Essays on Entrepreneurship and Pricing

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Niklas Karl Garnadt

im Herbstsemester 2019

Abteilungssprecher:	Prof. Dr. Hans Peter Grüner
Referent:	Prof. Klaus Adam, Ph.D.
Korreferent:	Prof. Konrad O. Stahl, Ph.D.
Tag der Verteidigung:	18.03.2020

Acknowledgements

I would like to thank my adviser Klaus Adam for his dedication as a mentor and supervisor, for giving me the freedom to pursue my own research interests and for valuable feedback throughout the various stages of my dissertation.

My co-author Konrad Stahl has also been an exceptional mentor and supervisor without whom it would not have been possible to write the second paper of my dissertation. I thank him for supporting my research interest in entrepreneurship, for his excellent writing advice, for his thoughtful and creative feedback on my research and for being an excellent motivator.

For the constructive discussions about our joint work on entrepreneurship and his deep insights on Swedish data I want to thank my co-author Joacim Tåg. I am also very thankful to Andreas Gulyas who put in a lot of effort to support me during the job market phase. The research in this thesis benefited from advice and constructive feedback on my research by the Mannheim Macro group, in particular Johannes Pfeifer and Matthias Meier, the participants of the Mannheim Macro Lunch, and Klaus Adam's Macro Colloquium.

I would also like to thank Susanne Putz and Nadine Scherer for administrative support, particularly in travel matters and during the job market phase. The CDSE administration, in particular Sandro Holzheimer, Marion Lehnert, Dagmar Röttsches, Golareh Khalilpour and Caroline Mohr also made Ph.D. my life easier.

I also want to thank my Ph.D. friends Alessandra Allocca, Thomas Bonczek, Florian Böser, Esteban Cattaneo, Tobias Etzel, Hanno Förster, Lukas Henkel, Ruben Hipp, Mario Meier, Sebastian Merkel, Timo Reinelt, Richard Senner, André Stenzel and many others who helped me with completing this thesis through helpful disucssions and by making graduate studies much more fun.

I am very thankful to my parents who supported me not only throughout this dissertation but my whole life. I also want to thank Patricia for her ongoing patience, encouragement and support without which this dissertation would not have been possible.

Contents

List of Figures

List of Tables

General Introduction

This dissertation studies questions about the interrelationship between the macroeconomic environment and entrepreneurship as well as questions on optimal pricing. It is comprised of three chapters which can be read independently. Chapters 1 and 2 are concerned with the effects of macroeconomic trends and fluctuations on entrepreneurship. In Chapter 1 I investigate the long run decline in firm creation rates in the US since the early 1980s. I document new empirical facts on the contemporaneous increase in average firm size and relate them to the decline in firm creation rates in a quantitative model. In the second Chapter which is co-authored with Konrad Stahl and Joacim Tåg we empirically analyze the relationship between variation in entrepreneurial quality and variation in firm quality and relate both to business cycle fluctuations. In Chapter 3 I analyze a theoretical model of optimal monopoly pricing in which the monopolist is uncertain about the demand he faces which introduces new tradeoffs into the optimal pricing problem.

Chapter 1 is titled "The Increasing Presence of Large Firms and the Decline in US Startup Rates". In this Chapter I propose a new explanation for the significant decline in firm formation rates in the US over the past 30 years. While this decline has raised concern about the health of the US economy, its causes are not yet fully understood. I argue that a significant part of it can be explained as an efficient response to size biased technological change. I document an increase in the size of large firms in the US since the mid 1980s contemporaneous to the decline in firm formation rates and show that large firms expanded by sharply increasing the number of establishments they operate. These changes in the organizational structure of large firms are consistent with improvements in information and communications technology which mainly benefitted large firms by reducing monitoring, coordination and distribution costs. I construct a simple industry dynamics model in which an expansion by large firms due to reductions in the cost of managing many establishments crowds out smaller firms that are responsible for most of firm turnover. While generating a decrease in startup rates and an increase in the size of large firms through an increase in the number of establishments they operate, welfare improves. These results counter the popular opinion that declining firm formation rates are necessarily a bad sign.

Chapter 2, which is joint work with Konrad Stahl from the University of Mannheim and Joacim Tåg from IFN Stockholm is titled "On the Cyclicality of Entrepreneur and Startup Characteristics". In this Chapter we investigate the relationship between variations in entrepreneur characteristics and startup characteristics in the cross section and over time. We establish substantial cross-sectional differences between firms started by different entrepreneur types, in particular regarding firm size and firm growth. As 40% of the variation in employment created at birth and 70% of employment created by age 5 is due to variation in the average size of firms, we investigate whether variation in the composition of entrepreneurs is an important driver of variation in average firm size. Decomposing variation in average firm size into variation within entrepreneur segments over time and into variation in the composition of entrepreneurs shows that between 13.5% and 23.5% of variation in startup size is due to variation in the composition of entrepreneurs. We also investigate whether these margins vary systematically over the business cycle. While it is well known that there is a significant drop in the number of new firms during recessions, there is a longstanding debate over whether these are of higher or lower quality. Recent research has documented for the U.S. that firms started in recessions have worse long run growth prospects i.e. are of lower quality. Employment creation by startups and its main margins, the number of firms and the average firm size, exhibit a similar cyclicality in Sweden as in the US. Employment creation by startups falls in times of high unemployment both due to a significant decrease in the number of startups and their average size. Disaggregating these measures by entrepreneur types, we show that the significant decline in the average size of startups on aggregate is due to two effects, on the one hand a higher share of firms started by low quality entrepreneurs and on the other hand a decrease in the size of firms started by high quality entrepreneurs.

In Chapter 3, which is titled "Demand Learning with Nonperishable Products - Multiunit Firms and Inventory Replenishment", I investigate how learning dynamics affect the optimal pricing problem of a monopolist. When setting optimal prices, firms are often uncertain about the actual demand they face. Thus in addition to solving the classic price-quantity tradeoff, firms also have to take into account how price setting affects the speed at which they learn about demand. While high prices increase revenue per product they also deter customers from buying, and thus reduce the speed at which firms can learn about demand through observed sales volumes. I extend the model of Mason and Välimäki (2011) from firms who only have a single unit in their inventory to multiunit firms, which gives rise to a much richer learning process. I prove several intuitive characteristics of the optimal pricing policy and show that higher uncertainty increases prices as firms don't want to set a suboptimal price and rather wait until they have better demand information in order to set the correct price.

