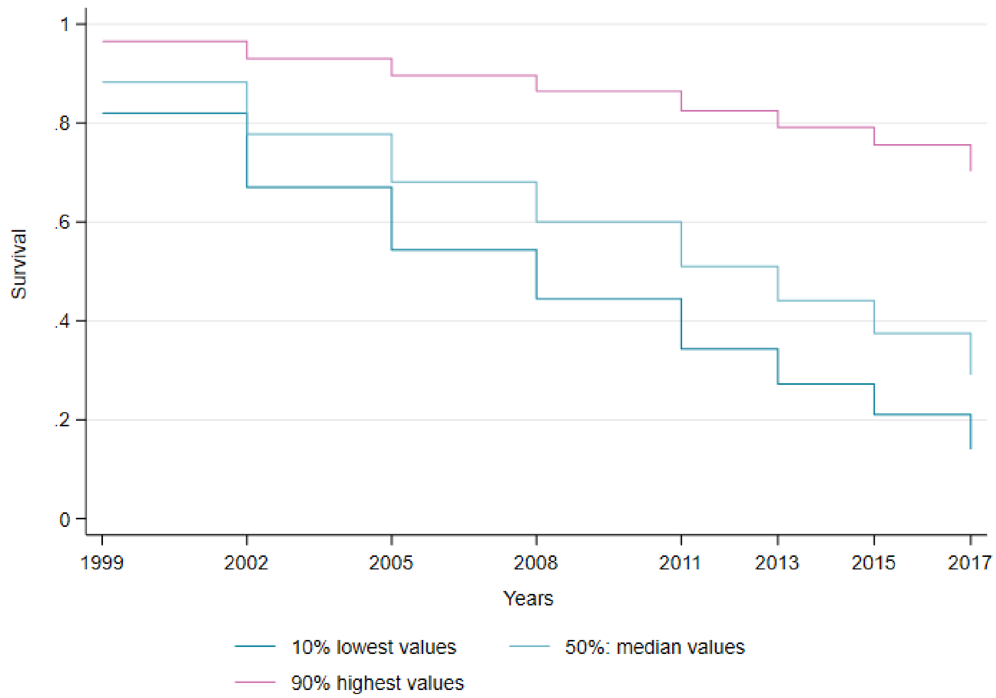


Figure 24: Hazard functions: variance due to education

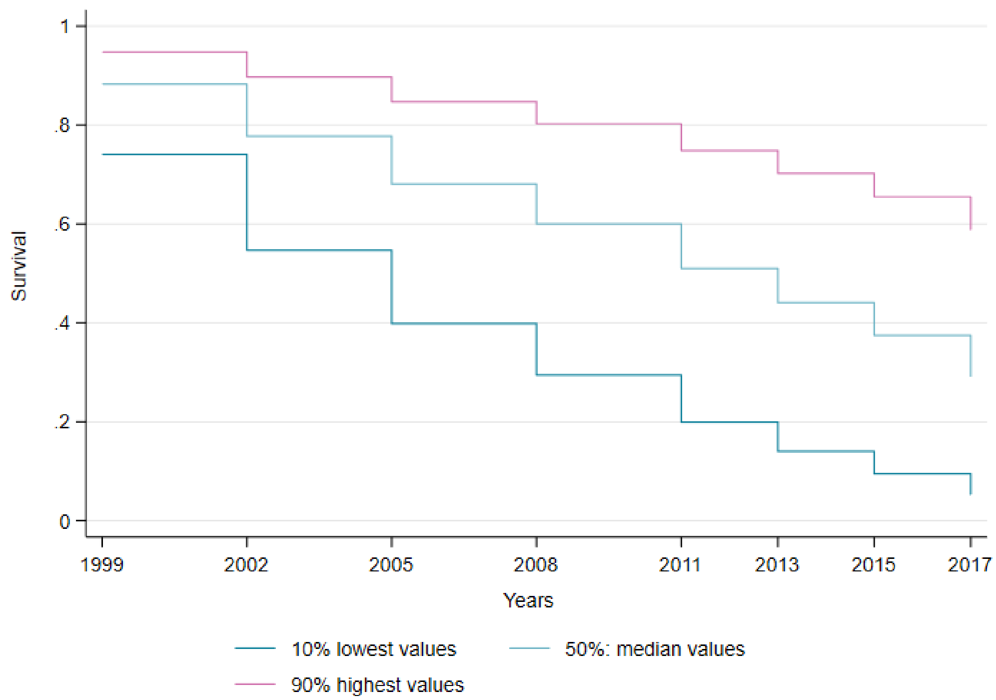
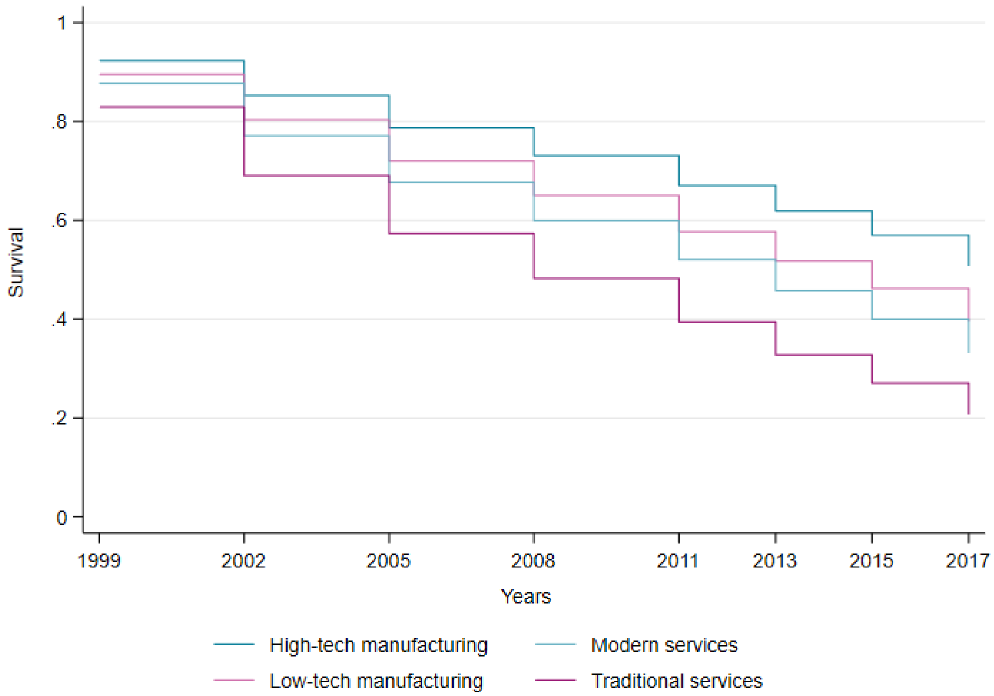


Figure 25: Sectoral differences



5.5.3 Sectoral differences

In this part we will split the survivor function into four subsectors: high-tech manufacturing, low-tech manufacturing, modern services, and traditional services. Figure 25 shows that there are quite some differences in the probability to maintain R&D over the entire time span. Firms in high-tech manufacturing are much more likely to keep their R&D activities than firms in any other subsector. Conversely, firms in traditional services are much more likely to exit from R&D than any other subsector. Interestingly, firms in low-tech manufacturing are somewhat more likely to still be R&D-active than firms in modern services. R&D is thus more important in manufacturing than in services also over time; the former firms are much less likely to exit from R&D. The shifts of the survivor functions visible in Figure 25 show that the subsector to which a firm belongs determines the survival of firms in R&D much less than the competitiveness shown in Figure 12 in Chapter 3.2. The survivor functions in Figure 12 show much broader variation than the survivor functions in Figure 25. This means that the firms' competitiveness is more important in determining eventual decisions to exit from R&D than the subsector to which the firms belong. The gap in survivor functions visible in Figure 25 is more comparable to the gap in survivor functions visible in Figure 15, where we compare deep vs. less deep innovators. Hence, differences in the subsectors to which firms belong yield a similar influence on the probability to stay R&D-active as differences between the two types of deep and less deep innovators.

5.6 International comparison with Netherlands

In this section, we discuss the results of the survival analysis for the Netherlands. Exploiting the comparative nature of the microdata, the estimations are run as close as possible to the estimations for Switzerland. In contrast to Switzerland, where we have data only in 2 and 3 year intervals, in the Netherlands we have data for every year. This results in a smoother survivor function than for Switzerland.

5.6.1 Baseline model

Figure 26 shows the baseline survivor function over the years 2001 to 2016. It is equivalent to Figure 19 in Chapter 5.3.1 for Switzerland. The explanatory variables are again held at their means and the displayed survivor function is thus relevant for an average firm. Like in Figure 19, the probability that firms maintain R&D decreases over time. Overall, the development of the survivor function in the Netherlands is similar to the development of the survivor function in Switzerland. The probability to maintain R&D decreases in Switzerland over the entire time span of almost 20 years to less than 0.3. We see a similar development in the Netherlands. However, Figure 26 shows that the development of the survivor function in the Netherlands is less linear than in Switzerland. The probability to still have R&D decreases very sharply in the first ten years, but this decrease is much less pronounced from 2010 onwards. The survivor function develops almost horizontally. In Figure 19 for Switzerland, in contrast, the survivor function does not decrease less sharply over time, and no flatting is observed. Hence, the rate of exiting from R&D has developed more favorably in the Netherlands than in Switzerland over the past 10 years. This is not a surprising result, given that over this time span the propensity of firms to conduct R&D has increased in the Netherlands whereas it has decreased in Switzerland. It indicates that the increase in the propensity to conduct R&D in the Netherlands is based on a less pronounced rate of exits from R&D.

5.6.2 Competitive and innovative: a diverging pattern

Figure 27 shows how competitiveness and innovativeness lead to a stark contrast in the probability to still be R&D active in the Netherlands, just as it was the case in Figure 17 for Switzerland. Figure 27 also combines two dimensions: the green line displays the survivor function for firms that are competitive, deep innovators, while the blue line displays the survivor function for firms that are less competitive, less deep innovators. The red line displays the survivor function for firms whose values for all variables are held at the median. Like in Figure 17 for Switzerland, Figure 27 shows that competitive, deep innovators are less likely to exit from R&D than less competitive, less deep innovators. However, in Figure 27 the divergence between the two groups is less strong than in Figure 17 for Switzerland. While in the Netherlands less competitive, less deep innovators show a somewhat less pronounced decline in the probability to exit from R&D, competitive, deep innovators show a more pronounced decline compared to Switzerland. This implies that in Switzerland competitive, deep innovators are more resistant in giving up R&D than in the Netherlands, while the opposite holds for less competitive, less deep innovators: they are more likely to give up R&D in Switzerland than in the Netherlands. This indicates that the more favorable trend of the survivor function in Figure 26 is mainly due to the less competitive, less deep innovators, who are comparatively less likely to exit R&D than in Switzerland. One possible reason for this difference is that in the Netherlands there is more support for innovation activities, to which we turn now.

Figure 26: Survivor function exit from R&D for the Netherlands

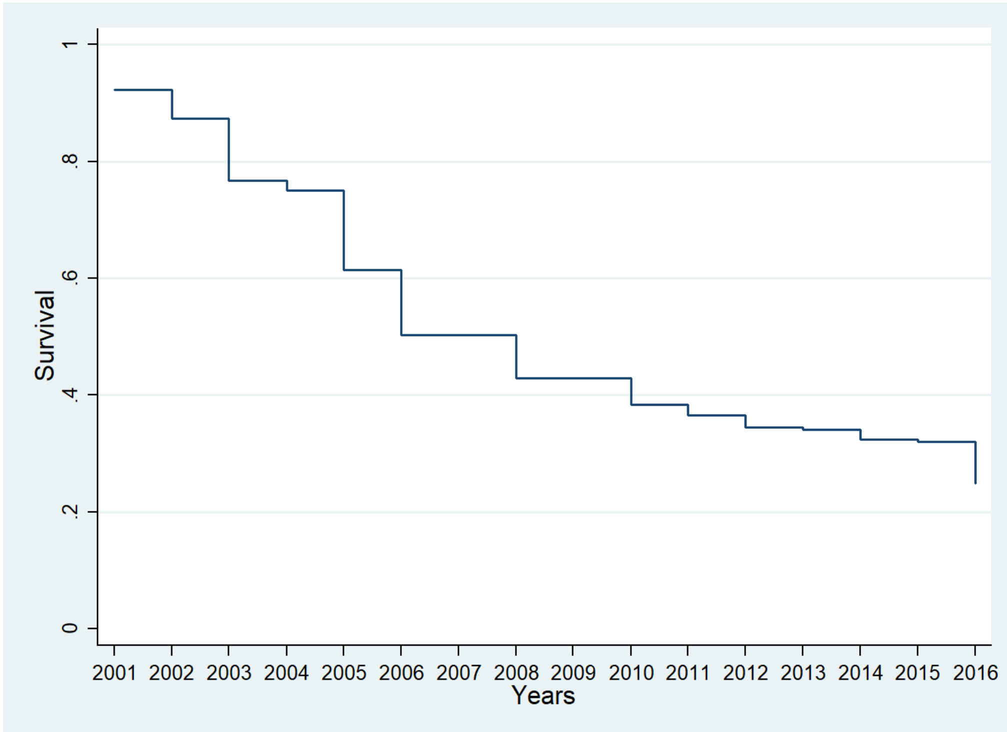
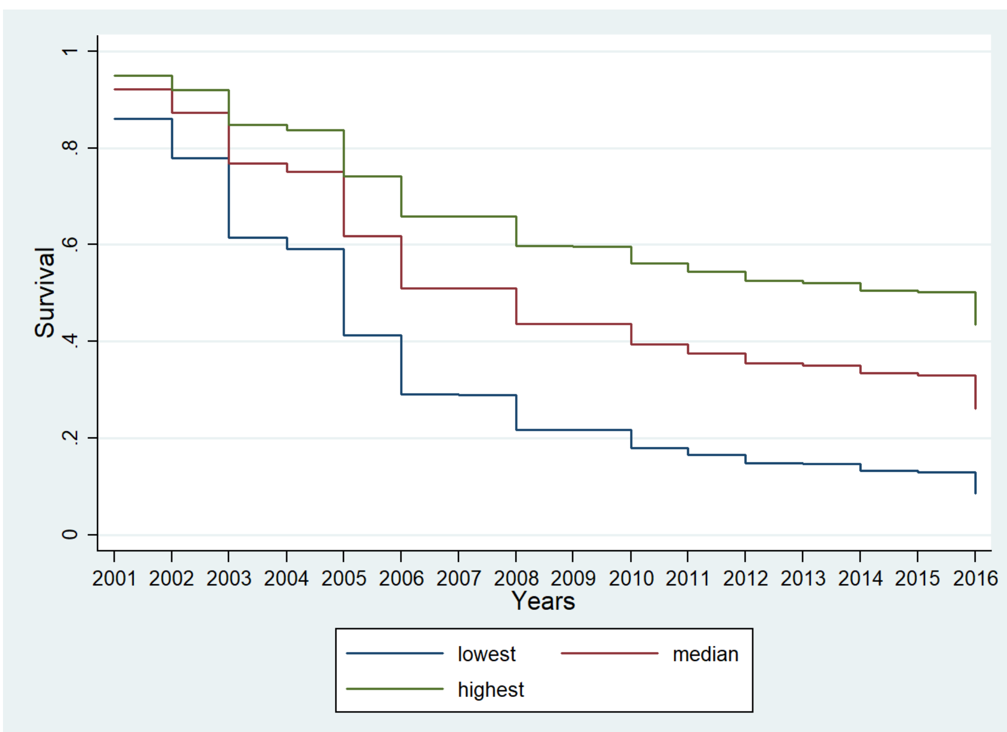


Figure 27: Survivor functions: A diverging pattern for the Netherlands



5.6.3 The innovation box

Differently to Switzerland, in the Netherlands the Dutch government provides incentives for firms to conduct R&D in the form of an "innovation box" and a "patent box". The innovation box tax incentive scheme works on the corporate income tax level, giving firms a lower tax rate on profits that are the result of immaterial assets. It started out in 2007 as a patent box, that is, the only immaterial asset that was allowed was a patent. This broadened to include R&D activities without a patent or other formal property rights in 2010 (renamed into innovation box). Figure 28 shows how the survivor functions develop over time for firms with and firms without access to the innovation box. There is a substantial difference between the two functions: firms that profit from the innovation box show a much less pronounced decrease in the probability to still be R&D active. The difference between the two survivor functions in Figure 28 is comparable to the difference in Figure 27 between competitive, deep innovators and less competitive, less deep innovators. The impact of the innovation box on the probability to still be R&D active might thus be substantial. Figure 28 suggests that the innovation box could have helped to maintain firms R&D active in the Netherlands. However, these results are not necessarily causal. The firms that profit from the innovation box might differ in observable and unobservable characteristics from those firms that cannot profit from it beyond the measurable observables we have included in the estimations, with these hidden characteristics in fact causing the difference. Nonetheless, we do observe a very stark difference between the two groups of firms, suggesting that the innovation box policy instrument might be quite effective in keeping firms in the R&D market.

5.6.4 Hampering factors R&D exit

Table 17 shows how the hampering factors influence the time until firms exit from R&D in the Netherlands. It is the equivalent of Table 10 for Switzerland. The only hampering factors that are statistically significantly related to the hazard rate are competition and lack of collaboration partners. They are both associated with a lower hazard rate to exit from R&D. This indicates that competition and own R&D are both beneficial to maintain R&D in the Netherlands. All other variables in Table 17 are not statistically significant. Most important, a shortage in financial resources, which is associated with a higher risk of exiting R&D in Table 10 for Switzerland, points even into the opposite direction in Table 17. Lack of financing is therefore not related to the exit from R&D activities in the Netherlands.