Chapter 1

The Increasing Presence of Large Firms and the Decline in US Startup Rates

1.1 Introduction

A strong entrepreneurial culture that fosters the creation of new firms and the constant reallocation of productive resources to their most efficient use is commonly considered to be a defining feature of the US economy. Accordingly, recent empirical studies documenting the secular decline in the rates of new firm formation and reallocation within most sectors and states in the US have garnered considerable attention by policymakers and the media.[1] Numerous studies have underscored the importance of entrepreneurship for the creation of new jobs, the introduction of new technologies and products as well as productivity growth. It is thus of obvious importance to understand whether firm formation rates declined as an efficient response to changes in the economic environment or as the result of policy distortions that inhibit firm creation and the reallocation of resources to more efficient uses.

In this paper, I argue that an essential part of the decline in firm creation rates can be explained as an efficient response to changes in information and communication technologies (ICT). The argument rests on three ideas: first, large firms benefit more than small firms from ICT investment. Thus improvements in ICT lead to an expansion of large firms while crowding out smaller competitors. Second, large firms are mainly active in mass-markets where products and services are standardized and thus more amenable to mass-production or codification. Third, as products are standardized, it is also relatively

[1] See Davis et al. (2006), Reedy and Strom (2012) for evidence on declining firm creation rates, Yellen (2014), Council of Economic Advisors (2016) for policymakers commenting on the declining firm creation rate and The Economist (2014), Harrison (2015), Smith (2016) for related media reports.

cheap for small firms to enter the mass-market compared to specialized markets which require investment in product-specific know-how. Thus small-firms on mass-markets exhibit higher entry and exit rates than on specialized markets and are the primary source of firm turnover. Expansion of large firms increases competition mainly in mass-markets, leading to a shift in the firm composition away from small mass-market firms. Through this compositional effect, the aggregate firm creation and destruction rates decrease.

The first contribution of this paper is to document several new stylized facts on the expansion of large firms with more than 1000 employees contemporaneous to the decline in firm formation rates since the 1980s. Using the Business Dynamics Statistics (BDS) data of the US Census, which covers all private sector employment in the United States, I show that the increase in average firm size in the US since the mid 1980s was caused roughly in equal parts by an increase in the *size* of large firms and an increase in the *share* of large and medium firms in the economy. In all sectors, large firms expanded by adding new establishments: compared to the mid-1980s, large firms in the 2010s operate 40 % more establishments. Within most sectors, the average number of employees per establishment declined for establishments operated by large firms, on average by 18%.[2] For the aggregate economy, the share of establishments operated by large firms has increased from around 10% in the 1980s to 16% in the 2010s.[3] The specific margins of expansion suggest that the optimal organization of large firms has changed significantly since the 1980s towards operating more but smaller establishments. This change in the organizational structure is consistent with the literature on improvements in ICT which led to reductions in monitoring, coordination and distribution costs, enabling large firms to more efficiently manage a large number of establishments. Together with the facts on the increasing size and share of large firms in the economy, this suggests that large firms increasingly enjoy a technological advantage relative to smaller firms.

As a second contribution in this paper, I quantify the effect of changes in the technology of large firms on the firm creation rate. To this end, I build an industry dynamics model with endogenous entry to study the contribution of changes in entry cost and population growth as well as changes in large firms' establishment level fixed costs and returns to scale to the observed decline in firm formation rates. The model features an industry with two distinct segments: a mass-market and a specialized market with two main differences between the segments.[4] First, standardization allows multi-establishment firms to operate

[2]The exception is the retail sector where large firms' average establishment size has increased by about 10%. As a comparison, in services large firms' average establishment size has decreased by 25% in manufacturing by 30%.

[3]This fact holds by and large within sectors. The exception is manufacturing, where the share of establishments operated by large firms has remained stable at 10%.

[4]The distinction between mass-markets and specialized markets can also be found in Holmes and Stevens (2014) and Bento (2014)

in the mass-market segment alongside small-scale single-establishment firms.[5] Second, whereas the higher degree of customization on the specialized market prevents replication across multiple establishments, the small firms on the specialized market exhibit lower exit rates than on the mass-market. Intuitively lower exit rates are due to specialized firms having invested in product-specific expertise making them more resilient to shocks.

The production technology exhibits decreasing returns to scale at the firm and the establishment level. Multi-establishment firms thus face a trade-off between expanding the size of their establishments and expanding the number of establishments they operate. Technological improvements specific to multi-establishment firms, in particular, IT-based improvements in back-office technology, offshoring and outsourcing of back-office tasks, and improvements in logistics of distributing intermediates, lead to a reduction in establishment level fixed cost and an increase in the returns to scale at the firm level.[6] As large firms expand their production, the competitive pressure for small firms in the mass-market increases. The expected value of starting a multi-establishment firm or a niche market firm increases relative to the value of starting a single establishment mass-market firm. The change in relative firm values induces a shift in the composition of entrants towards multi-establishment and niche-market single-establishment firms. I show that this compositional shift is the primary driver behind the decrease in firm creation rates. The decline of small mass-market firms is also consistent with observations from the retail, manufacturing and service sectors.

I calibrate the model to match salient features of the business dynamics statistics in the late 1980s. Subsequently, I recalibrate the model to the mid-2000s to quantify the relative contribution of changes in population growth, technology and startup costs to the observed decline in firm creation rates. The stylized facts on changes in firm and establishment size of large firms discipline the changes in the technology of multi-establishment firms and imply that firm-level returns to scale have slightly increased and establishment-level fixed costs have fallen. Furthermore, I find that startup costs for multi-establishment firms have increased, dampening the shift away from small-scale single establishment firms and the decline in firm formation rates. Decomposing the decline in firm creation rates I find that 20% is due to decreases in establishment level fixed cost, 40% is due to increases in firm-level returns to scale and 40% is due to lower population growth. From a welfare perspective the positive effects of increases in the returns to scale and lower establishment fixed costs of large firms dominate the increase in startup costs and imply an increase in per capita output by 7.3% relative to the late 1980s calibration, and this despite a lower

[5]The single establishment firms in the mass market are assumed to be firms owned by entrepreneurs whose primary motive is to be self-employed but not to grow. Hurst and Pugsley (2011) show that a significant fraction of entrepreneurs fit into this category of non-growth oriented entrepreneurs who directly derive utility from being self-employed.

[6]In Section 1.2.5 I discuss related research on these size-biased technological improvements and their connection to the stylized facts.

firm creation rate. While there already exists research which suggests that firm formation rates declined as an efficient response to changes in the economic environment, in contrast to this prior work I can match the stylized facts on changes in the organization of large firms.

As a final contribution, I use variation across MSAs to provide empirical support for the predictions of the model. According to the model, expansion by large firms increases competition causing lower firm startup rates. I show that the regions which experienced the most substantial increase in the fraction of establishments belonging to large firms between the late 1980s and the 2000s also experienced the largest drop in firm startup rates.

The rest of this paper is structured as follows. I discuss my contributions to the related literature next. In Section 1.2 I present the stylized facts on startup rates and the increasing importance and changing structure of large firms. A variant of the model of Bento (2014) is extended to multi-establishment firms and endogenous entry and exit of firms in Section 1.3. I calibrate the model in Section 1.4 and use it to calculate the relative contributions to the observed decline in firm formation rates, on the one hand of changes in entry cost and population growth and on the other hand of changes in large firms' establishment level fixed costs and returns to scale. Cross-sectional data is used in Section 1.5 to show that the model's prediction that an increase in the presence of large firms is correlated with a lower firm formation rate is consistent with the data. I conclude with Section 1.6.

Related literature This paper is related to three strands of literature. First, it relates to the literature documenting the decline in firm formation rates and business dynamism in the United States since the 1980s. Second, it is also related to the literature on increasing firm size and concentration in the United States. Last, my paper builds on the literature on models of industry dynamics with within industry differentiation between mass and niche markets.

Davis et al. (2006) were the first to note the decline in firm formation rates. Subsequent research, among others by Reedy and Strom (2012), Hyatt and Spletzer (2013), Davis and Haltiwanger (2014), Decker et al. (2014), Hathaway and Litan (2014a,b,c), Molloy et al. (2016) has documented that the decrease in business dynamism has been pervasive across sectors and regions in the US. Statistical decompositions show that compositional shifts towards older and larger firms can explain only a limited fraction of the decrease as most of the decrease in dynamism has occurred within groups defined by age, size, and industry. However, Decker et al. (2014) and Molloy et al. (2016) show that the decline in firm formation rates is an important driver of the decline in dynamism. The widespread decline in firm formation and dynamism suggests that structural forces common to most industries and geographies are responsible for this decline. In this paper, I propose such a structural force, namely size biased technological change.

However, there is no consensus on the underlying sources of the decline in firm formation rates. Hathaway and Litan (2014c) use variation in startup rate declines across Metropolitan Statistical Areas (MSA) and investigate several potential causes of the decline in business formation with OLS regressions. They find a strong regression to the mean effect: MSAs with large initial startup rates experienced the most substantial declines. They find a very weak correlation with business consolidation. In contrast to my analysis, they use average firm size in an MSA as a measure of business consolidation. I document that the average size of large firms' establishments decreased which could result in lower average firm size in an MSA even though large firms account for a higher fraction of local economic activity. To circumvent this problem, I use the fraction of establishments that belong to large firms as a measure of business consolidation and find that the MSAs which enjoyed a stronger increase in this measure experienced a stronger decrease in firm creation rates.

Davis and Haltiwanger (2014) propose more and stricter regulation, in particular on the labor markets, as a candidate explanation of the decrease in firm formation rates which would manifest themselves in higher entry, higher fixed, or higher adjustment costs. In a calibrated occupational choice model Kozeniauskas (2017) finds that higher entry costs are the primary driver of decreasing entrepreneurship rates. Using data on federal regulations across industries, Goldschlag and Tabarrok (2018) find that stricter regulations at the industry level do not explain lower startup rates.[7]

Demographic explanations are explored by Karahan et al. (2018) who look at the role of declining labor force growth and by Engbom (2017) and Kopecky (2017) who measure the effect of an aging population on firm formation rates and business dynamism. Jiang and Sohail (2017) and Salgado (2018) investigate the role of skill-biased technological change, noting that entrepreneurship rates have decreased more for the highly educated. Decreases in corporate tax rates are explored in Neira and Singhania (2017), declining interest rates in Chatterjee and Eyigungor (2018).

Summarizing, these underlying sources can be grouped into two groups: policy distortions that lead to an inefficient decrease in firm formation rates or exogenous technological or demographic developments which lead to an efficient downward adjustment of the firm formation rate. I show in this paper that a sizable fraction of the decline in startup rates can be attributed to an efficient adjustment to technological change specific to large firms, which at the same time led to an increase in the size of large firms. In contrast to the

[7]Furthermore stricter regulation should also negatively affect establishment creation by incumbents which is not observed in the data (Hathaway et al., 2014).

prior literature suggesting an efficient downward adjustment, this explanation can match the new stylized facts on changes in the organization of large firms.

The increase in firm size I document, especially for the largest firms, is related to the literature on increasing concentration and the superstar firm hypothesis. Autor et al. (2017) show that concentration in many industries increased significantly over the past 30 years. They propose the superstar firm hypothesis as an explanation which posits that technological improvements have advantaged large firms over smaller ones. Gutiérrez and Philippon (2017) and Grullon et al. (2017) relate the increase in employment concentration of large firms to increased market power and changes in the product market structure which allegedly has led to less competition.[8] Rossi-Hansberg et al. (2018) however, show that while concentration has increased at the national level, the expansion of large firms into local markets usually lead to a decrease in local measures of concentration. I contribute to this literature by showing that an essential part of the increase in average firm size can be attributed to the increasing size of large firms. Furthermore, I document that large firms grew primarily by expanding the number of establishments they operate, while at the same time reducing the average number of employees per establishment. These changes in the structure of large firms lend themselves to specific technological explanations[9] which I show are consistent with a decrease in firm formation rates.