Figure 28: Survivor functions Innovation Box

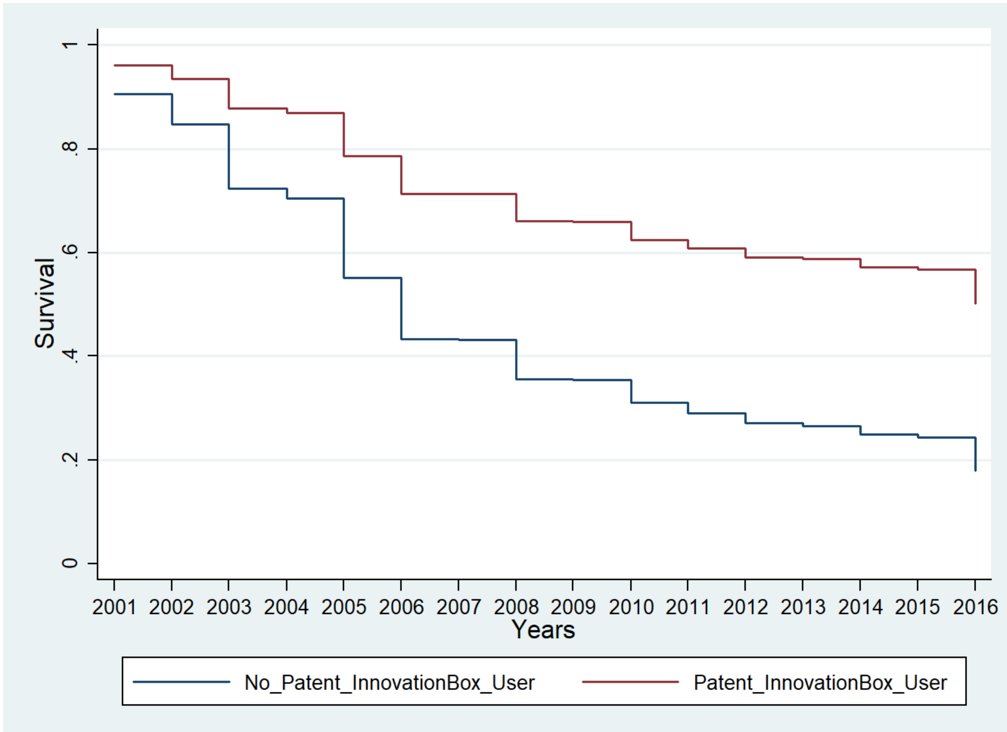


Table 17: Hampering factors R&D exit

VARIABLES	(1) Model I	(2) Model II	(3) Model III	(4) Model IV	(5) Model V	(6) Model VI	(7) Model VII
High costs	0.909						1.046
	(0.084)						(0.142)
Competition		0.842*					0.910
		(0.088)					(0.104)
Market risk			0.894				0.923
			(0.085)				(0.108)
Lack of finance				0.887			0.992
				(0.082)			(0.136)
Lack of skilled personnel					0.867		0.969
					(0.078)		(0.113)
Lack of collab. partners						0.828*	0.871
						(0.090)	(0.102)
Observations	4,648	3,191	4,291	4,656	4,648	2,947	2,934
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*** p<0.01, ** p<0.05, * p<0.1

6 The Determinants of Firms' R&D Decision

In the following we will analyze factors determining the firm's R&D decision. We will first analyze simple correlations in a linear probability model between the R&D decision and firm characteristics affecting in-house R&D success, imitation success, and competition. In contrast to the survival analysis in Chapter 5, the linear probability model is derived from the structural model presented in this chapter. Hence, it allows us to estimate parameters that directly map to a theory about the firm's R&D decision. We will then present this structural model, which allows us to separately identify the effects of the difficulty of making innovations (in-house R&D success), cost of R&D, and technology diffusion (imitation) on the R&D decision. We will use the estimated structural model to perform counterfactual policy simulations and to investigate the effectiveness of R&D policies. A more detailed discussion of the structural model can be found in Appendix B.

6.1 Firms' Profits

Let the profit, $\pi_i(t)$, of firm i time t be given by (König et al., 2020, 2016)

$$\pi_i(t) = \psi \times \underbrace{A_i(t)}_{\text{productivity}} - \underbrace{c_i(t)}_{\text{R\&D cost}} \tag{1}$$

where $\psi > 0$ is a constant. Firms can increase their productivity, $A_i(t)$, (and profits, $\pi_i(t)$), due to in-house R&D or imitating other, more productive firms. The firm's decision to do R&D is based on the expected profits in Equation (1) from choosing in-house R&D or imitation, respectively.

6.2 R&D Decision and Linear Probability Model

Consider a firm i with in-house R&D success probability p_i. Denote by D_i the decision variable of the firm to do R&D. Then one can show that

$$D_i = \underbrace{p_i}_{\substack{\text{innovation}\\ \text{potential}}} \underbrace{-\tilde{\kappa}e^{\theta(\bar{a}-a_i)}}_{\substack{\text{competition}\\ \text{effect}}} \underbrace{-q(1-F_{a_i})}_{\substack{\text{imitation}\\ \text{potential}}} \tag{2}$$

where a_i is the log-productivity of firm i, $\bar{a}$ the average log-productivity, $1 - F_{a_i}$ is the fraction of firms with a productivity larger than a_i, q is the imitation success probability, and $\tilde{\kappa}$ the relative cost of R&D. Equation (6.2) can be estimated with ordinary least squares (OLS) and we will refer to this as the linear probability model (LPM). In an extended model we can also allow for imitation of firms that fail to do successful in-house R&D. We refer to this as "passive imitation" with success probability δ. This model can be estimated with non-linear least squares (NLS).

6.3 Structural Endogenous Growth Model

The simple model introduced in Section 6.2 is indicative for measuring some basic correlations with the R&D decision. But this model might suffer from an endogeneity bias due to productivity (and other variables) being affected by the innovation decision (and vice versa). We therefore introduce a structural model that takes into account the endogenous evolution of the firms' productivities from their innovation and imitation decisions. We characterize the evolution of the productivity distribution with heterogeneous firms (random effects) in terms of their in-house R&D success probabilities.

For the purpose of estimating the parameters of this model, we choose the parameters in such a way as to fit the distributions predicted by the model to the empirical distributions. More precisely, we consider a set of moments (measuring the absolute differences between the model and the data) constructed from the log-productivity distribution and the R&D profile (i.e. the fraction of firms conducting R&D for a given log-productivity level). We then search for the parameters that minimize the distance between the targeted empirical moments and the stationary distribution of the model. This estimation approach is referred to as Simulated Method of Moments (SMM) procedure.

6.4 Estimation Results

In this section we report the estimation results for both, the linear probability model (LPM/NLS) introduced in Section 6.2 and the structural growth model (SMM) introduced in Section 6.3.

The estimation results pooled across the years 1996 to 2017 can be seen in Table 18. Table 19 shows the Nonlinear Least Squares (NLS) parameter estimates across the years from 1999 to 2017.

Table 18: Estimation results pooled across the years 1996 to 2017.

		LPM	NLS		SMM
		w/o passive imitation (1)	w/o firm characteristics (2)	with firm characteristics (3)	with random effects (4)
Innovation	(p)	0.5880*** (0.0203)	0.4999*** (0.0178)	–	0.7962*** (0.1199)
Cost	$(\widetilde{\kappa})$	0.1495*** (0.0211)	0.0737*** (0.0154)	0.0126*** (0.0083)	1.7999*** (0.7443)
Imitation	(q)	0.0781*** (0.0157)	0.6798*** (0.0652)	0.8912*** (0.1000)	0.6986*** (0.1141)
Passive Imitation	(δ)	–	1.0437*** (0.0229)	1.0197*** (0.0141)	0.4307* (0.2382)
Const.	(β_0)	–	–	-3.5608*** (0.1637)	–
Higher Education	(β_1)	–	–	0.0042*** (0.0013)	–
Export (yes/no)	(β_2)	–	–	1.4907*** (0.0417)	–
Competitors: 6-10	(β_3)	–	–	-0.0664 (0.0488)	–
Competitors: 11-15	(β_4)	–	–	-0.0917 (0.0624)	–
Competitors: 16-50	(β_5)	–	–	-0.2738*** (0.0637)	–
Competitors: > 50	(β_6)	–	–	-0.8569*** (0.0628)	–
Technological Potential	(β_7)	–	–	0.3816*** (0.0189)	–
ln(firm age)	(β_8)	–	–	0.0469** (0.0228)	–
KTT Universities	(β_9)	–	–	0.2782*** (0.0187)	–
KTT Competitors	(β_{10})	–	–	-0.0734*** (0.0198)	–

Notes: Model (1) corresponds to the Linear Probability Model (LPM) discussed in Section B.5. Models (2) and (3) correspond to a Nonlinear Least Squares (NLS) estimation procedure with an innovation decision variable as in Equation (8) as a dependent variable. Model (4) is the Simulated Methods of Moments (SMM) estimation algorithm discussed in Section C for which we set $\tilde{a} = 0.45$. In all models we set $\theta = 0.2$ following König et al. (2020). Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses for models (1)–(3). P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. *

Table 19: Linear Probability Model (LPM) and Nonlinear Least Squares (NLS) estimation results across years. The pooled estimation results can be found in columns (1) and (2) of Table 18, respectively.

		1996 (1)	1999 (2)	2002 (3)	2005 (4)	2008 (5)	2011 (6)	2013 (7)	2015 (8)	2017 (9)
		Linear Probability Model (LPM) w/o Passive Imitation								
Innovation	(p)	0.6214***	0.6152***	0.5024***	0.5201***	0.4532***	0.4207***	0.4356***	0.4242***	0.3894***
		(0.1084)	(0.2900)	(0.0204)	(0.0195)	(0.0197)	(0.0188)	(0.0193)	(0.0202)	(0.0217)
Innov. Cost	($\widetilde{\kappa}$)	0.0537	0.1946	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		(0.0862)	(0.1642)	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(1.0000)
Imitation	(q)	0.0631	0.0038	0.1644***	0.2625***	0.1991***	0.1532***	0.2255***	0.2519***	0.2253***
		(0.0747)	(0.1444)	(0.0346)	(0.0322)	(0.0355)	(0.0339)	(0.0362)	(0.0379)	(0.0410)
		Nonlinear Least Squares (NLS) with Passive Imitation								
Innovation	(p)	0.5638***	0.4281***	0.4677***	0.4938***	0.3940***	0.4042***	0.3615***	0.3653***	0.3105***
		(0.0187)	(0.1113)	(0.0952)	(0.0804)	(0.0156)	(0.1051)	(0.0607)	(0.0325)	(0.0747)
Innov. Cost	($\widetilde{\kappa}$)	0.0000	0.0000	0.0000	0.0109	0.0000	0.0040	0.0000	0.0000	0.0000
		(0.0086)	(0.1216)	(0.1040)	(0.0754)	(0.0148)	(0.1154)	(0.0652)	(0.0209)	(0.0807)
Imitation	(q)	0.9706***	0.8991*	0.7855***	0.7547***	0.9651***	0.7411***	0.8419***	0.9340***	0.9298***
		(0.0403)	(0.4596)	(0.2862)	(0.1818)	(0.0355)	(0.3153)	(0.1780)	(0.0805)	(0.1865)
Pass. Imit.	(δ)	0.9790***	0.9694***	0.9161***	0.8903***	0.9249***	0.9297***	0.8300***	0.9134***	0.8828***
		(0.0145)	(0.1221)	(0.1117)	(0.0988)	(0.0330)	(0.1342)	(0.1021)	(0.0424)	(0.0831)

Notes: P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. * Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses.

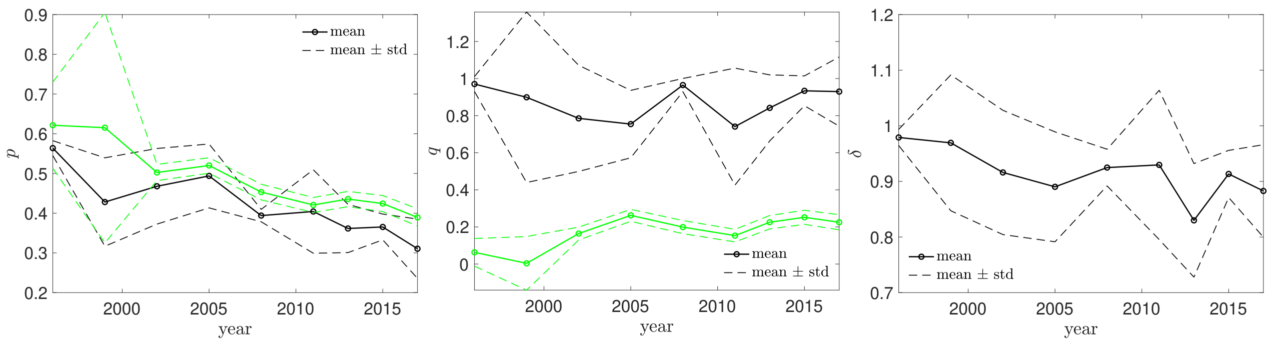


Figure 29: Nonlinear least squares model parameter estimates from Table 19 across different years. The black curve (–) indicates the model with passive imitation (δ), while the green curve (–) indicates the model without passive imitation ($\delta = 0$).

Year-wise parameter estimates from the Nonlinear Least Squares (NLS) model are shown in Figure 29. We observe that the in-house R&D success probability (p) is decreasing over the years, suggesting that it becomes harder to make innovations. Also the imitation parameters q and δ show a weakly decreasing (or stabilizing) trend, indicating attenuated or constant trend in technology diffusion. This is consistent with the recent analysis in Lucking et al. (2019) who show that the diffusion of technological knowledge has been more or less constant over time.

Figure 30 shows the influence of various firm characteristics on the firm's innovation decision that affect the in-house R&D success probability (i.e. observable firm characteristics affecting the in-house R&D success probability p_i in Equation (2)): Technological potential is the worldwide privately and publicly available technological knowledge available to the firm for bringing about marketable innovations. This includes basic scientific knowledge; knowledge of key technologies (e.g. nanotechnology, semiconductor technology, biotechnology, IT, audiovisual techniques, etc.) that is used to implement innovations. Higher education measures the number of people with a degree higher than vocational professional education. We observe that higher education, export orientation, the technological potential and the knowledge and technology transfer (KTT) between universities and the industry have increasing importance on the innovation decision.

Figure 31 shows the effect of competition on the innovation decision. Across all years, competition tends to have a negative effect on innovation (cf. "Schumpeterian effect"; Aghion et al. (2014)), which is becoming stronger the more intense competition is.

Table 20 shows the SMM parameter estimates across the years from 1996 to 2017.

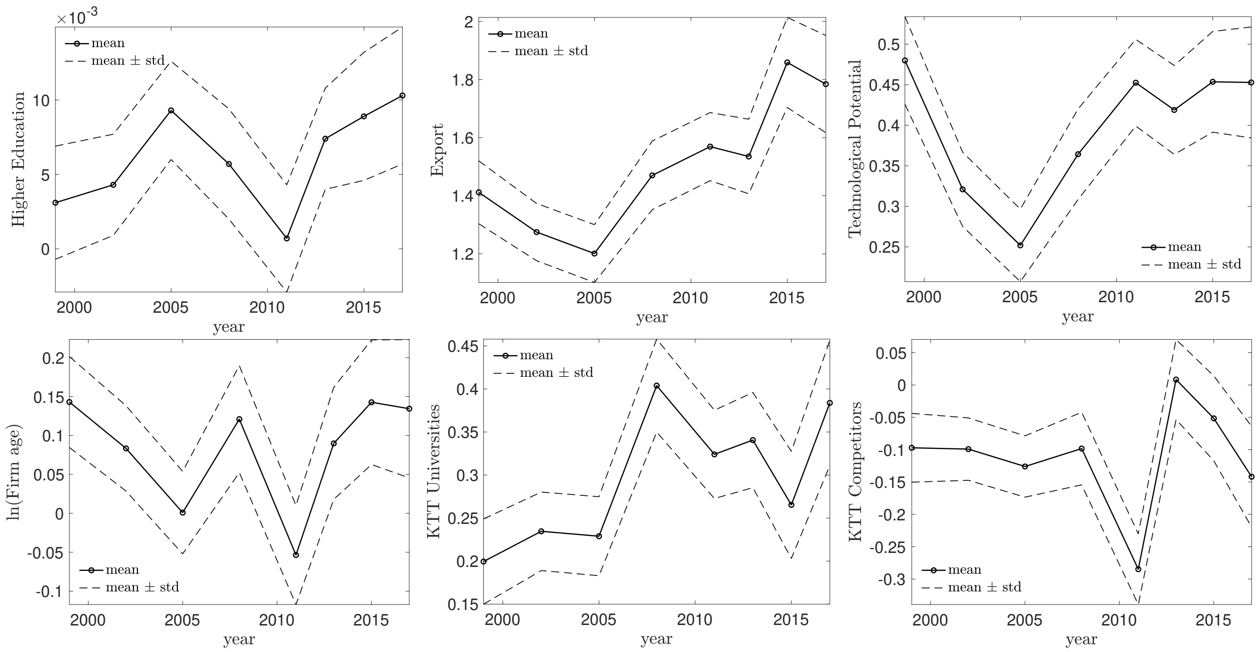


Figure 30: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics education, export (yes/no), technological potential, log firm age, KTT universities, and KTT competitors, across different years. The pooled estimation results can be found in column (3) of Table 18.