Technically I contribute to the literature on industry dynamics with within industry differentiation between mass and niche market firms. Holmes and Stevens (2014) show that incorporating niche market firms into a model of international trade is essential to generate the observed responses of US firms to trade liberalization. The critical insight is that many small firms are not just small because they had a bad productivity draw as in standard heterogeneous firms models (Hopenhayn (1992), Melitz (2003)) but that they operate in a market which only allows for small firms. Starting a niche-market firm thus limits the growth opportunities but on the other hand, protects firms from the competition of large mass-market firms. This force is at work as well in my model, where improvements in the technology of large mass market firms lead to a shift away from small mass market firms towards niche market firms. Bento (2014) develops a static general equilibrium version of this model to show that accounting for niche-market firms operating in markets with low substitutability of products exacerbates efficiency losses from high entry costs. I contribute to this literature by incorporating exit and endogenous entry

[8]Decreasing competition and increasing markups have been investigated by De Loecker and Eeckhout (2017). Traina (2018) questions their measure of variable costs and argues that the increase in markups has been much lower.

[9]For example, decreases in spatial frictions as found by Giroud (2013) for manufacturing firms in the US and Behrens and Sharunova (2015) for firms in Canada would lead to an expansion of firms through the creation of new establishments in new markets. This would also be consistent with decreasing local concentration as in Rossi-Hansberg et al. (2018).

into an industry dynamics model with niche and mass-market firms.[10] I show that the distinction between niche and mass-market firms is essential to match the decline in entry rates and that the shift from small mass market to small niche market firms is a potential explanation for declines in business dynamism within firm-age-industry cells.

1.2 Empirical analysis

The rate of firm formation in the US has been falling steadily since the 1980s. After describing the data I use for the empirical analysis and revisiting this fact, this section presents three stylized facts on contemporaneous changes in firm characteristics.

Fact 1 Average firm size has increased since the 1980s. This increase has been caused both by an increase in the size of large firms and an increase in the share of medium and large firms.

Fact 2 Large firms have increased their size by expanding the number of establishments per firm while at the same time the average establishment size of large firms has decreased.

Fact 3 Large firms account for an increasing share of establishments in the aggregate economy, within most sectors and almost all US metropolitan areas.

Fact 1 is related to recent papers on increasing concentration and the superstar firm hypothesis which show that an increasing share of economic activity is accounted for by large firms.[11] To my knowledge the opposing changes in the average number of establishments per firm and the average number of employees as well the dominance of the establishment expansion margin stressed by fact 2 have not been documented in the literature except for two contemporaneous contributions by Aghion et al. (2019) and Hsieh and Rossi-Hansberg (2019). Fact 3 on the increasing share of establishments operated by large firms is related to the notion of increased business consolidation across US metros in Hathaway and Litan (2014b) as well as recent research by Rossi-Hansberg et al. (2018) on the pro-competitive effects of the expansion of large firms across the US.

[10] Sedláček and Sterk (2017) build a dynamic general equilibrium business cycle model with endogenous entry to analyze the effect of business cycle shocks on the niche-market vs. mass-market composition of entrants. However, these firms all operate in the same market, i.e., the competition effect from new niche market firms is the same as the competition effect from new mass-market firms. Niche vs. mass market in their terminology measures the ability of firms to become large.

[11] See, e.g. Autor et al. (2017), De Loecker and Eeckhout (2017) and Gutiérrez and Philippon (2017)

1.2.1 Data description

In order to analyze business formation rates and the firm size distribution, I use the publicly available Business Dynamics Statistics (BDS) data product of the US Census. The BDS is based on the US Census' Longitudinal Business Database (LBD) which is a yearly panel containing establishment level information on the employment of nearly all private US establishments with at least one payroll employee in the week of March 12th each year. The Census is also able to aggregate individual establishments into firms, where a firm is a collection of all establishments under common ownership. The public use BDS data contains yearly data on, e.g. the number of firms, the number of establishments and employment, aggregated along different strata as, e.g. sectors, MSAs, firm size groups or firm age groups. The data run from 1976 to 2014.[12]

In this data set, firm age is assigned according to the age of the oldest establishment in a firm. Establishments age naturally one year at a time and are assigned age 0 in the year when they first report positive payroll employment. I will denote age 0 firms as startups; thus a startup is a firm which is responsible for genuinely new business activity at a new location. M&A activity or legal reorganization of firms thus will not be classified as a startup.

1.2.2 Changing entry and exit patterns

In Figure 1.1 I plot the time series of the aggregate firm formation rate between 1980 and 2013. I define the firm formation rate in the year t as the number of new (age 0) firms in the year t divided by the total number of firms in the previous year $t-1$. The firm formation rate declined in steps, from about 13% in the 1980s to 11% in the 1990s, gradually going down to 10% in the 2000s and dropping to about 8% after the great recession. The trendline in Figure 1.1 is a LOESS smoothed trend. This decline was not confined to a specific sector but occurred across all broad sectors of the economy.[13] By 2012, as the US recovered from the great recession, the firm startup rate of every broad sector was at least 25% below the level of 1980.

The decline in the firm startup rate was not countervailed by an increase in the size of new firms. The employment-weighted firm startup rate, i.e., the number of new jobs created by new firms in the year t divided by the total number of jobs in the US economy in the previous year, shows a very similar time series behavior as the unweighted firm

[12] I will be using data starting in 1979, due to data problems in 1977 and 1978 as noted by Moscarini and Postel-Vinay (2012). Furthermore, when using MSA data, I will start in 1981 due to data problems for Texas MSAs in 1979 and 1980.