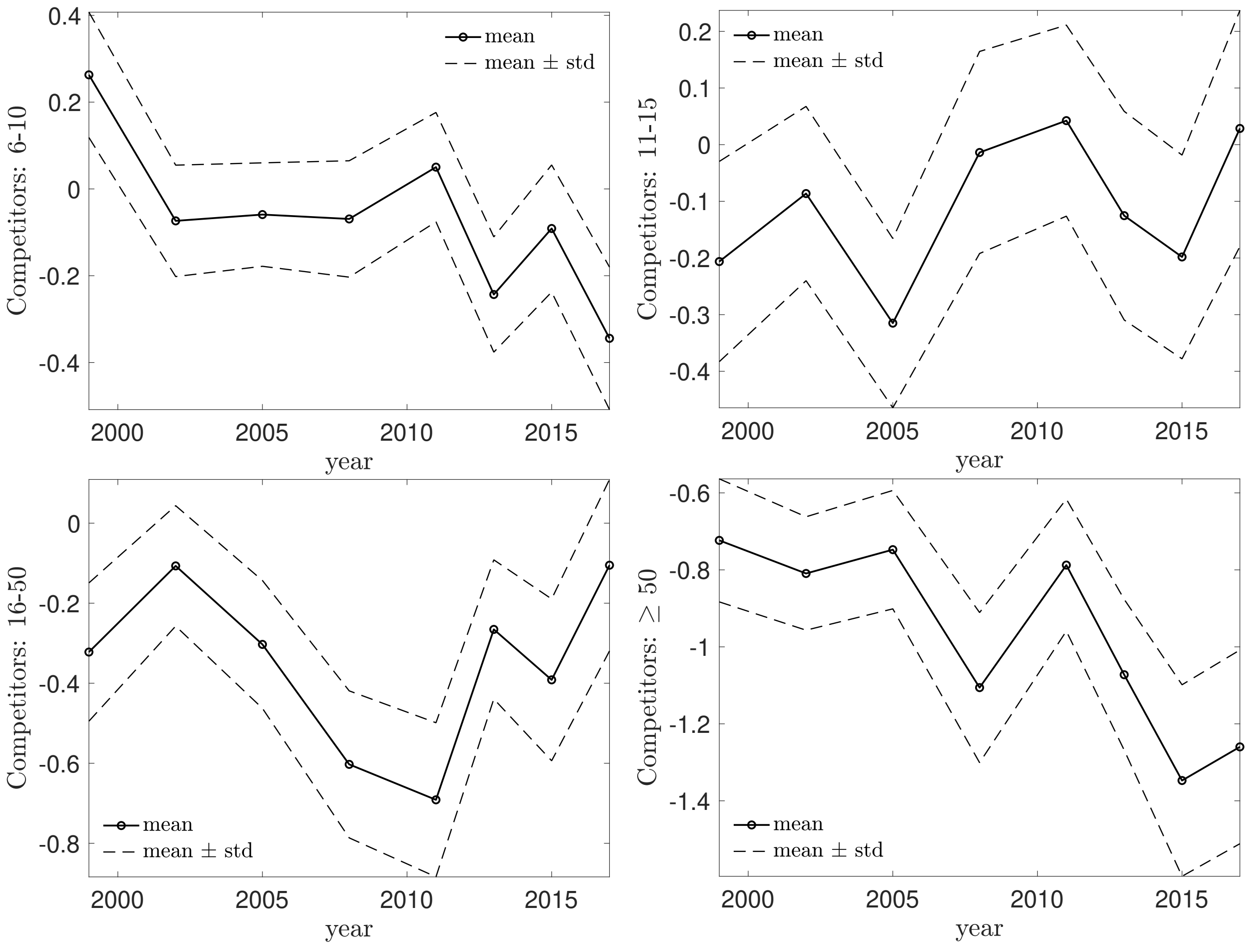


Figure 31: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics related to competition across different years. The pooled estimation results can be found in column (3) of Table 18.

Table 20: SMM estimation results across years. The pooled estimation results can be found in column (4) of Table 18.

		Simulated Method of Moments (SMM)								
		1996	1999	2002	2005	2008	2011	2013	2015	2017
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Imitation	(q)	0.7257***	0.7978***	0.7810***	0.8329***	0.7890***	0.7845***	0.7703***	0.8017***	0.8370***
		(0.1057)	(0.0790)	(0.0634)	(0.0757)	(0.0938)	(0.0631)	(0.0770)	(0.0594)	(0.0777)
Passive Imitation	(δ)	0.9096***	0.9116***	0.8943***	0.9092***	0.8951***	0.8928***	0.8740***	0.9142***	0.8866***
		(0.0631)	(0.0591)	(0.0627)	(0.0577)	(0.0630)	(0.0605)	(0.0806)	(0.0512)	(0.0655)
Innovation	$(\overline{p})$	0.7032***	0.6358***	0.6119***	0.6419***	0.5343***	0.6016***	0.5569***	0.5420***	0.5271***
		(0.0844)	(0.0925)	(0.0897)	(0.0808)	(0.0846)	(0.0918)	(0.0923)	(0.0742)	(0.0913)
Cost	$(\widetilde{\kappa})$	0.8497**	1.0493**	0.9719**	1.0307*	0.9750**	1.0821*	0.9757**	1.0632*	0.9823**
		(0.3943)	(0.5353)	(0.4733)	(0.5327)	(0.4887)	(0.5665)	(0.4815)	(0.5475)	(0.4838)

Notes: P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. * Statistically significant at 10% level.

Figure 32 shows the SMM parameter estimates across different years as reported in Table 20. The figure suggests a decreasing trend in the innovation success probability p, a weakly decreasing trend in the passive imitation success probability δ and a weakly increasing trend in the imitation success probability q while the innovation cost parameter $\widetilde{\kappa}$ remains largely stable over the years.

The robust declining trend in the in-house R&D success probabilities indicates that it becomes more difficult for firms to successfully make innovations. This declining trend also influences the innovation decision of the firms. As it becomes less likely for a firm to succeed with R&D, while the R&D costs remain the same, fewer firms decide to conduct R&D, in particular those that are further away from the frontier and that have more to gain from imitation.

6.5 Manufacturing and Services Sectors

Year-wise parameter estimates for the manufacturing and services sector from the Nonlinear Least Squares model are shown in Figure 33. As for the pooled sample, we observe that the in-house R&D success probability (p) is decreasing over the years, suggesting that it becomes harder to make innovations. Moreover, we find that in the services sector the estimated p is roughly half of what it is for the manufacturing sector. This indicates that it is more difficult to make innovations in the services sector than in the manufacturing sector. More detailed estimation results can be found in Appendix E.

6.6 Summary of the Estimation Results for Switzerland

To summarize our findings from our micro-regressions at the firm level based on survey data, we can extend the list of stylized facts in Section 4.2 further by observing that the decision of a firm to conduct R&D is affected by the following variables:

(4) Firm size has a positive effect on the R&D decision,

(5) Number of competitors has a negative effect on the R&D decision,

(6) Technological potential[4] has a positive effect on the R&D decision,

(7) Industry-University collaborations have a positive effect on the R&D decision.

(8) The in-house R&D success probability is decreasing over the years, suggesting that it becomes harder to make innovations.

[4] By technological potential, we mean the worldwide private and public technological knowledge that can be used to produce marketable innovations in your field of activity. This includes: scientific basic knowledge, knowledge about key technologies (e.g. nanotechnology, semiconductor technology, biotechnology, computer science, audiovisual techniques, etc.), which is suitable for implementation in innovations, technological and/or organizational knowledge specific to your field of activity.

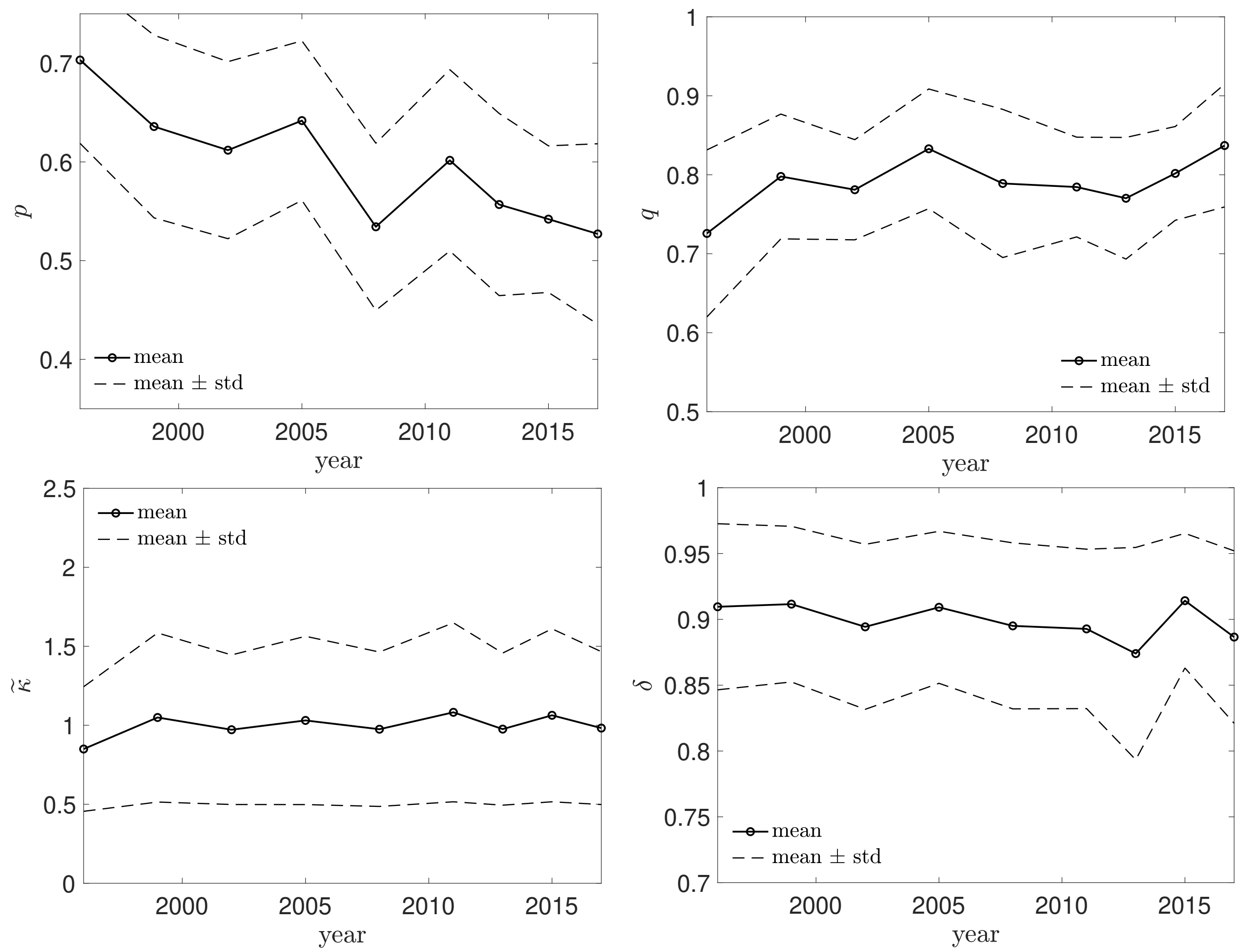


Figure 32: SMM parameter estimates across different years as reported in Table 20.

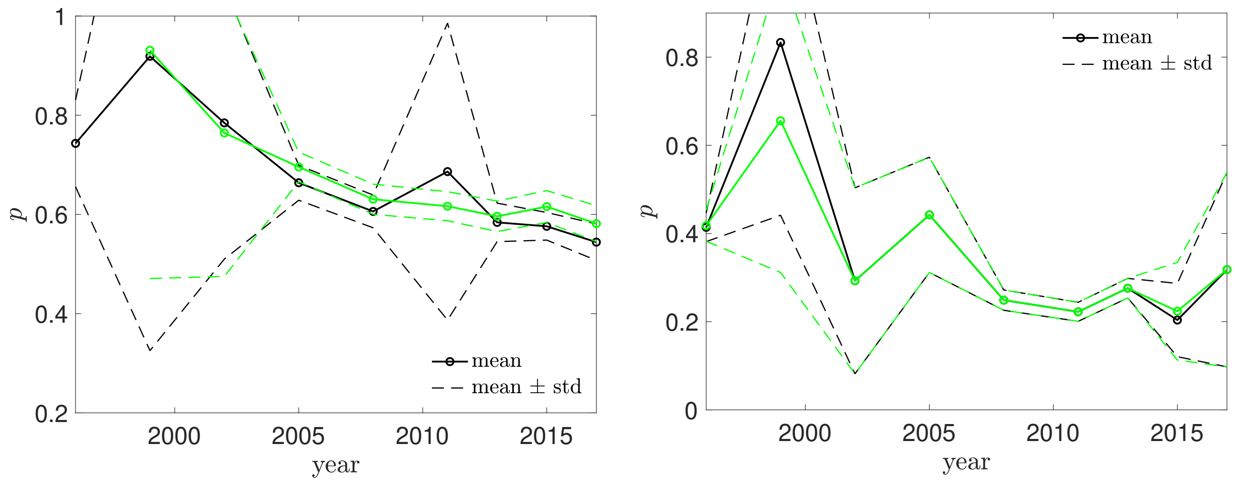


Figure 33: Nonlinear least squares model parameter estimates from Table 23 across different years for the manufacturing (left panel) and services (right panel) sector. The black curve (−) indicates the model with passive imitation (δ), while the green curve (−) indicates the model without passive imitation ($\delta = 0$).

6.7 Estimation Results for the Netherlands

In the following we provide a comparison of the estimation results we have obtained for Switzerland with the Netherlands. As already noted in Section 5.6.3, differently to Switzerland, the Dutch government actively promotes engaging in R&D activities through a favorable corporate tax system (in the form of an “innovation box” and a “patent box”). The estimation results pooled across the years 2000 to 2016 for the Netherlands with the innovation box and the patent box included as additional regressors can be seen in Table 21. We find that both, the innovation box and the patent box, have a positive and significant impact on the R&D decision of the firm. This shows that R&D policy incentives can affect the decision of firms to conduct R&D. The other estimates are similar to what we have observed for Switzerland.

Table 22 shows the Nonlinear Least Squares (NLS) parameter estimates across the years from 2000 to 2016 for the Netherlands.

Table 21: Estimation results for the Netherlands pooled across the years 2000 to 2016.

		LPM	NLS	
		w/o passive imitation	w/o firm characteristics	with firm characteristics
		(1)	(2)	(3)
Innovation	(p)	0.5958***	0.5622***	–
		(0.0160)	(0.0131)	
Imitation	(q)	0.4223***	0.6229***	0.8760***
		(0.0124)	(0.0365)	(0.0926)
Passive Imitation	(δ)	–	0.4906***	0.9660***
			(0.0592)	(0.0176)
Const.	(β_0)	–	–	-0.9049**
				(0.1224)
Higher Education	(β_1)	–	–	1.3316***
				(0.1028)
Export (yes/no)	(β_2)	–	–	0.8231***
				(0.0435)
Competitors	(β_3)	–	–	0.8614***
				(0.0451)
Patent Box	(β_4)	–	–	0.6808***
				(0.3117)
Innovation Box	(β_5)	–	–	1.7532***
				(0.1382)
ln(firm age)	(β_6)	–	–	-0.1093**
				(0.0252)
No. Obs.		40,781	40,781	12,486

Notes: Model (1) corresponds to the Linear Probability Model (LPM) discussed in Section B.5. Models (2) and (3) correspond to a Nonlinear Least Squares (NLS) estimation procedure with an innovation decision variable as in Equation (8) as a dependent variable. In all models we set $\theta = 0.2$ following König et al. (2020). Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses for models (1)–(3). P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. *

Table 22: Linear Probability Model (LPM) and Nonlinear Least Squares (NLS) estimation results across years for the Netherlands.

		2000 (1)	2002 (2)	2004 (3)	2006 (4)	2008 (5)	2010 (6)	2012 (7)	2014 (8)	2016 (9)
		Linear Probability Model (LPM) w/o Passive Imitation								
Innovation	(p)	0.4024	0.5401	0.4978	0.6036	0.4392	0.3963	0.5693	0.3674	0.4187
		(0.0354)	(0.0172)	(0.0420)	(0.1221)	(0.0626)	(0.0609)	(0.1750)	(0.1789)	(0.1458)
Innov. Cost	$(\widetilde{\kappa})$	0.0000	0.0512	0.0000	0.0440	0.0000	0.0000	0.0000	0.0000	0.0000
		(0.0000)	(0.0151)	(0.0000)	(0.1464)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Imitation	(q)	0.4314	0.3187	0.4258	0.4832	0.4821	0.4559	0.3077	0.6540	0.5262
		(0.0357)	(0.0244)	(0.0377)	(0.0762)	(0.0442)	(0.0514)	(0.1139)	(0.1131)	0.0985
		Nonlinear Least Squares (NLS) with Passive Imitation								
Innovation	(p)	0.4125	0.4477	0.4754	0.6416	0.4360	0.3914	0.5715	0.4220	0.4329
		(0.0178)	(0.0177)	(0.0332)	(0.1511)	(0.0711)	(0.0426)	(0.0477)	(0.0908)	(0.1231)
Innov. Cost	$(\widetilde{\kappa})$	0.0000	0.0000	0.0000	0.0561	0.0000	0.0000	0.0000	0.0000	0.0000
		(0.0000)	(0.0000)	(0.0000)	(0.1802)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Imitation	(q)	0.8547	0.7864	0.6510	0.2858	0.4145	0.7161	1.0336	1.0676	0.7272
		(0.0585)	(0.0542)	(0.1119)	(0.1305)	(0.1465)	(0.1386)	(0.1023)	(0.1357)	(0.2420)
Pass. Imit.	(δ)	0.7245	0.8387	0.5174	-1.4394	-0.2708	0.5427	0.8074	0.5638	0.3697
		(0.0735)	(0.0441)	(0.1639)	(1.6194)	(0.6470)	(0.1821)	(0.0883)	(0.1119)	(0.3103)
No. Obs.		4,178	5,445	4,627	3,761	4,242	2,408	1,827	1,715	1,926

Notes: P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. * Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses.