[13] I use the sector definition of the BDS dataset which is based on the SIC industrial classification. The sectors and their abbreviations are agriculture (AGR), mineral extraction (MIN), construction (CON), manufacturing (MAN), transportation, communications and utilities (TCU), wholesale trade (WHO), retail trade (RET), finance, insurance and real estate (FIRE) and services (SRV).

formation rate. The weighted firm startup rate dropped from 3.75% in the early 1980s to 3% in the 1990s and early 2000s to 2% in the aftermath of the great recession. The sectoral picture shows more differences in the individual time series than for the unweighted firm startup rate. While the employment-weighted firm startup rate is mostly flat until before the great recession in finance, manufacturing and transportation, communications and utilities, there is a notable decline since the end of the 1980s in services, retail and wholesale as well as construction. Notwithstanding these differences, by 2012 all sectors had an employment-weighted firm startup rate of at least 25% below the level in 1980.

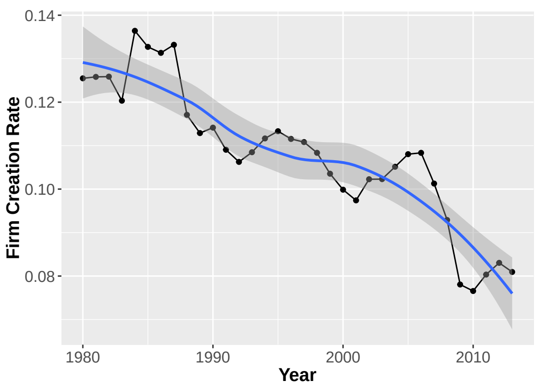


Figure 1.1: Firm creation rate

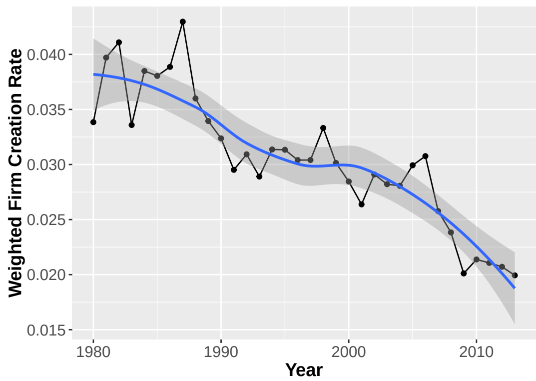


Figure 1.2: Employment weighted firm creation rate

Compared to the firm creation rate, the firm exit rate has remained slightly more stable between 1980 and 2013. In Figure 1.3 I plot the time series of the firm exit rate. The firm exit rate is defined as the number of firms that had positive employment in the year $t-1$ but have no positive employment in the year t, divided by the number of firms in the year $t-1$. Firm exit rates declined from slightly above 9% in the 1980s to around 8% just before and after the great recession. At the sectoral level, this decline is most pronounced in retail while manufacturing has only experienced a slight downward trend.

The weighted firm exit rate shows a stronger downward trend. The weighted firm exit rate is measured as the number of jobs destroyed by exiting firms divided by the total number of jobs in the previous year. The aggregate weighted firm exit rate declined from around 3% in the 1980s to 2.5% in the late 1990s and 2000s and declined slightly more after the great recession. This decline is stronger than for the unweighted firm exit rate which suggests that the size of exiting firms relative to the size of continuing firms declined between 1980 and 2010.

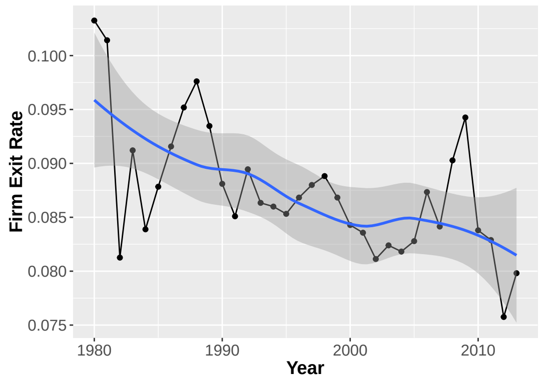


Figure 1.3: Firm exit rate

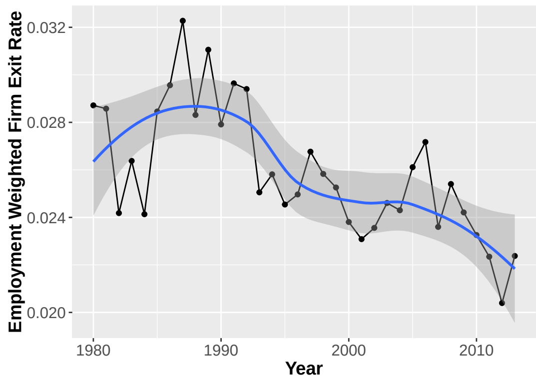


Figure 1.4: Employment weighted firm exit rate

1.2.3 Changes in the firm size distribution

Increasing average firm size Contemporaneous with the decline in firm creation and firm exit rates, the characteristics of incumbent firms also changed significantly. Figure 1.5 displays the average number of employees per firm in the US economy since the 1980s. Average firm size increased by around 15% from 20 employees in the early 1980s to around 23 employees since the 2000s. This increase is all the more notable as the reallocation of economic activity away from manufacturing towards services acted as a drag on the increase in the average number of jobs per firm. Holding sectoral shares constant at their early 1980s levels, the average number of employees per firm would have increased by an additional 5% to around 24 by 2014 as shown in Figure 1.6. Thus the increase in average firm size was caused by within sector increases in firm size, in particular in retail and services.[14]

To shed more light on the increase in average firm size, consider a partition of the firm population into different firm size groups, e.g., firms with less than 100, firms with 100 to 1000 and firms with more than 1000 employees. The average firm size in the economy is given by the sum of average firm sizes within each such group g weighted by the share of firms that belong to each group:

$$mean(size)_t = \sum_g share_{g,t} \times mean(size)_{g,t}$$

Using this formulation, allows me to decompose the change in average firm size into changes in the share of each group and changes in the average size of each group, as well as a covariance term. A decomposition of the change between the mid-1980s and the

[14]In manufacturing average firm size decreased. In particular, the size of large firms with more than 1000 employees decreased. One candidate explanation is increased offshoring of labor-intensive activities by large firms to low wage countries, e.g. in East Asia. Offshoring would reduce the measured size of large firms as the BDS data only covers US-based employment of firms.