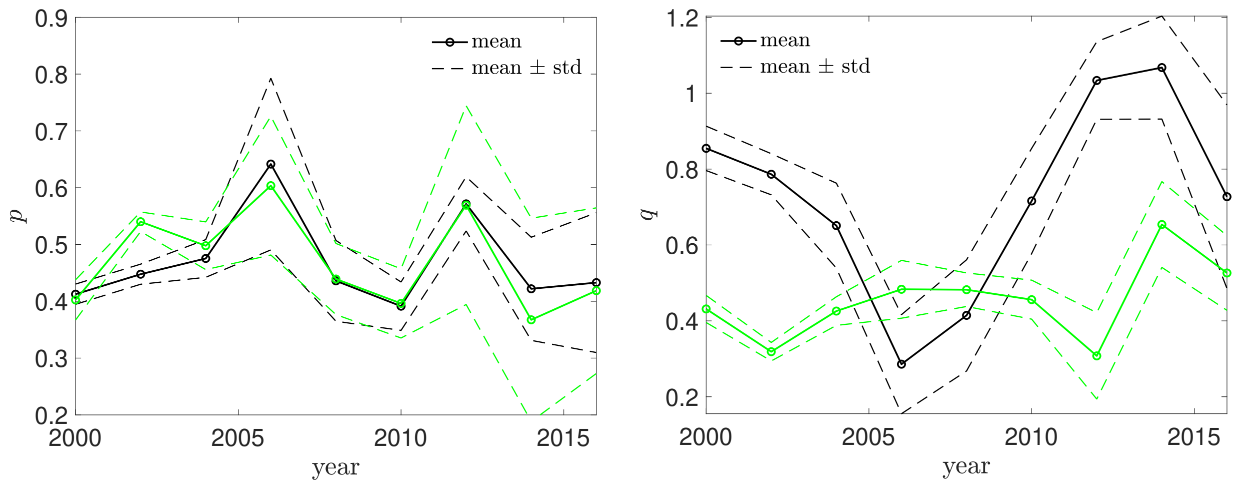


Figure 34: Nonlinear least squares model parameter estimates from Table 22 across different years for the in-house R&D success probability p (left panel) and the imitation success probability q (right panel) for the Netherlands. The black curve (–) indicates the model with passive imitation (δ), while the green curve (–) indicates the model without passive imitation ($\delta = 0$).

Corresponding year-wise parameter estimates for the Netherlands from the Nonlinear Least Squares model are shown in Figure 34. Similar to Switzerland, we observe that the in-house R&D success probability (p) is sightly decreasing (or stagnating) after 2005, suggesting that it becomes harder to make innovations. At the same time, after 2005, the imitation success probability is increasing. This seems to suggest that firms substitute innovation with imitation.

7 Counterfactual Analyses and R&D Policy Implications

In the following we analyze various counterfactuals to understand how the productivity growth rate and dispersion depend on the in-house R&D success probability, the imitation success probability and the R&D costs.

7.1 Sensitivity Analysis

Figure 35 shows changes in the productivity growth rate ν and the productivity variance σ^2 when changing the in-house R&D success probability p and the imitation success probability q, respectively. The changes are computed relative to benchmark scenario in which p and q are set to their estimated values. The remaining parameters (other than p and q) are set to their estimates in column (8) in Table 20. We find that the growth rate ν is monotonically increasing with both, increasing p and q. Setting $p = 1$ increases the growth rate by 14%, while setting $q = 1$ increases the growth rate by 6%. Hence, policies that aim at improving the in-house R&D success probability seem to be more effective than those that increase the imitation success probability (technology diffusion). However, this comes with an increase in productivity dispersion. Setting $p = 1$ increases the productivity variance by 80%, while increasing q reduces it.

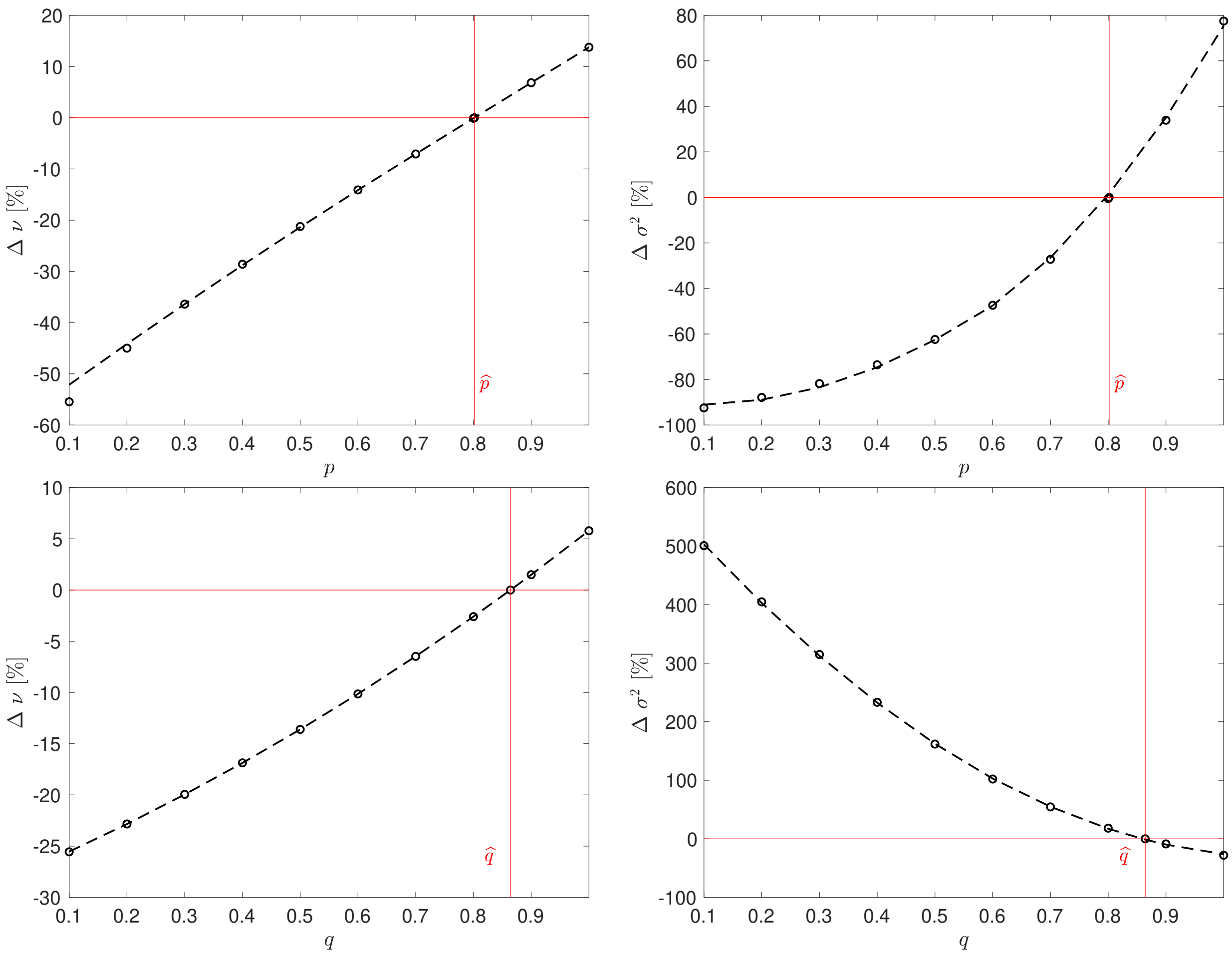


Figure 35: Changes in the productivity growth rate ν and the productivity variance σ^2 when changing the in-house R&D success probability p and the imitation success probability q, respectively, relative to their estimated values with the remaining parameters set to their estimates in column (9) in Table 20. Vertical lines indicate the estimated values $\widehat{p}$ and $\widehat{q}$ for p and q, respectively.

7.2 R&D Funding

We consider a subsidy, $s_i(t) \in [0,1]$, to the R&D costs of the firm. This subsidy can reflect, for example, R&D tax credits. More specifically, firm i's current (period t) profits are given by

$$\pi_i(t) = \psi A_i(t) - (1 - s_i(t))c_i(t), \tag{3}$$

where we consider two possible cases: In the first case, we consider a uniform subsidy identical for all firms, $s_i(t) = s$, while in the latter case, we consider a subsidy only for the firms below the threshold.

Figure 36 shows the productivity growth rate ν and the productivity variance σ^2 as a function of the R&D subsidy $s \in [0,1]$ relative to the case without a subsidy ($s = 0$). The parameters of the model are set to their estimated values in column (9) in Table 20. We find that the growth rate can be increased by up to 0.05% when introducing a full subsidy to the R&D costs of the firms. At the same time this subsidy reduces the productivity dispersion by 0.05%. The increase in the growth rate is due to more firms (namely those below the threshold) conducting R&D when their R&D costs are subsidized, which also leads to a reduction in the productivity dispersion as these firms increase their productivities and thus move closer to the productivity frontier.

Figure 37 shows the ratio of the cost of the R&D subsidy when all firms receive it (C_{uniform}), or when only the firms below the threshold receive it ($C_{\text{threshold}}$), as a function of the R&D subsidy $s \in [0,1]$. While the impact on the growth rate for both policies is the same, we observe that $C_{\text{threshold}}$ is several orders of magnitude smaller than C_{uniform}. However, the implementation of a threshold dependent subsidy requires information about the location of the threshold along the productivities of the firms.

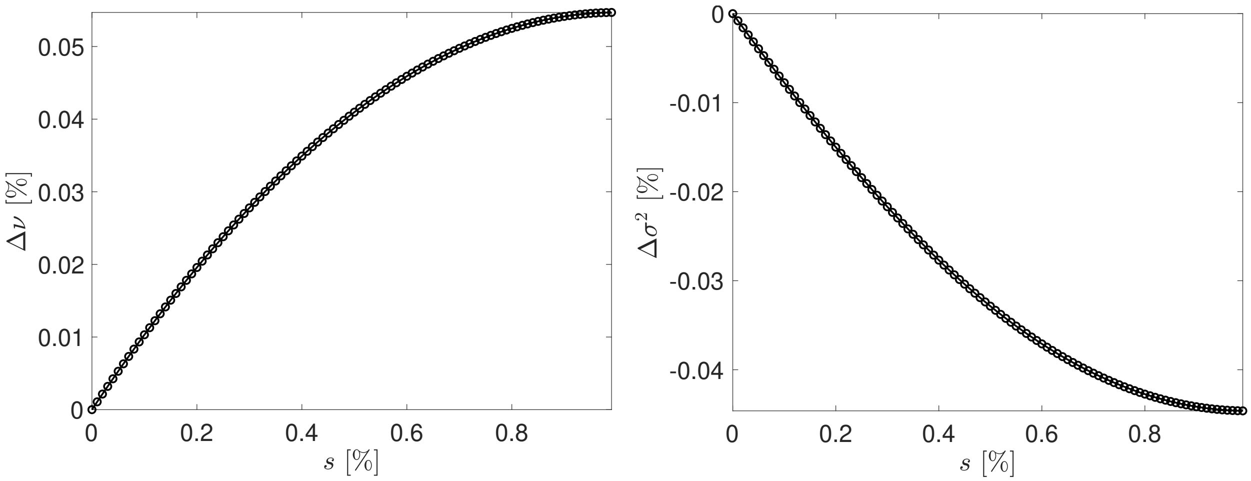


Figure 36: Changes in the productivity growth rate ν and the productivity variance σ^2 as a function of the R&D subsidy $s \in [0, 1]$ relative to the case without a subsidy ($s = 0$). The parameters of the model are set to their estimated values in column (9) in Table 20.

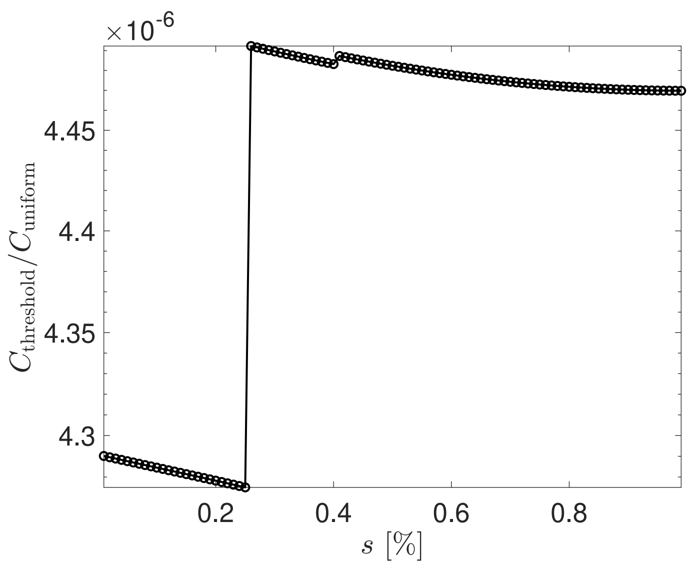


Figure 37: Ratio of the cost of the R&D subsidy when all firms receive it (C_{uniform}), or when only the firms below the threshold receive it ($C_{\text{threshold}}$), as a function of the R&D subsidy $s \in [0, 1]$.

8 Conclusion

We have analyzed the distribution of productivity and the innovation decision of firms in Switzerland across a panel of 20 years of observation. Our analysis suggests that productivity growth is stagnating and the dispersion of productivity is increasing. Our structural model estimates indicate that this is due to increasing difficulties of firms to make innovations, as we observe a decline in the in-house R&D success probabilities of the firms across the years. Moreover, we observe weakly declining technology diffusion among firms, as evidenced by a weakly declining trend in the imitation success probability.

We then use our calibrated model to investigate the impact of improving the in-house innovation success probabilities versus the imitation success probabilities. We find that policies that aim at improving the in-house R&D success probability seem to be more effective than those that increase the imitation success probability (technology diffusion). However, this comes with an increase in the productivity dispersion (inequality).

Finally, we investigate the effectiveness of a subsidy to firms' R&D costs. While this can be achieved relatively cost-effectively if the subsidy is provided to firms below the threshold only (and not to firms above the threshold that would have performed R&D also without the subsidy), the effect on improving the growth rate are rather moderate. Policies that improve the in-house R&D success probability seem to be more effective. Such policies could support firms to access international markets, help in increasing the share of high skilled workers or foster university-industry collaborations.

References

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, 4(1):37–74.

Aghion, P., Akcigit, U., and Howitt, P. (2014). What Do We Learn From Schumpeterian Growth Theory? *NBER Working Paper No. 18824*.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2019). A theory of falling growth and rising rents. Technical report, National Bureau of Economic Research.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2):701–728.

Akcigit, U., Grigsby, J., Nicholas, T., and Stantcheva, S. (2018). Taxation and innovation in the 20th century. Technical report, National Bureau of Economic Research.

Akcigit, U., Hanley, D., and Stantcheva, S. (2019). Optimal taxation and R&D policies. Technical report, National Bureau of Economic Research.

Akcigit, U. and Stantcheva, S. (2020). Taxation and innovation: What do we know? Technical report, National Bureau of Economic Research.

Akicigit, U. and Ates, S. (2019). Ten facts on declining business dynamism and lessons from endogenous growth theory. *NBER Working paper*, 25755.

Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business & Economic Statistics*, 14(3):353–366.

Andrews, D., Criscuolo, C., and Gal, P. N. (2015). Frontier firms, technology diffusion and public policy: Micro evidence from oecd countries. *OECD Productivity Working Papers*, (2).

Andrews, D., Criscuolo, C., and Gal, P. N. (2016). The best versus the rest: The global productivity slowdown, divergence across firms and the role of public policy. *OECD Publishing, No. 5*.

Arqué-Castells, P. and Mohnen, P. (2015). Sunk costs, extensive R&D subsidies and permanent inducement effects. *The Journal of Industrial Economics*, 63(3):458–494.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Bartelsman, E. J. and Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic Literature*, 38(3):569–594.

Bartelsman, E. J. and Wolf, Z. (2018). Measuring productivity dispersion. In *The Oxford Handbook of Productivity Analysis*.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, 33(3):163–84.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics: methods and applications*. Cambridge University Press.

Cavenaile, L., Celik, M. A., and Tian, X. (2019). Are markups too high? competition, strategic innovation, and industry dynamics. *Mimeo*.

Chen, Z., Liu, Z., Suárez Serrato, J. C., and Xu, D. Y. (2018). Notching R&D investment with corporate income tax cuts in China. *National Bureau of Economic Research, Working Paper No. 24749*.

Cleves, M., Gould, W., Gould, W. W., Gutierrez, R., and Marchenko, Y. (2008). *An introduction to survival analysis using Stata*. Stata Press.

Cohen, W. and Klepper, S. (1996). A reprise of size and R&D. *The Economic Journal*, 106(437):925–951.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Dorn, D., Katz, L. F., Patterson, C., Van Reenen, J., et al. (2017). Concentrating on the fall of the labor share. *American Economic Review*, 107(5):180–85.

Eeckhout, J. and Jovanovic, B. (2002). Knowledge spillovers and inequality. *American Economic Review*, 92(5):1290–1307.

Jones, B. F. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317.

König, M., Song, Z. M., Storesletten, K., and Zilibotti, F. (2020). From imitation to innovation where is all that Chinese R&D going? National Bureau of Economic Research (NBER) Working Paper No. 27404.

König, M. D., Lorenz, J., and Zilibotti, F. (2016). Innovation vs. imitation and the evolution of productivity distributions. *Theoretical Economics*, 11:1053–1102.

Lucking, B., Bloom, N., and Van Reenen, J. (2019). Have R&D spillovers declined in the 21st century? *Fiscal Studies*, 40(4):561–590.

Peters, B., Roberts, M. J., Vuong, V. A., and Fryges, H. (2017). Estimating dynamic R&D choice: an analysis of costs and long-run benefits. *The RAND Journal of Economics*, 48(2):409–437.

Rammer, C. and Schubert, T. (2016). Concentration on the few? R&D and innovation in German firms between 2001 and 2013. Technical report, Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis.

Rammer, C. and Schubert, T. (2018). Concentration on the few: mechanisms behind a falling share of innovative firms in Germany. *Research Policy*, 47(2):379–389.

Spescha, A. and Woerter, M. (2020). Starker Trend zur Konzentration von Forschungs- und Entwicklungsausgaben. *Die Volkswirtschaft*, 6:53–55.

Appendix

A Technical information about the Cox model

The attractive feature of the Cox model is that the baseline hazard h0(t) has no particular parameterization (Cleves et al., 2008). One does not even need to estimate it. It is possible to estimate it from the data though. Thus, the Cox model makes no assumption about the shape of the hazard rate over time, which could be increasing, decreasing, or take any other form. However, it is assumed to be the same for all firms (Cleves et al., 2008). The Cox model is called proportional because the hazard of a subject is multiplicatively proportional to the baseline hazard h0(t). The Cox model parameterizes how the covariates alter the baseline hazard function. More specifically, the Cox model assumes that the covariates multiplicatively shift the baseline hazard function h0(t). In the Cox model, the covariates measure how they either extend or shorten the risk of failure. For example, whether higher productivity either extends or shortens the risk of firms to exit from R&D.

Important issues in survival analysis are censoring and truncation. Censoring describes the instant that a firm does not exhibit an exit from R&D during analysis time. Right censoring means that the firm is still R&D active when the analysis ends; left censoring means that the firm has never been R&D active. Both kinds of censoring are easily handled by Stata's "stset" function. Interval censoring concerns inexact measurement of failure times, such as when subjects report data only in specific, regular intervals. In our case, the period where the firms fail or enter are arguably measured inexactly, somewhere between, for example, the years 2008 and 2011. However, because the intervals do not overlap, there is no problem for the Cox model in handling failure times; the Cox model is merely less efficient. More exact failure times, for instance exits recorded every year, would provide more information and thus allow for estimates that are more precise. Truncation, in contrast, describes the instant when we have no information on firms in certain periods. This is a problem in our case because of the non-response of firms. Left truncation, which means that firms join the analysis later on, is addressed through explicitly designating firms in the analysis as such. This is straightforward to implement in Stata. Interval truncation is a concern when there are gaps between the different responses to the survey. Fortunately, there are no cases of interval truncation in the data. Finally, right truncation is a concern when subjects are included in the sample because of their failure times. This is not an issue either in our case, because our sample is representative of the economy; it is not the case that certain firms are included purposefully in the sample because they either exit or enter R&D.

Important in survival analysis is that subjects who have equal values at time t also face the same risk (Cleves et al., 2008). For instance, in an analysis of the effects of smoking on cancer, the onset of the risk to develop cancer is in the year the person has started smoking. Which calendar year the person has started smoking is less relevant and is allowed to differ across different persons. The survival analysis starts for every person in the year he or she has started smoking, and the analysis time is the number of years the person has smoked (one might need to control for the age of the person though). In contrast, in our case the onset of risk for all firms has to be the fixed to the year 1996, irrespective of whether the firms enter the analysis already in 1996 or whether they enter it in later calendar years. This is important, because the risk for exiting R&D, proxied by unobserved covariates that we cannot measure, may have increased since the year 1996. Firms entering the analysis later on could thus face right from the outset a higher risk for exiting from R&D than those firms that have started out in the year 1996 already. By fixing the onset of risk to the calendar year 1996, we can hold constant the risk profile over the entire analysis timespan of 23 years from 1996-2017, meaning that each year is associated with a specific risk for the firms to exit R&D. Of course, each firm then only contributes to the measurement of the risk profile for the years it is part of the analysis. A firm that enters in, for example, 2005 only contributes to how the risk to exit from R&D has evolved from the year 2005 on.

In the following we discuss a structural model for the decision of firms to conduct R&D and the resulting productivity dynamics. We explain in detail the estimation algorithm, estimation results, a discussion of the structural model and the counterfactual policy simulation. In particular, the structural model is introduced in Section B, the estimation algorithms and estimation results are discussed in Section C. The policy implications that are based on the estimated model can be found in Section 7.

B A Structural Endogenous Growth Model

In the following we introduce a structural model that takes into account the endogenous evolution of the firms' productivities from their innovation and imitation decisions. Our theoretical model builds on König et al. (2016) and its extension in König et al. (2020).

B.1 Firms' Profits and Production

The final good, denoted by $Y(t)$, is produced by a representative firm using labor and a set of intermediate goods $x_i(t)$, $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. Its technology is represented by the following production function:

$$Y(t) = \frac{1}{\alpha} L^{1-\alpha} \sum_{i=1}^{N} A_i\,(t)^{1-\alpha}\, x_i(t)^{\alpha}, \quad \alpha \in (0,1),$$

where t denotes time, x_i is the intermediate good i, and A_i is the technology level of industry i. We normalize the labor force to unity, $L = 1$. The final good can be used for consumption, as an input to R&D, and also as an input to the production of intermediate goods. Its price is set to be the numeraire. The profit maximization program yields the following inverse demand function for intermediate goods:

$$p_i(t) = \left(\frac{A_i(t)}{x_i(t)}\right)^{1-\alpha}.$$

The profit earned by the incumbent in any intermediate sector i is then proportional to productivity,[5]

$$\pi_i(t) = (p_i(t) - 1)\, x_i(t) - c_i\,(t) = \psi A_i(t) - c_i\,(t)\,, \tag{4}$$

[5] Each intermediate good i is produced by a technology leader who has access to the best technology. By this best-practice technology the marginal cost of producing any intermediate input equals one unit of the final good. The leader is subject to the potential competition of a fringe of firms that can produce the same input albeit at a higher constant marginal cost, χ, where $1 < \chi \leq 1/\alpha$. Note that a higher value of χ indicates less competition. Bertrand competition implies that each technology leader monopolizes its market, sets the price equal to the unit cost of the fringe, $p_i(t) = \chi$, and sells the quantity $x_i(t) = \chi^{-\frac{1}{1-\alpha}} A_i(t)$. Namely, the equilibrium entails a limit price strategy and an inactive fringe as in Acemoglu et al. (2006).

where we have denoted by $\psi \equiv \frac{\chi-1}{\alpha}\chi^{-\frac{1}{1-\alpha}}$, with $\frac{d\psi}{d\chi} > 0$, and $c_i(t)$ is the cost of innovation given by (Chen et al., 2018; König et al., 2020)[6]

$$c_i(t) = \begin{cases} \kappa \overline{A}(t)^{\theta} A_i(t)^{1-\theta} & \text{if} \quad i \text{ innovates}, \\ \\ 0 & \text{if} \quad i \text{ imitates}. \end{cases} \tag{5}$$

Here, $\overline{A}(t)$ denotes the average productivity at time t (which can be related to the average wage rate in the economy, cf. König et al. (2020)) and $\kappa > 0$, $\theta \in [0, 1]$ are cost parameters. Note that we assume the R&D cost to be proportional to a geometric combination of A_i and $\overline{A}$, where the latter is in turn proportional to the average wage rate in the economy. If $\theta = 1$, the R&D costs are independent of TFP. In reality, one might expect R&D costs to vary across firms, being possibly higher for high-productivity firms that are closer to the technological frontier. For instance, more productive firms may be forced to divert managerial and labor resources that have a higher opportunity cost. Our specification captures in a flexible way this possibility. For our empirical analysis we will set $\theta = 0.2$ following König et al. (2020).

B.2 Innovation vs. Imitation

Productivity is measured along a quality ladder, $A_i \in \{\widetilde{A}, \widetilde{A}^2, \widetilde{A}^3, \ldots\}$. Firms can increase their productivity by a factor $\widetilde{A}$ along the ladder via two alternative channels: through costly in-house R&D (innovation) or through imitating other firms' technologies (diffusion).

Imitation A firm pursuing the imitation strategy is randomly matched with another firm in the empirical distribution. If the firm is matched with a more productive firm, its productivity increases by one notch with probability $q \in [0, 1]$ and remains constant with probability $1 - q$. If the firm is matched with a less productive firm, it retains its initial productivity. Because of random matching, the probability that an imitating firm with log-productivity $a = \log(A)$ moves up the productivity ladder equals $q \sum_{j=1}^{\infty} P_{a+j} = q(1 - F_a)$, where P_a denotes the log-productivity distribution (probability mass function; pmf) and $F_a = \sum_{j=0}^{a} P_j$ is the corresponding cumulative distribution function (cdf).

Innovation A firm can discover something genuinely new that is unrelated to the knowledge set of other firms. We label this process in-house innovation, and denote by $p_i \in [0, 1]$ the probability of success through in-house innovation. We assume p_i to be drawn from an i.i.d. uniform distribution with support $p_i \in [\underline{p}, \overline{p}]$, where $0 \leq \underline{p} < \overline{p} < 1$. The realization of p_i is observed at the beginning of each period t, before firms choose whether to innovate or imitate.

If innovation fails, the firm gets a second chance to improve its technology via (passive) imitation. However, in this case the probability of success is different from that of a firm actively pursuing imitation, being equal to $\delta q(1 - F_a)$. Thus, the total probability of success of a firm pursuing innovation is $p_i + (1 - p_i)\delta q(1 - F_a)$.

[6] Observe that any homothetic production function with Hicks neutral productivity admits this representation (cf. Chen et al., 2018).

B.3 Innovation Decision and Threshold

We assume that firms choose whether to innovate through in-house R&D or to imitate other firms based on a standard value-maximization objective. In our environment, this is equivalent to maximizing the expected profit in every period t. In turn, Equation (4) shows that the profit is linearly increasing in the technology level.

Let $\mathbb{E}_i^{\text{in}}[\pi_i(t+\Delta t)|\cdot]$ and $\mathbb{E}_i^{\text{im}}[\pi_i(t+\Delta t)|\cdot]$ denote the expected profit for a firm i, from choosing in-house R&D or imitation, respectively. The probability of success in innovating through in-house R&D is given by $p_i(t) \in [\underline{p}, \overline{p}]$, with $0 \leq \underline{p} \leq \overline{p} \leq 1$. The probabilities $p_i(t)$ are i.i.d. and realized at the beginning of each period t.

Firm i chooses innovation whenever, conditional on its current productivity $A_i(t)$ and the state of $p_i(t)$ we have that

$$\mathbb{E}_i^{\text{in}}[\pi_i(t+\Delta t)|\, A_i(t), p_i(t), P(t)] = \mathbb{E}_i^{\text{in}}\left[\psi A_i(t+\Delta t) - \kappa \overline{A}(t)^\theta A_i(t)^{1-\theta} \middle| A_i(t), p_i(t), P(t)\right] > \mathbb{E}_i^{\text{im}}\left[\pi_i(t+\Delta t)|\, A_i(t), P(t)\right] = \mathbb{E}_i^{\text{im}}\left[\psi A_i(t+\Delta t)|\, A_i(t), P(t)\right]. \quad (6)$$

The expected profit from imitation is given by

$$\mathbb{E}_i^{\text{im}}\left[\pi_i(t+\Delta t)|\, A_i(t), P(t)\right] = q\left(1 - F_{a_i(t)}(t)\right)\psi A_i(t)\widetilde{A} + \left(1 - q\left(1 - F_{a_i(t)}(t)\right)\right)\psi A_i(t),$$

while the expected profit from innovation is given by

$$\begin{aligned}
&\mathbb{E}_i^{\text{in}}[\pi_i(t+\Delta t)|\, A_i(t), p_i(t), P(t)] = p_i(t)\psi A_i(t)\widetilde{A} \\
&+ (1-p_i(t))\left\{\delta\left[q\left(1-F_{a_i(t)}(t)\right)\psi A_i(t)\widetilde{A} + \left(1 - q\left(1-F_{a_i(t)}(t)\right)\right)\psi A_i(t)\right] + (1-\delta)\psi A_i(t)\right\} \\
&- \kappa\overline{A}(t)^\theta A_i(t)^{1-\theta}.
\end{aligned} \quad (7)$$

In terms of log-productivities $a_i(t) \equiv \ln A_i(t)$, $\overline{a}(t) \equiv \ln \overline{A}(t)$ and $\ln \widetilde{A} \equiv \widetilde{a}$,[7] we can write

$$\begin{aligned}
&\mathbb{E}_i^{\text{in}}[\pi_i(t+\Delta t)|\, a_i(t), p_i(t), P(t)] = p_i(t)\psi e^{(a_i(t)+\widetilde{a})} \\
&+ (1-p_i(t))\left\{\delta\left[q\left(1-F_{a_i(t)}(t)\right)\psi e^{a_i(t)+\widetilde{a}} + \left(1 - q\left(1-F_{a_i(t)}(t)\right)\right)\psi e^{a_i(t)}\right] + (1-\delta)\psi e^{a_i(t)}\right\} \\
&- \kappa e^{\theta\overline{a}(t)} e^{(1-\theta)a_i(t)},
\end{aligned}$$

and

$$\mathbb{E}_i^{\text{im}}\left[\pi_i(t+\Delta t)|\, a_i(t), P(t)\right] = \psi e^{a_i(t)}\left(1 + q\left(1-F_{a_i(t)}(t)\right)\left(e^{\widetilde{a}} - 1\right)\right).$$

In the following we denote by $\overline{\pi}_i^{\text{im}}(a_i(t), P(t)) \equiv \mathbb{E}_i^{\text{im}}[\pi_i(t+\Delta t)|\, a_i(t), P(t)]$ and $\overline{\pi}_i^{\text{in}}(a_i(t), p_i(t), P(t)) \equiv \mathbb{E}_i^{\text{in}}[\pi_i(t+\Delta t)|\, a_i(t), p_i(t), P(t)]$. The indicator function for whether firm i conducts in-house R&D or pursues imitation is given by $\chi^{\text{im}}(a,p,P) = \mathbb{1}_{\left\{\overline{\pi}_i^{\text{im}}(a,P) > \overline{\pi}_i^{\text{in}}(a,p,P)\right\}}$. We also define the indicator

[7] Observe that $A_i \in \{\widetilde{A}, \widetilde{A}^2, \widetilde{A}^3, \ldots\} = \{e^{\widetilde{a}}, e^{2\widetilde{a}}, e^{3\widetilde{a}}, \ldots\}$.

function for innovation as $\chi^{\text{in}}(a,p,P) \equiv 1 - \chi^{\text{im}}(a,p,P)$. Further, denoting by

$$\widetilde{\kappa} = \frac{\kappa}{\psi(e^{\widetilde{a}} - 1)}$$

we can write

$$\chi^{\text{im}}(a,p,P) = 1 - \chi^{\text{in}}(a,p,P) = \begin{cases} 1 & \text{if } p < \frac{(1-\delta)q(1-F_a)+\widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q(1-F_a)}, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

B.4 Innovation Decision and Comparative Statics

To derive some further intuition for the innovation decision we consider in the following the special case of $\delta = 0$. Then Equation (8) reduces to

$$\chi^{\text{im}}(a,p,P) = 1 - \chi^{\text{in}}(a,p,P) = \begin{cases} 1 & \text{if } p < q(1-F_a) + \widetilde{\kappa}e^{\theta(\overline{a}-a)}, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The R&D decision in Equation (9) is consistent with the empirical observations (4) - (7) in Section 4.2. Consider a firm with in-house R&D success probability p_i and denote by

$$D_i = p_i - \widetilde{\kappa}e^{\theta(\overline{a}-a_i)} - q(1 - F_{a_i}). \tag{10}$$

Then the firm conducts innovation ($\chi^{\text{in}}(a,p,P) = 1$) if $D_i > 0$ and does imitation ($\chi^{\text{im}}(a,p,P) = 1$) if $D_i \leq 0$. From a comparative statics analysis we find that

$$\begin{aligned} \frac{\partial D_i}{\partial a_i} &= \theta\widetilde{\kappa}e^{\theta(\overline{a}-a_i)} + qf_{a_i} > 0, && \text{(size)}, \\ \frac{\partial D_i}{\partial \widetilde{\kappa}} &= -e^{\theta(\overline{a}-a_i)} < 0, && \text{(competition)}, \\ \frac{\partial D_i}{\partial p_i} &= 1 > 0, && \text{(tech. potential / univ. collab.)}. \end{aligned}$$

In particular, firms with a higher log-productivity, a_i, tend to be larger - see the profit function in Equation (4)[8] - and according to Equation (10) have a higher probability to do R&D ($\frac{\partial D_i}{\partial a_i} > 0$). More intense competition - empirical observation (5) in Section 6.6 - can be captured by a higher value of the competition parameter $1/\psi$ and according to Equation (10) this leads to a lower probability to do R&D ($\frac{\partial D_i}{\partial \widetilde{\kappa}} < 0$ where $\widetilde{\kappa} = \kappa/(\psi(e^{\widetilde{a}} - 1))$ is increasing with $1/\psi$). Note that the term $e^{\theta(\overline{a}-a_i)}$ is smaller the further the firm's log-productivity is above the average log-productivity $\overline{a}$, and hence diminishing the competition effect through $\widetilde{\kappa}$. Moreover, a higher value of the in-house R&D success probability p_i, due to a higher technological potential of the firm or university collaborations, leads to a higher probability to conduct R&D ($\frac{\partial D_i}{\partial p_i} > 0$), consistent with empirical observations (6) and (7) in Section 6.6. Finally, the term $q(1 - F_{a_i})$ reduces the likelihood of the firm conducting R&D, which is increasing with the imitation success probability q and is higher for firms lagging further behind in their log-productivity due to higher values of $1 - F_{a_i}$.