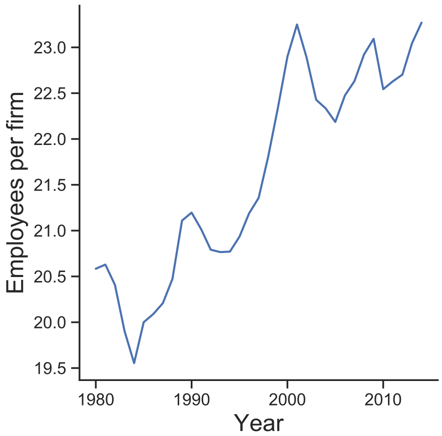


Figure 1.5: Average number of jobs per firm

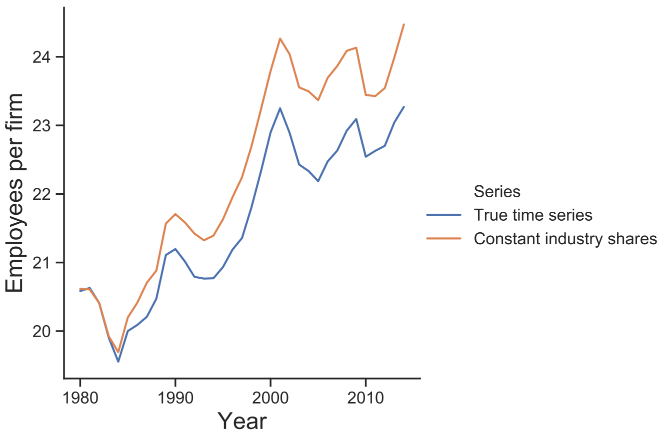


Figure 1.6: Average number of jobs per firm and counterfactual number of jobs per firm if industry shares had remained constant.

mid-2000s shows that both margins are essential. There has both been a shift towards larger firm size groups, in particular from small to medium-sized firms, and the average size of large firms has increased. In Table 1.1 I show the changes along both margins for the periods which are used in the calibration of the model in Section 1.4. There has been a marked increase in the average size of firms in the group with the largest firms as well as a shift in shares towards medium and large firms. In Table 1.2 I show the decomposition of the change in average firm size for these years. About 1/3 of the increase in average firm size is due to the shift towards larger firms, and 2/3 of the increase is due to the increasing size of large firms.[15]

Period	1987-1989		2005-2007	
	$share_{g,t}$	$size_{g,t}$	$share_{g,t}$	$size_{g,t}$
1 - 99 employees	98.12 %	8.48	97.95 %	8.38
100 - 999 employees	1.68 %	239.44	1.84 %	239.76
1000+ employees	0.194 %	4403.01	0.198 %	5002.31

Table 1.1: Firm size statistics in the late 1980s and mid 2000s

	1987 - 1989	2005 - 2007	Δ	Share
Average size	20.88	22.55	1.67	100 %
Share 1980s, size t		21.95	1.07	63.8 %
Share t, size 1980s		21.46	0.58	34.4 %

Table 1.2: Decomposition of change in average firm size between late 1980s and mid 2000s

[15]Depending on the specific years used for the comparison between 1/3 and 2/3 of the overall change in average firm size can be explained by each margin.

Changes in the characteristics of large firms In the previous paragraphs, the increasing size of large firms emerged as an important driver of the increasing average firm size since the 1980s. In the following paragraphs, I will analyze the margins along which the size of large firms increased. The BDS data contains information on the number of employees employed at firms in each size bracket and the number of establishments operated by firms in each size bracket. The number of employees employed at a firm can either change because of changes in the average number of employees employed at a firm's establishments or because of a change in the number of establishments operated by the firm. The aggregate time series are displayed in Figures 1.7 and 1.8. Since the mid-1980s, the average number of establishments operated by large firms has increased significantly from around ca. 70 to ca. 100. On the other hand, the average number of employees per establishment has decreased for large firms from around 65 to around 50. However, the increase in the average number of establishments operated by large firms dominated the decrease in the average number of employees per establishment, causing the number of employees per firm to increase by 15% since the late 1980s.

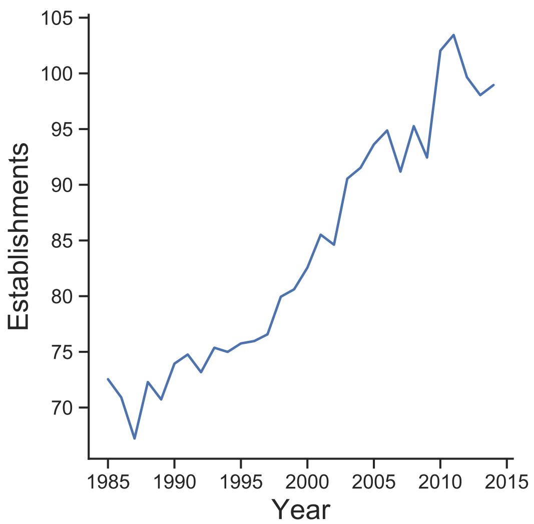


Figure 1.7: Average number of establishments per firm for large firms

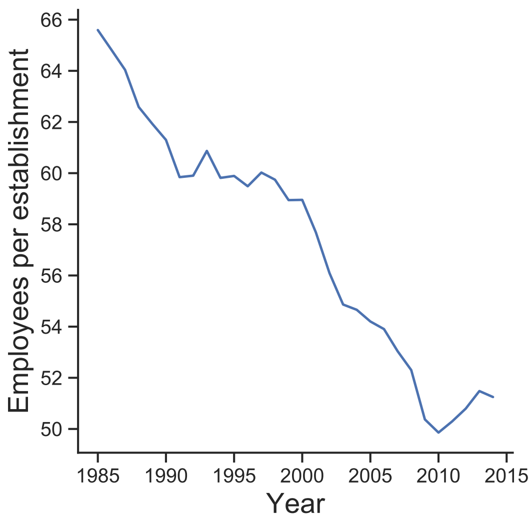


Figure 1.8: Average number of employees per establishment for large firms

The aggregate pattern by and large also holds within sectors. Table 1.3 displays the number of employees per firm, establishments per firm and employees per establishment for firms with more than 1000 employees disaggregated by SIC sector for the late 1980s and the 2010s.[16] First, the average size of large firms increased in most sectors. Second, the average number of establishments operated by large firms increased in every sector. The service sector stands out in particular, as large firms roughly doubled the average number of establishments under management. Finally, in most sectors the increase in

[16]The 2010s were chosen as the number of employees in the largest size bracket are censored in the BDS data for several sectors in the years 2005-2007.

the number of establishments under management was accompanied by a decrease in the average number of employees per establishment. The only sector that experienced a non-negligible increase in the number of employees per establishment is the retail sector, probably due to the increasing presence of big-box retail store chains.[17]

Sector	Employees per firm			Establishments per firm			Employees per establishment		
	'87-'89	'12-'14	Change	'87-'89	'12-'14	Change	'87-'89	'12-'14	Change
AGR	421	962	(+)	8.50	24.43	(+)	50.65	39.37	(-)
MIN	1015	1314	(+)	11.31	16.26	(+)	89.79	80.81	(-)
CON	953	1189	(+)	7.58	9.43	(+)	125.77	126.08	(+)
MAN	3373	2517	(-)	11.96	12.40	(+)	281.94	203.07	(-)
TCU	2953	2387	(-)	40.71	52.64	(+)	72.58	45.41	(-)
WHO	747	1163	(+)	22.21	33.69	(+)	33.62	34.52	(+)
RET	4184	5417	(+)	141.18	164.23	(+)	29.64	32.99	(+)
FIRE	2135	1744	(-)	63.88	75.06	(+)	33.51	23.25	(-)
SRV	2144	3171	(+)	18.95	37.50	(+)	113.22	84.36	(-)

Table 1.3: Characteristics of large firms with more than 1000 employees

To further illustrate the point that the main driver of increases in firm size come from within sectors, Table 1.4 displays four counterfactual scenarios for the change in the average number of employees between the late 1980s and the 2010s. Similar to the decomposition above, average firm size can be written as the weighted average of average firm sizes within each sector s as:

$$mean(size)_t = \sum_s share_{s,t} \times \frac{estabs_{s,t}}{firms_{s,t}} \times \frac{emp_{s,t}}{estabs_{s,t}}$$

where the average firm size within a sector is equal to the average number of establishments per firm in the sector times the number of employees per establishment in the sector. The first counterfactual considered in column (a) of Table 1.4 holds the share of firms belonging to each sector at their 1987-89 values and thus considers a scenario where only the average firm size within each sector had changed. If only the size of large firms within each sector had changed and the sectoral composition had remained constant, the average size of large firms had increased to 2674 employees per firm, accounting for 2/3 of the total increase. The counterfactual in column (b) considers the opposite case, i.e., how the average firm size of large firms would have changed if only the composition had changed, but the average size of large firms had remained at their late 1980s levels. The average size of large firms would have decreased slightly. Considering that on their own, changes in the average size within sectors and the sectoral composition account for around 60% of the change in average firm size leaves the remaining 40% to be accounted for by a covariance

[17] For a discussion of trends in the retail sector see Foster et al. (2006)

term. Indeed the service sector which expanded most, also experienced a large increase in average employment of large firms and the manufacturing sector which shrank most experienced a large decrease in average employment of large firms.

Columns (c) and (d) consider the contributions of changes in the number of employees per establishment and the number of establishments per firm holding the sectoral composition at the level of the late 1980s. Column (c) shows the average firm size which would result if only the number of employees per establishment had changed within each sector, while the number of establishments per firm and the sectoral composition were at their late 1980s level. The average firm size would have decreased significantly by 443 to 1990 employees per firm reiterating the finding from above that the establishment size of large firms declined significantly. Column (d) shows the average firm size which would have resulted if only the number of establishments per firm would have changed within each sector, while the number of employees per establishment and the sectoral composition had remained at their late 1980s level. The average firm size would have increased by 865 to 3299 employees per firm, more than double the actual increase.

Altogether Table 1.4 underlines the results obtained from the aggregate data. The average number of employees of large firms increased due to an increase in the size of large firms within sectors. The within sector increase in the size of large firms was the result of an increase in the average number of establishments operated by large firms which more than compensated for the decrease in the average number of employees per establishment.

Period	Av. size	(a)	(b)	(c)	(d)
1987-89	2433				
2012-14	2796	2674	2419	1990	3299
Δ	362	240	-15	-443	865
Share		66.3 %	- 4.1 %	-122.5 %	238.9 %

Table 1.4: Decomposition of the changing firm size of large firms

1.2.4 The presence of large firms across the US

Overall, large firms increased average employment and the average number of establishments per firm between the 1980s and the 2000s. The increase in firm size was especially pronounced in non-tradable sectors, while tradable sectors like mining and manufacturing show a different pattern. Because average establishment size declined significantly for large firms, a good measure for the importance of large firms in the economy is the share of establishments that are operated by large firms. In Figure 1.9 I plot for the aggregate economy the share of establishments that are operated by large firms. Since the 1980s this share has increased by around 6 percentage points or 60% from 10% to 16% reflecting the fact that large firms primarily expanded by increasing the number of establishments as documented in the previous subsection. Changes in the sectoral composition do not drive this fact. In Figure 1.10 I plot the actual share of establishments operated by large firms and a counterfactual share, holding the industry composition fixed at its early 1980s level. The counterfactual is computed by multiplying the actual share of establishments operated by large firms in each industry by the industry's share of all establishments in the 1980s. This exercise shows that the share of establishments operated by large firms would have been even higher, had the industry composition not changed. Thus the increase in the share of establishments that are operated by large firms is driven solely by within sector changes.