[8] Firm size x_i is proportional to productivity, a_i, as $x_i = \chi^{-\frac{1}{1-\alpha}} exp(a_i)$. See also Footnote 5.

B.5 Linear Probability Model Approximation

We can estimate Equation (10) directly by specifying the in-house R&D success probability as

$$p_i = \Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta} + \varepsilon_i\right), \tag{11}$$

where $\Lambda : \mathbb{R} \to [0,1]$, $x \mapsto 1/(1+e^{-x})$, is the logistic function, $\mathbf{y}_i$ are observable covariates of firm i that affect the firm's ability to successfully conduct R&D (such as technological potential and university collaborations) and let ε_i be a small zero-mean random term. In leading orders of ε_i, we can write $p_i = \Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta} + \varepsilon_i\right) \approx \Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta}\right) + \Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta}\right) \Lambda\left(-\mathbf{y}_i^\top \boldsymbol{\beta}\right) \varepsilon_i$, and hence Equation (10) can be written as

$$D_i \approx \underbrace{\Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta}\right)}_{\substack{\text{innovation}\\ \text{potential}}} \underbrace{-\widetilde{\kappa} e^{\theta(\overline{a}-a_i)}}_{\substack{\text{competition}\\ \text{effect}}} \underbrace{-q(1-F_{a_i})}_{\substack{\text{imitation}\\ \text{potential}}} + \widetilde{\varepsilon}_i, \tag{12}$$

where we have denoted by $\widetilde{\varepsilon}_i = \Lambda\left(\mathbf{y}_i^\top \boldsymbol{\beta}\right) \Lambda\left(-\mathbf{y}_i^\top \boldsymbol{\beta}\right) \varepsilon_i$. If we assume that the $\widetilde{\varepsilon}_i$ are identically and independently zero-mean normally distributed and D_i is replaced with the observed innovation decisions (χ^{in}), then Equation (12) represents a Linear Probability Model (LPM) that can be estimated with an Ordinary Least Squares (OLS) method. In a similar way we can estimate the decision variable of Equation (8) using a Nonlinear Least Squares (NLS) estimation algorithm. In both estimation methods we can correct for heteroscedastic standard errors.

These models are indicative for measuring some basic correlations but might suffer from an endogeneity bias due to productivity (and other variables) being affected by the innovation decision (and vice versa). In the following sections we introduce a structural model that takes into account the endogenous evolution of the firms' productivities from their innovation and imitation decisions.

B.6 Law of Motion of the Productivity Distribution

We consider an environment in which firms decide to conducting in-house R&D or to imitate other firms by maximizing the expected profit in every period t (see Section B.3). These decisions determine how the distribution of productivity evolves over time t. The following proposition provides a complete characterization of the evolution of the productivity distribution with heterogeneous firms in terms of their in-house R&D success probabilities, $p_i(t) \in [\underline{p}, \overline{p}]$.

Proposition 1. *The evolution of the* log*-productivity distribution, $P_a(t)$, $a \in \mathcal{A}$, is given by the following system of integro-differential equations*

$$\begin{aligned}\frac{\partial P_a(t)}{\partial t} = \int_{[\underline{p},\overline{p}]} g(dp) \Big[&\left(\chi^{im}(a-1,p,P) + \delta(1-p)\chi^{in}(a-1,p,P)\right) q(1-F_{a-1}(t))P_{a-1}(t) \\ &- \left(\chi^{im}(a,p,P) + \delta(1-p)\chi^{in}(a,p,P)\right) q(1-F_a(t))P_a(t) \\ &+ \chi^{in}(a-1,p,P)pP_{a-1}(t) - \chi^{in}(a,p,P)pP_a(t)\Big],\end{aligned} \tag{13}$$

where $g : [\underline{p}, \overline{p}] \to [0,1]$ *is the density function of a random variable over the interval* $[\underline{p}, \overline{p}]$ *and*

$$\chi^{im}(a, p, P) = 1 - \chi^{in}(a, p, P) = \begin{cases} 1 & \text{if } p < \frac{(1-\delta)q(1-F_a)+\widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q(1-F_a)}, \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

with $F_a = \sum_{a'=1}^{a} P_{a'} = \sum_{a'=1}^{a} P_{a'}$, *and the average* log*-productivity given by* $\overline{a} = \sum_{a=1}^{\infty} F_a$.

The proof of Proposition 1 can be found in Appendix F.[9] We can compute the evolution of the productivity distribution $P_a(t)$ by numerically solving the system of ordinary differential equations provided in Equation (17) for a given initial condition $P_a(0)$. This will be important for the estimation of the model discussed in Section C.

B.7 Uniformly Distributed R&D Success Probabilities

For uniformly distributed in-house R&D success probabilities we can further simplify Equation (13) in Section B.6 as follows:

Lemma 1. *Assuming a uniform distribution of the in-house R&D success probabilities with support* $[0, \overline{p}]$ *allows us to write Equation (13) as follows*

$$\begin{aligned}
&\frac{\partial P_a(t)}{\partial t} = \\
&\frac{1}{\overline{p}} \Bigg[q(1-F_{a-1}(t))P_{a-1}(t) \left(\min\{C(a-1,P), \overline{p}\} + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a-1,P)(2-C(a-1,P)) \right) \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} \right) \\
&-q(1-F_a(t))P_a(t) \left(\min\{C(a,P), \overline{p}\} + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a,P)(2-C(a,P)) \right) \mathbb{1}_{\{C(a,P)<\overline{p}\}} \right) \\
&-\frac{1}{2} P_a(t) \left(\overline{p}^2 - C(a,P)^2 \right) \mathbb{1}_{\{C(a,P)<\overline{p}\}} + \frac{1}{2} P_{a-1}(t) \left(\overline{p}^2 - C(a-1,P)^2 \right) \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} \Bigg],
\end{aligned} \tag{15}$$

where we have denoted by

$$C(a, P) \equiv \frac{(1-\delta)q(1-F_a) + \widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q(1-F_a)}. \tag{16}$$

Note that $C(a, P)$ is non-negative and decreasing in a. Let the threshold productivity be defined as $a^* = \{\min a \in \mathcal{A} : C(a, P) < \overline{p}\} = \{\max a \in \mathcal{A} : C(a, P) \geq \overline{p}\}$. Then for all $a \leq a^*$ we have that $\min\{C(a-1, P), \overline{p}\} = \min\{C(a, P), \overline{p}\} = \overline{p}$ and $\mathbb{1}_{\{C(a,P)<\overline{p}\}} = \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} = 0$ so that we can write Equation (15) for all $a \leq a^*$ as follows

$$\frac{\partial P_a(t)}{\partial t} = q(1-F_{a-1}(t))P_{a-1}(t) - q(1-F_a(t))P_a(t).$$

Then for $a = a^* + 1$ we have that $\min\{C(a-1, P), \overline{p}\} = \overline{p}$ but $\min\{C(a, P), \overline{p}\} = C(a, P)$ and

[9] A more detailed analysis of Equation (13) can be found in Appendix B.7.

$\mathbb{1}_{\{C(a,P)<\overline{p}\}} = 1$ but $\mathbb{1}_{\{C(a-1,P)<\overline{p}\}} = 0$ so that for $a = a^* + 1$ we can write Equation (15) as follows

$$\frac{\partial P_a(t)}{\partial t} = \frac{1}{\overline{p}} \left[q(1-F_{a-1}(t))P_{a-1}(t)\overline{p} - q(1-F_a(t))P_a(t) \left(C(a,P) + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a,P)(2-C(a,P)) \right) \right) \right.$$
$$\left. - \frac{1}{2} P_a(t) \left(\overline{p}^2 - C(a,P)^2 \right) \right],$$

For all $a > a^* + 1$ we have that $\min\{C(a-1,P), \overline{p}\} = C(a-1,P)$ and $\min\{C(a,P), \overline{p}\} = C(a,P)$ and $\mathbb{1}_{\{C(a,P)<\overline{p}\}} = \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} = 1$ so that Equation (15) for all $a > a^* + 1$ can be written as

$$\frac{\partial P_a(t)}{\partial t} = \frac{1}{\overline{p}} \left[q(1-F_{a-1}(t))P_{a-1}(t) \left(C(a-1,P) + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a-1,P)(2-C(a-1,P)) \right) \right) \right.$$
$$- q(1-F_a(t))P_a(t) \left(C(a,P) + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a,P)(2-C(a,P)) \right) \right)$$
$$\left. - \frac{1}{2} P_a(t) \left(\overline{p}^2 - C(a,P)^2 \right) + \frac{1}{2} P_{a-1}(t) \left(\overline{p}^2 - C(a-1,P)^2 \right) \right]. \tag{17}$$

We can compute the evolution of the productivity distribution $P_a(t)$ by numerically solving the system of ordinary differential equations provided in Equation (17) for a given initial condition $P_a(0)$.

B.8 Productivity Distribution and Comparative Statics

In Section 4.2 we have documented two empirical observations regarding the productivity distribution among firms in Switzerland and the fraction of firms conduction R&D. First, we observe an increasing threshold productivity for firms to conduct R&D, and second, an increasing dispersion of productivity across firms.

Increasing Threshold A shift to the right of the threshold $a^*(t) = \{\max a \in \mathcal{A} : \chi^{\text{im}}(a) = 1\}$ as illustrated in Figure B.1 results in a reduction in the number of R&D active firms. The reasons for this shift could be, for example, due to higher innovation costs. We explore this possibility in the estimation Section 6.4.

Increasing Dispersion An increasing dispersion of the productivity distribution $P_a(t)$ as illustrated in Figure B.2 could be triggered by, for example, higher productivity gains from innovation. We explore this possibility in the estimation Section 6.4.

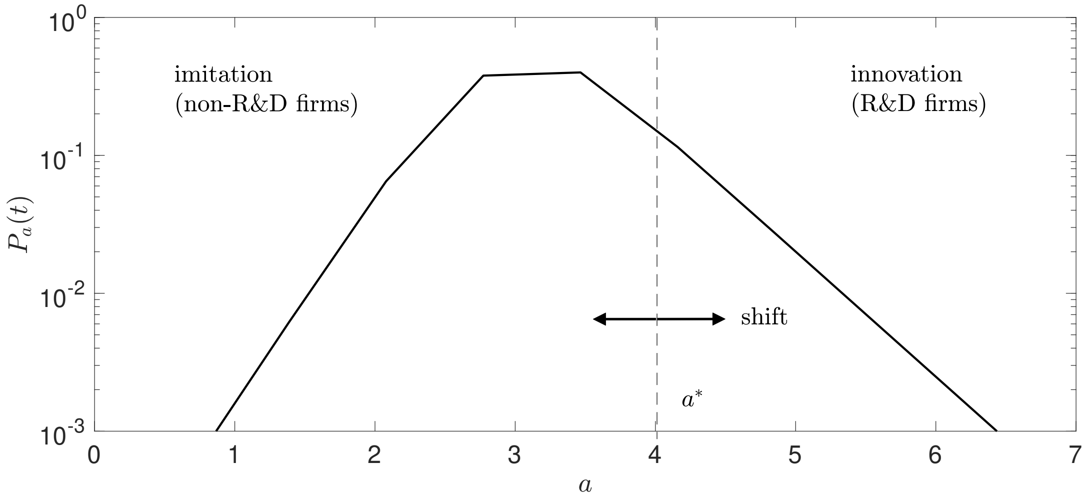


Figure B.1: An illustration of the threshold, $a^*(t)$.

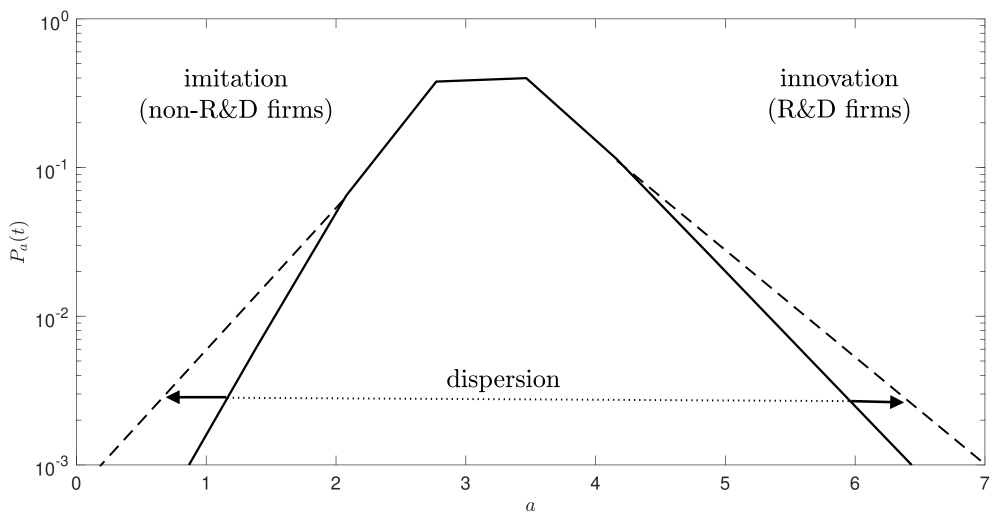


Figure B.2: An illustration of the dispersion of the productivity distribution, $P_a(t)$.

C Structural Estimation: Productivity distribution and R&D decision

The simple regression models seen in Section B.5 are indicative for measuring some basic correlations related to the R&D decision of firms. But this might suffer from an endogeneity bias due to productivity (and other variables) being affected by the innovation decision (and vice versa).

We consider a set of moments constructed from the log-productivity distribution, P_a, and the innovation decision indicator variable, χ^{in}. We then search for the parameters, θ, that minimize the distance between the targeted empirical moments and the stationary distribution of the model. Note that the tractability of the model is crucial for our approach. Our Simulated Method of Moments (SMM) (see e.g. Cameron and Trivedi, 2005) approach requires simulating the model many times, and calculating the distance from the targeted moments. We minimize the weighted sum of the distance between the empirical and simulated moments:

$$\hat{\theta} = \arg\min_{\theta} h(\theta)' W^{-1} h(\theta),$$

where $h_m(\theta) = g_m(\theta) - \frac{1}{K}\sum_k^K g_{m,k}$ and $\frac{1}{K}\sum_k^K g_{m,k}$ is the moment averaged across K samples. Denote Ω the variance-covariance matrix of the bootstrapped moments. Under the null hypothesis, Ω is proportional to the variance-covariance matrix of the simulated moments. We use the identity matrix as the benchmark weighting matrix to avoid the potential small-sample bias (see, e.g., Altonji and Segal (1996)). We also obtain results from using the optimal weighting matrix as a robustness check. The difference between the true and estimated parameter follows asymptotically a normal distribution with mean zero and the variance-covariance matrix of V, where $V = \left(DW^{-1}D'\right)^{-1}$ and $D = \frac{\partial h(\theta)}{\partial \theta}|_{\theta=\hat{\theta}}$. The variance of the estimated parameters are on the diagonal of V.

The first set of moments derives from the productivity distribution P_a. Figure 5 shows the empirical productivity distribution across firms in Switzerland that we target where we measure productivity as value added per employee. The goodness of fit of the model with the empirical distribution for the pooled sample across years can be found in Appendix D.

The second set of moments is constructed from the fraction of firms conducting R&D for a given log-productivity level a and the log-productivity distribution P. Following the innovation decision in Equation (14) and assuming uniform draws of p in the interval $[\underline{p}, \overline{p}]$ (see also Appendix B.7) this is given by

$$H_a(P) \equiv \int_{[\underline{p},\overline{p}]} \chi^{\text{in}}(a,p,P)dp = \int_{[\underline{p},\overline{p}]} \mathbb{1}_{\{p>C(a,P)\}}dp = (\overline{p} - \max\{C(a,P), \underline{p}\})\mathbb{1}_{\{C(a,P)<\overline{p}\}},$$

where $C(a,P)$ is defined in Equation (16) in Lemma 1. The goodness of fit of the model for the pooled sample across years can be found in Appendix D. In particular, a comparison of the empirical log-productivity innovation profile H pooled across the years 1997 to 2017 for Switzerland and the prediction by the model can be seen in the bottom panel in Figure D.1. The parameter estimates correspond to column (3) in Table 18.

D Goodness-of-Fit

Figure D.1 shows a comparison of the empirical log-productivity distribution P_a and the empirical log-productivity innovation profile H_a with the prediction by the model. The parameter estimates correspond to column (4) in Table 18. The data are pooled across the years 1996 to 2017 for Switzerland.

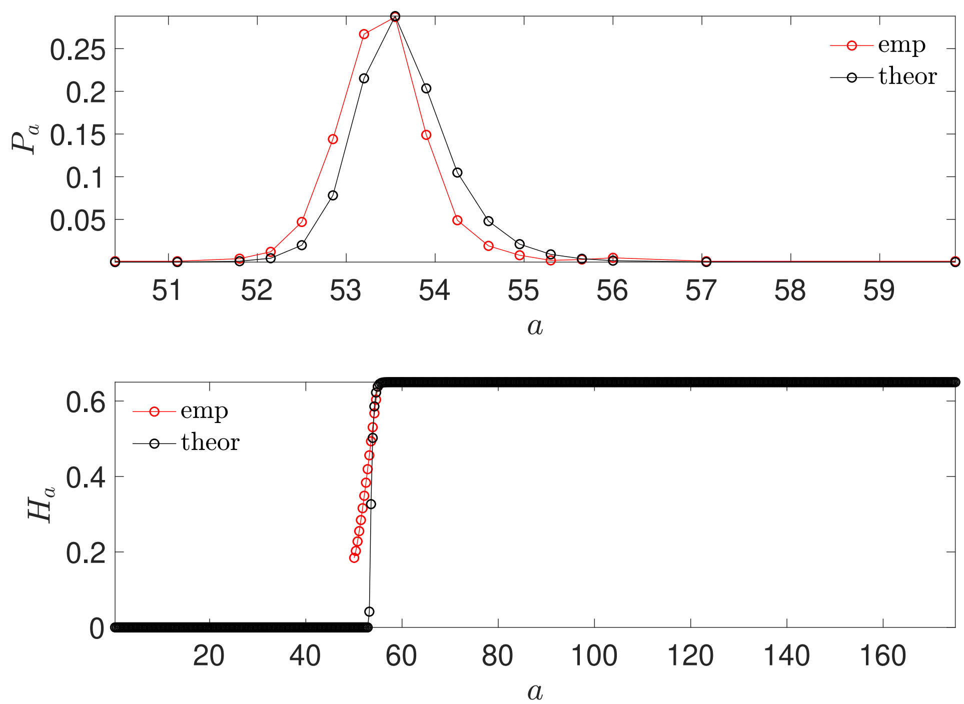


Figure D.1: (Top panel) Comparison of the empirical log-productivity distribution P_a and the prediction by the model. (Bottom panel) Comparison of the empirical log-productivity innovation profile H_a and the prediction by the model. The parameter estimates correspond to column (3) in Table 18. The data are pooled across the years 1996 to 2017 for Switzerland.

E Manufacturing and Services Sectors

In this section we perform the econometric analysis for the linear probability model restricted to firms in the manufacturing and the services sectors only. We find, however, that the main estimates and time trends remain roughly the same as the ones obtained for the full sample including all firms across all sectors.

E.1 Manufacturing Sector

Table 23 shows the Nonlinear Least Squares (NLS) parameter estimates across the years from 1999 to 2017 for the manufacturing sector.

Table 23: Linear Probability Model (LPM) and Nonlinear Least Squares (NLS) estimation results across years for the manufacturing sector.

		1996	1999	2002	2005	2008	2011	2013	2015	2017
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Linear Probability Model (LPM) w/o Passive Imitation								
Innovation	(p)	1.4868	0.9309	0.7643	0.6958	0.6307	0.6166	0.5962	0.6159	0.5815
		(0.2829)	(0.4601)	(0.2892)	(0.0297)	(0.0304)	(0.0293)	(0.0307)	(0.0321)	(0.0362)
Innov. Cost	$(\widetilde{\kappa})$	0.8236	0.2977	0.0760	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		(0.4087)	(1.8527)	(4.4861)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Imitation	(q)	-0.1106	0.0178	0.1844	0.2806	0.2000	0.1600	0.1850	0.2538	0.2657
		(0.1532)	(0.2181)	(0.1487)	(0.0501)	(0.0576)	(0.0546)	(0.0601)	(0.0631)	(0.0711)
		Nonlinear Least Squares (NLS) with Passive Imitation								
Innovation	(p)	0.7433	0.9188	0.7846	0.6639	0.6061	0.6864	0.5840	0.5761	0.5442
		(0.0875)	(0.5934)	(0.2746)	(0.0352)	(0.0330)	(0.2994)	(0.0388)	(0.0278)	(0.0367)
Innov. Cost	$(\widetilde{\kappa})$	0.0072	0.2851	0.1126	0.0000	0.0000	0.0978	0.0000	0.0000	0.0000
		(11.4608)	(2.3757)	(2.7240)	(0.0000)	(0.0000)	(3.4587)	(0.0000)	(0.0000)	(0.0000)
Imitation	(q)	0.9398	0.0551	0.3326	0.6847	0.7466	0.4913	0.5310	0.9293	0.8190
		(0.1687)	(1.1280)	(0.5672)	(0.2282)	(0.2876)	(0.6319)	(0.6103)	(0.0796)	(0.1957)
Pass. Imit.	(δ)	0.9743	0.9466	0.7133	0.8047	0.8912	0.9873	0.7811	0.9325	0.8866
		(0.0819)	(8.8994)	(0.9345)	(0.1861)	(0.1357)	(0.5621)	(0.4865)	(0.0431)	(0.1116)

Notes: P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. * Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses.

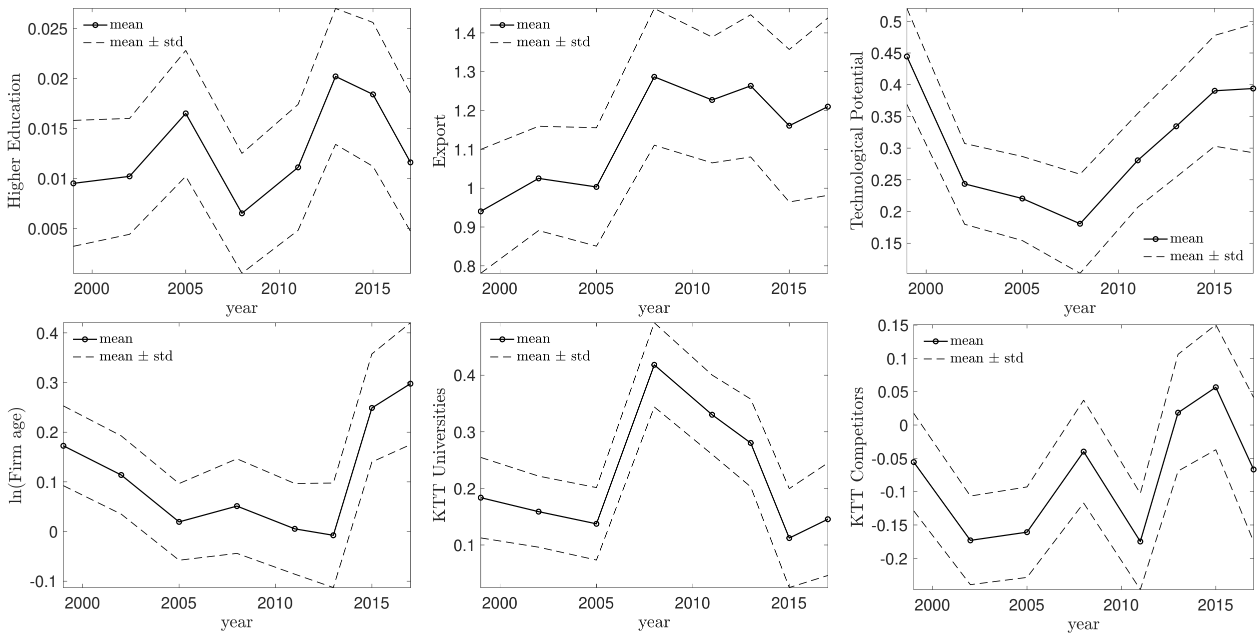


Figure E.1: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics higher education, export (yes/no), technological potential, log firm age, KTT universities, and KTT competitors, across different years for the manufacturing sector.

Figure E.1 shows the influence of various firm characteristics on the firm's innovation decision in the manufacturing sector. We observe that higher education, export orientation, the technological potential and the knowledge and technology transfer (KTT) collaborations have a stable or increasing importance on the innovation decision.

Figure E.2 shows the effect of competition on the innovation decision in the manufacturing sector. Similar to the pooled sample across sectors we find that across all years, competition tends to have a negative effect on innovation.

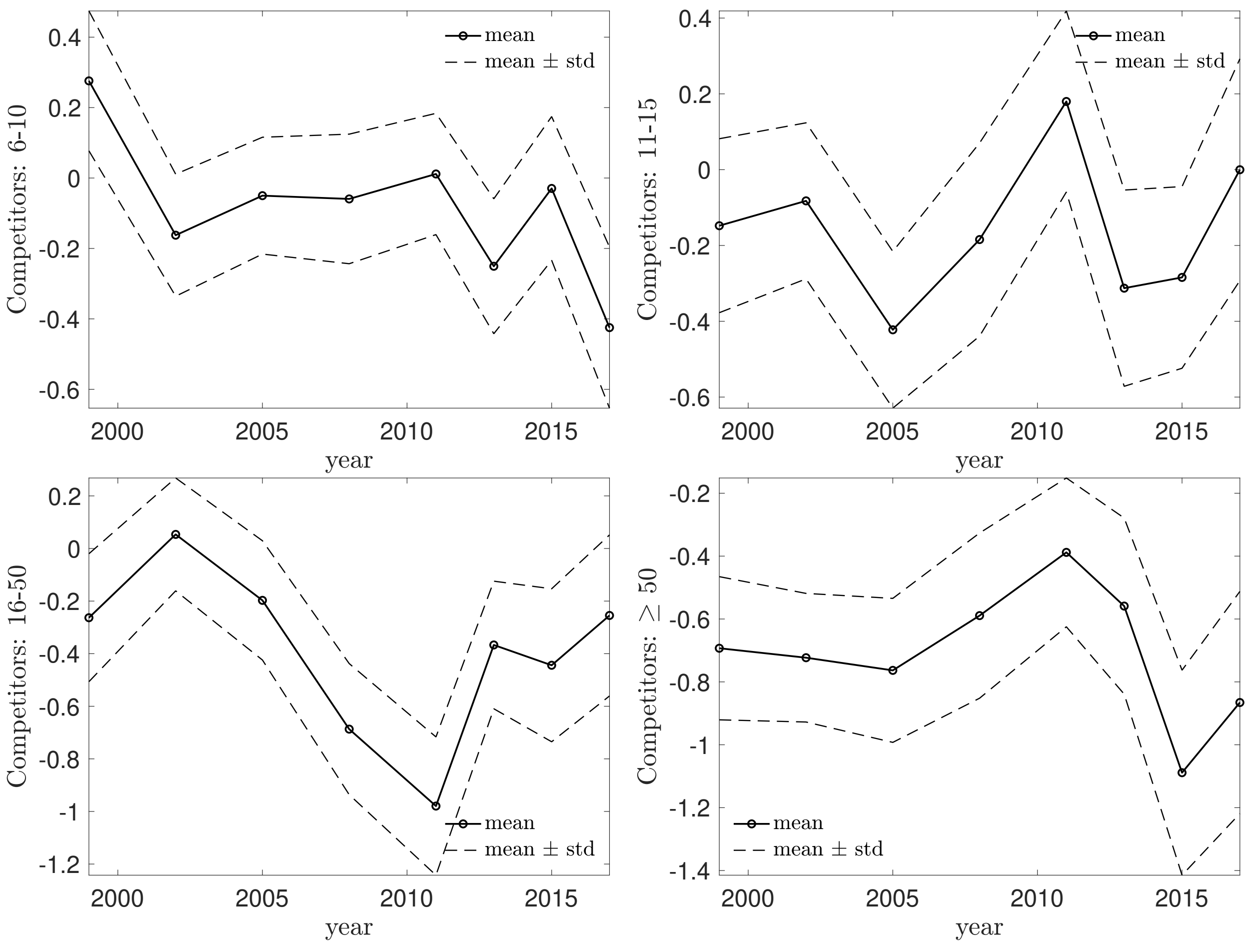


Figure E.2: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics related to competition across different years for the manufacturing sector.

E.2 Services Sector

Table 24 shows the Nonlinear Least Squares (NLS) parameter estimates across the years from 1999 to 2017 for the services sector.

Table 24: Linear Probability Model (LPM) and Nonlinear Least Squares (NLS) estimation results across years for the services sector.

		1996	1999	2002	2005	2008	2011	2013	2015	2017
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Linear Probability Model (LPM) w/o Passive Imitation								
Innovation	(p)	0.4177	0.6558	0.2927	0.4423	0.2489	0.2224	0.2760	0.2237	0.3180
		(0.0352)	(0.3442)	(0.2109)	(0.1307)	(0.0230)	(0.0216)	(0.0222)	(0.1108)	(0.2205)
Imitation	(q)	0.0286	-0.0617	0.0734	0.1587	0.1403	0.0903	0.2140	0.1669	0.1024
		(0.0568)	(0.1820)	(0.1284)	(0.0920)	(0.0406)	(0.0388)	(0.0409)	(0.0844)	(0.1360)
Innov. Cost	($\widetilde{\kappa}$)	0.0000	0.4445	0.0130	0.1037	0.0000	0.0000	0.0000	0.0067	0.1145
		(0.0000)	(0.9490)	(19.8088)	(1.3419)	(0.0000)	(0.0000)	(0.0000)	(13.5547)	(2.3055)
		Nonlinear Least Squares (NLS) with Passive Imitation								
Innovation	(p)	0.4140	0.8334	0.2927	0.4423	0.2489	0.2224	0.2760	0.2037	0.3180
		(0.0324)	(0.3920)	(0.2109)	(0.1307)	(0.0230)	(0.0216)	(0.0222)	(0.0831)	(0.2205)
Innov. Cost	($\widetilde{\kappa}$)	0.0000	0.6597	0.0130	0.1037	0.0000	0.0000	0.0000	0.0015	0.1145
		(0.0000)	(0.7259)	(19.8087)	(1.3419)	(0.0000)	(0.0000)	(0.0000)	(41.0406)	(2.3055)
Imitation	(q)	0.8422	-0.2355	0.0734	0.1587	0.1403	0.0903	0.2140	0.5685	0.1024
		(0.9560)	(0.1929)	(0.1284)	(0.0920)	(0.0406)	(0.0388)	(0.0409)	(0.4043)	(0.1360)
Pass. Imit.	(δ)	0.9922	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8378	0.0000
		(6.6033)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(1.8827)	(0.0000)

Notes: P-values are computed under the assumption of an asymptotic normal distribution of the estimators: *** Statistically significant at 1% level. ** Statistically significant at 5% level. * Statistically significant at 10% level. Robust (i.e. heteroskedasticity consistent) standard errors in parentheses.

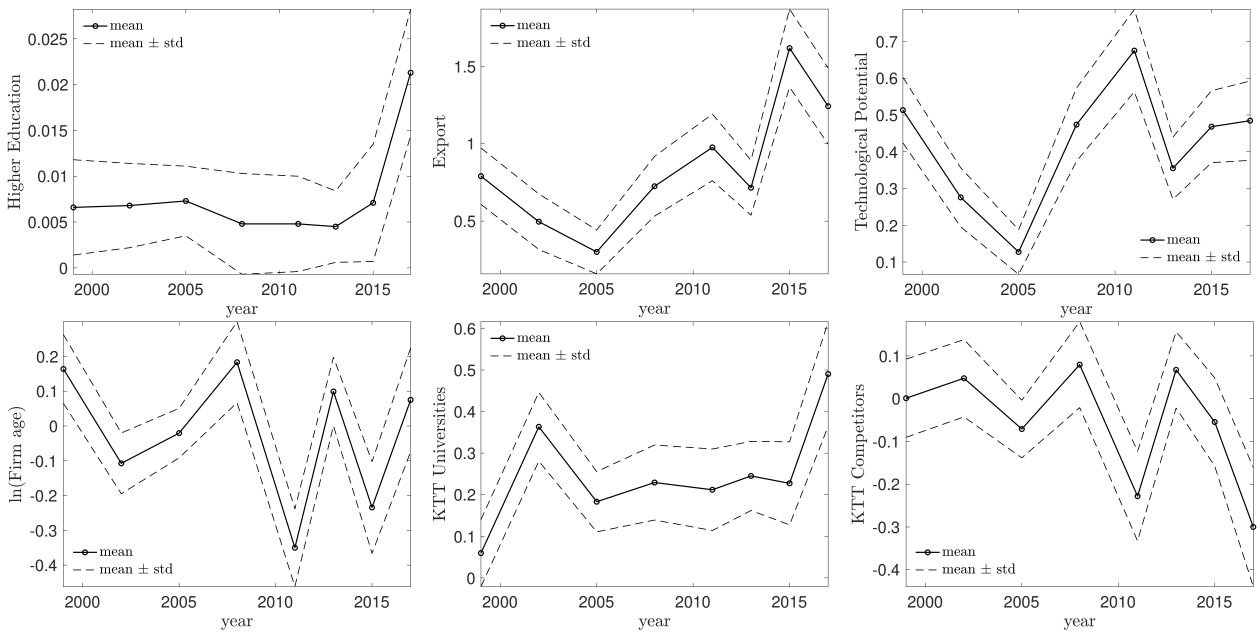


Figure E.3: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics higher education, export (yes/no), technological potential, log firm age, KTT universities, and KTT competitors, across different years for the services sector.

Figure E.3 shows the influence of various firm characteristics on the firm's innovation decision in the services sector. We observe that higher education, export orientation, the technological potential and the knowledge and technology transfer (KTT) university collaborations have a stable or increasing importance on the innovation decision. Collaborations with competitors have a declining influence.

Figure E.4 shows the effect of competition on the innovation decision in the services sector. Similar to the pooled sample across sectors we find that across all years, competition tends to have a negative effect on innovation.

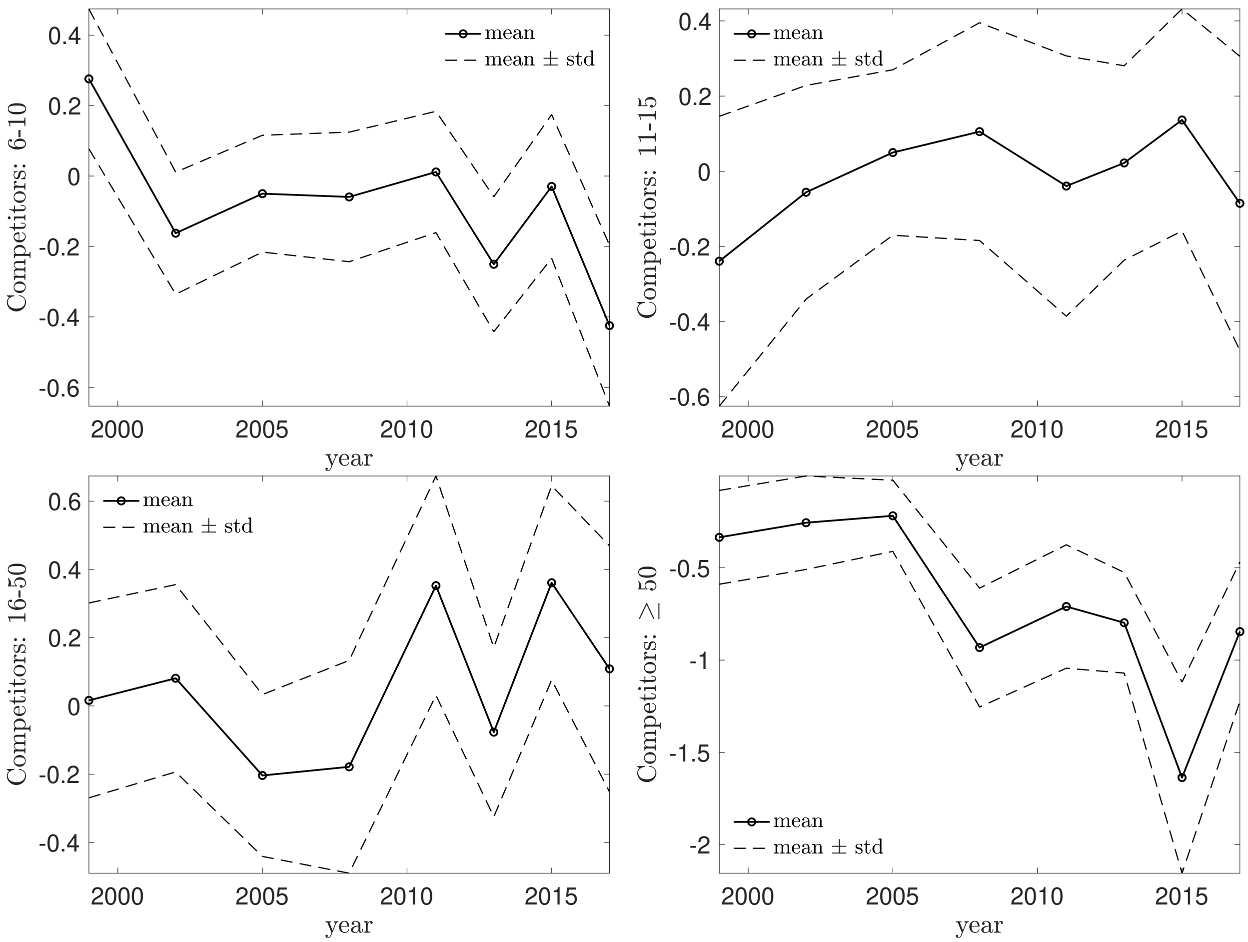


Figure E.4: Nonlinear Least Squares (NLS) model parameter estimates for the firm characteristics related to competition across different years for the services sector.

E.3 R&D Funding Costs

We consider a subsidy, $s_i(t) \in [0,1]$, to the R&D costs of the firm. More specifically, firm i's current (period t) profits are given by

$$\pi_i(t) = \psi A_i(t) - (1 - s_i(t))c_i(t), \tag{18}$$

where we consider two possible cases: In the first case, we consider is a uniform subsidy identical for all firms, $s_i(t) = s$, while in the latter case, we consider a subsidy only for the firms below the threshold, that is

$$s_i(t) = s \cdot \chi^{\text{im}}(a_i, p, P(t)) \tag{19}$$

where $s \in [0,1]$ is the fraction of the R&D cost subsidized by the government. The long-run cost of the uniform subsidy is given by

$$\begin{aligned} C_{\text{uniform}}(s) = \lim_{t\to\infty}\sum_{i=1}^{n} s_i(t)c_i(t) &= \lim_{t\to\infty}\sum_{i=1}^{n} s_i(t)c_i(t) \\ &= \lim_{t\to\infty} \kappa\overline{A}(t)^{\theta}\sum_{i=1}^{n} s_i(t)A_i\left(t\right)^{1-\theta} \\ &= \lim_{t\to\infty} s\kappa e^{\theta\bar{a}(t)}\sum_{i=1}^{n} e^{(1-\theta)a_i(t)} \\ &= \lim_{t\to\infty} s\kappa n e^{\theta\bar{a}(t)}\sum_{b=1}^{\infty} e^{(1-\theta)b}P_b(t), \end{aligned} \tag{20}$$

while the the long-run cost of the threshold subsidy is given by

$$\begin{aligned} C_{\text{threshold}}(s) = \lim_{t\to\infty}\sum_{i=1}^{n} s_i(t)c_i(t) &= \lim_{t\to\infty}\sum_{i=1}^{n} s_i(t)c_i(t) \\ &= \lim_{t\to\infty} \kappa\overline{A}(t)^{\theta}\sum_{i=1}^{n} s_i(t)A_i\left(t\right)^{1-\theta} \\ &= \lim_{t\to\infty} s\kappa e^{\theta\bar{a}(t)}\sum_{i=1}^{n} \chi^{\text{im}}(a_i, p, P(t))e^{(1-\theta)a_i(t)} \\ &= \lim_{t\to\infty} s\kappa n e^{\theta\bar{a}(t)}\sum_{b=1}^{\infty} \chi^{\text{im}}(b, p, P(t))e^{(1-\theta)b}P_b(t), \end{aligned} \tag{21}$$

where the log-productivity distribution $P_a(t)$ is determined by Equation (13), the imitation decision variable $\chi^{\text{im}}(b, p, P)$ without the subsidy is given by Equation (14) while the decision variable for the firms that receive the subsidy because they are below the threshold is given by

$$\widetilde{\chi}^{\text{im}}(a, p, P, s) = 1 - \widetilde{\chi}^{\text{in}}(a, p, P, s) = \begin{cases} 1 & \text{if } p < \frac{(1-\delta)q(1-F_a) + s\widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q(1-F_a)}, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

F Proofs

Proof of Proposition 1. For simplicity, we consider the case in which all firms have the same in-house R&D success probability p. The generalization to heterogeneous probabilities is straight forward. The evolution of the log-productivity distribution $P_a(t)$ can be written as

$$\begin{aligned} P_a(t+\Delta t) - P_a(t) = &\left(\chi^{\text{im}}(a-1,p,P) + \delta\chi^{\text{in}}(a-1,p,P)(1-p)\right) q(1-F_{a-1}(t))P_{a-1}(t) \\ &- \left(\chi^{\text{im}}(a,p,P) + \delta\chi^{\text{in}}(a,p,P)(1-p)\right) q(1-F_a(t))P_a(t) \\ &+ \chi^{\text{in}}(a-1,P)pP_{a-1}(t) - \chi^{\text{in}}(a,P)pP_a(t), \end{aligned} \tag{23}$$

where δ is the passive imitation probability, and

$$\chi^{\text{im}}(a,p,P) = 1 - \chi^{\text{in}}(a,p,P) = \begin{cases} 1 & \text{if } p < \frac{(1-\delta)q(1-F_a)+\widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q(1-F_a)}, \\ 0 & \text{otherwise,} \end{cases} \tag{24}$$

with the average log-productivity given by $\bar{a} = \sum_{a=1}^{\infty} F_a$. The first term in Equation (23) corresponds to the case that a firm with log-productivity $a-1$ is selected, times the indicator that it wants to imitate, $\chi^{\text{im}}(a-1,p,P) = 1$, or that it wants to innovate, failed to do so and then engages in passive imitation, $\chi^{\text{in}}(a-1,p,P)(1-p)\delta$, times the probability that it draws a firm with log-productivity larger than $a-1$ and successfully improves its log-productivity by one unit with probability q. The second term corresponds to the events that a firm with log-productivity a is selected and successfully imitates. The third term corresponds to the case that a firm with log-productivity $a-1$ is selected, wants to innovate, $\chi^{\text{in}}(a-1,p,P) = 1 - \chi^{\text{im}}(a-1,p,P) = 1$, and succeeds to improve its log-productivity by one unit with probability p. The fourth term corresponds to the case that a firm with log-productivity a is selected, wants to innovate, $\chi^{\text{in}}(a,p,P) = 1 - \chi^{\text{im}}(a,p,P) = 1$, and succeeds with probability p. Finally, one can check from Equation (23) that for all $t \geq 0$: $\sum_{a=1}^{\infty}(P_a(t+\Delta t) - P_a(t)) = 0$. □

Proof of Lemma 1. Assuming a uniform distribution of the in-house R&D success probabilities allows us to write Equation (13) as follows

$$\begin{aligned} \frac{\partial P_a(t)}{\partial t} = \frac{1}{\overline{p}-\underline{p}} \int_{[\underline{p},\overline{p}]} \Big[&\left(\chi^{\text{im}}(a-1,p,P) + \delta(1-p)\chi^{\text{in}}(a-1,p,P)\right) q(1-F_{a-1}(t))P_{a-1}(t) \\ - \left(\chi^{\text{im}}(a,p,P) + \delta(1-p)\chi^{\text{in}}(a,p,P)\right) q(1-F_a(t))P_a(t) &+ \chi^{\text{in}}(a-1,p,P)pP_{a-1}(t) - \chi^{\text{in}}(a,p,P)pP_a(t) \Big] dp. \end{aligned}$$

Using the fact that

$$\int_{[\underline{p},\overline{p}]} \chi^{\text{im}}(a,p,P)dp = \int_{[\underline{p},\overline{p}]} \mathbb{1}_{\{p<C(a,P)\}}dp = (\min\{C(a,P),\overline{p}\} - \underline{p})\mathbb{1}_{\{C(a,P)>\underline{p}\}},$$

and

$$\int_{[\underline{p},\overline{p}]} p\chi^{\text{in}}(a,p,P)dp = \int_{[\underline{p},\overline{p}]} p\mathbb{1}_{\{p>C(a,P)\}}dp$$
$$= \int_{[\max\{\underline{p},C(a,P)\},\overline{p}]} pdp\mathbb{1}_{\{C(a,P)<\overline{p}\}}$$
$$= \frac{1}{2}\left(\overline{p}^2 - (\max\{\underline{p},C(a,P)\})^2\right)\mathbb{1}_{\{C(a,P)<\overline{p}\}},$$

and

$$\int_{[\underline{p},\overline{p}]} (1-p)\chi^{\text{in}}(a,p,P)dp = \int_{[\underline{p},\overline{p}]} (1-p)\mathbb{1}_{\{p>C(a,P)\}}dp$$
$$= \int_{[\max\{\underline{p},C(a,P)\},\overline{p}]} (1-p)dp\mathbb{1}_{\{C(a,P)<\overline{p}\}}$$
$$= \frac{1}{2}\left. p(2-p)\right|^{\overline{p}}_{\max\{\underline{p},C(a,P)\}}\mathbb{1}_{\{C(a,P)<\overline{p}\}}$$
$$= \frac{1}{2}\left(\overline{p}(2-\overline{p}) - \max\{\underline{p},C(a,P)\}(2-\max\{\underline{p},C(a,P)\})\right)\mathbb{1}_{\{C(a,P)<\overline{p}\}},$$

where we have denoted by

$$C(a,P) \equiv \frac{(1-\delta)q(1-F_a) + \widetilde{\kappa}e^{\theta(\overline{a}-a)}}{1-\delta q\,(1-F_a)}, \tag{25}$$

we can write

$$\begin{aligned}
\frac{\partial P_a(t)}{\partial t} = \frac{1}{\overline{p}-\underline{p}}\Big[& q(1-F_{a-1}(t))P_{a-1}(t)\Big((\min\{C(a-1,P),\overline{p}\} - \underline{p})\mathbb{1}_{\{C(a-1,P)>\underline{p}\}} \\
&+ \delta\frac{1}{2}\left(\overline{p}(2-\overline{p}) - \max\{\underline{p},C(a-1,P)\}(2-\max\{\underline{p},C(a-1,P)\})\right)\mathbb{1}_{\{C(a-1,P)<\overline{p}\}}\Big) \\
&- q(1-F_a(t))P_a(t)\Big((\min\{C(a,P),\overline{p}\} - \underline{p})\mathbb{1}_{\{C(a,P)>\underline{p}\}} \\
&+ \delta\frac{1}{2}\left(\overline{p}(2-\overline{p}) - \max\{\underline{p},C(a,P)\}(2-\max\{\underline{p},C(a,P)\})\right)\mathbb{1}_{\{C(a,P)<\overline{p}\}}\Big) \\
&- \frac{1}{2}P_a(t)\left(\overline{p}^2 - (\max\{\underline{p},C(a,P)\})^2\right)\mathbb{1}_{\{C(a,P)<\overline{p}\}} \\
&+ \frac{1}{2}P_{a-1}(t)\left(\overline{p}^2 - (\max\{\underline{p},C(a-1,P)\})^2\right)\mathbb{1}_{\{C(a-1,P)<\overline{p}\}}\Big].
\end{aligned} \tag{26}$$

From Equation (25) we see that $C(a,P) \geq 0$. Further, assuming that $\underline{p} = 0$ allows us to write

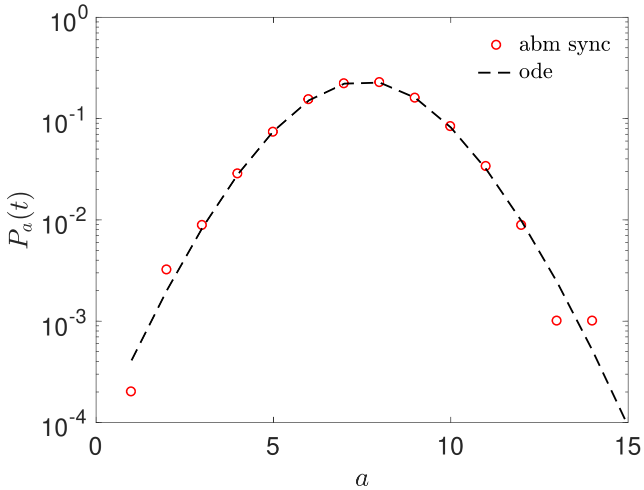


Figure F.1: The log-productivity distribution $P_a(t)$ from a Monte Carlo simulation of the stochastic process indicated with circles. The dashed line indicates the solution of the ODE of Equation (27).

$$
\begin{aligned}
\frac{\partial P_a(t)}{\partial t} = \frac{1}{\overline{p}} \Big[& q(1 - F_{a-1}(t))P_{a-1}(t) \Big(\min\{C(a-1,P), \overline{p}\} \\
& + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a-1,P)(2 - C(a-1,P))\right) \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} \Big) \\
& - q(1 - F_a(t))P_a(t) \Big(\min\{C(a,P), \overline{p}\} \\
& + \delta \frac{1}{2} \left(\overline{p}(2-\overline{p}) - C(a,P)(2 - C(a,P))\right) \mathbb{1}_{\{C(a,P)<\overline{p}\}} \Big) \\
& - \frac{1}{2} P_a(t) \left(\overline{p}^2 - C(a,P)^2\right) \mathbb{1}_{\{C(a,P)<\overline{p}\}} \\
& + \frac{1}{2} P_{a-1}(t) \left(\overline{p}^2 - C(a-1,P)^2\right) \mathbb{1}_{\{C(a-1,P)<\overline{p}\}} \Big] .
\end{aligned} \tag{27}
$$

□

Figure F.1 shows the log-productivity distribution $P_a(t)$ from a numerical solution of Equation (27).