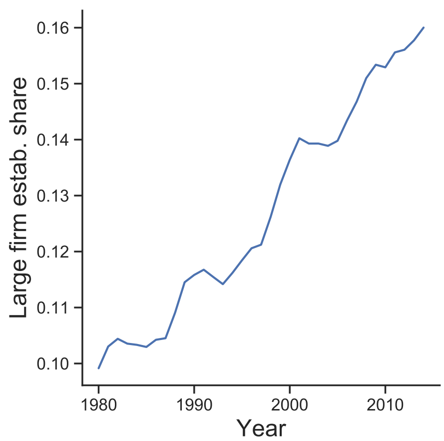


Figure 1.9: Share of establishments that are operated by large firms.

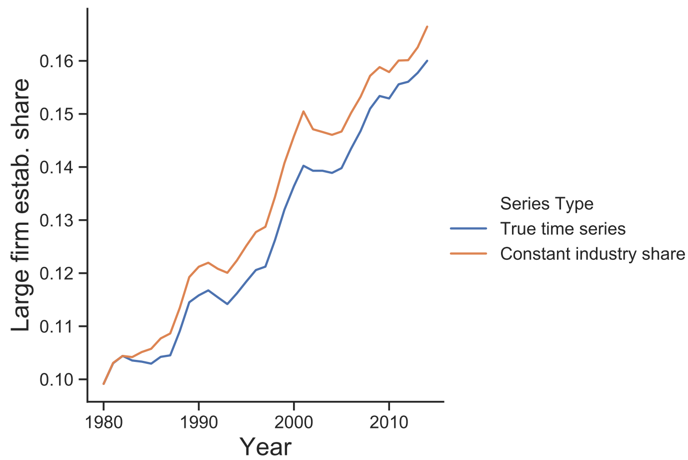


Figure 1.10: Share of establishments that are operated by large firms and counterfactual share of establishments that had been operated by large firms if industry shares had remained constant.

In Table 1.5 I further illustrate the pervasiveness of this phenomenon across sectors. While the share of large firms has increased slightly in some sectors but decreased in others, the share of establishments operated by large firms has increased in all sectors. In particular, sectors where establishments are likely to serve local demand, as output is relatively non-tradable, such as retail and services, experienced a strong increase in the share of establishments operated by large firms by around 50% and 100% respectively. The share of employees employed at large firms increased in most sectors but to a lesser extent than the share of establishments. This fact is driven by the decrease in establishment size which was observed in most sectors in the previous subsection. However, the share of establishments operated by large firms seems to be the most relevant measure for measuring the importance of large firms, in particular in the light of several labor-saving technologies with high fixed cost which have been implemented to a larger extent by large firms. Implementing new labor-saving technology would reduce the number of employees needed to deliver a service. On the other hand, lower costs of providing the service would make it profitable for large firms to expand into markets which were previously unprofitable, increasing the number of establishment managed by the firm. In the next section, I discuss evidence for these kinds of technologies.

Sector	Share of firms			Share of establishments			Share of employment		
	'87-'89	'12-'14	Change	'87-'89	'12-'14	Change	'87-'89	'12-'14	Change
AGR	0.15%	0.10%	(-)	1.3%	2.4%	(+)	8.7%	11.3%	(+)
MIN	1.52%	1.41%	(-)	13.8%	17.4%	(+)	53.1%	49.7%	(-)
CON	0.09%	0.12%	(+)	0.7%	1.1%	(+)	8.1%	12.9%	(+)
MAN	1.09%	1.07%	(-)	10.9%	11.0%	(+)	56.7%	51.2%	(-)
TCU	0.71%	0.90%	(+)	21.0%	29.9%	(+)	60.7%	62.8%	(+)
WHO	0.67%	0.69%	(+)	11.4%	16.5%	(+)	28.3%	37.7%	(+)
RET	0.19%	0.25%	(+)	18.7%	26.2%	(+)	40.6%	50.4%	(+)
FIRE	0.45%	0.54%	(+)	21.0%	25.0%	(+)	50.5%	52.8%	(+)
SRV	0.25%	0.29%	(+)	4.2%	9.1%	(+)	32.4%	43.4%	(+)

Table 1.5: Firm shares, establishment shares and employment shares of firms with more than 1000 employees in each sector.

The increasing role of large firms in the economy can also be seen geographically by analyzing changes in the presence of large firms across local markets. I identify local markets with the US Census' metropolitan statistical areas (MSAs), which are groupings of contiguous counties with a high population density and a high degree of economic integration. This analysis relates closely to contemporaneous work by Rossi-Hansberg et al. (2018) who analyze the competition effects of entry of large firms into new local markets.

I use two measures to analyze the presence of large firms in local markets. First, one can measure the percentage of large firms among all firms that are active in an MSA. A second measure is the percentage of all establishments in an MSA that belong to large firms. To get a feeling for these measures, consider the following example: Assume that, e.g. WalMart is not yet present in the New York metro area but is about to open a new store there. Then both the percentage of large firms active in the New York metro and the percentage of establishments operated by large firms in the New York metro would increase. If WalMart opens a second store in New York, only the percentage of establishments operated by large firms will increase, the percent of active firms that are large would not increase. Changes in the percentage of large firms among all firms in an MSA proxy for the expansion of large firms into new markets. Changes in the percentage of establishments in an MSA which are operated by large firms proxies for changes in the local market penetration by large firms. As explained above, the second measure seems to be the more relevant measure for the extent of competition from large firms in local markets.

In Figures 1.11 and 1.12 I plot the median of both measures across all MSAs over time. In 1980, in the median MSA, firms with nationally more than 1000 employees made up 5.7% of all active firms. This number increased to over 8.3% by 2013. At the same time, these firms also operated a larger fraction of establishments in the median MSA. The fraction of establishments operated by large firms increased from about 10% in 1980 to more than 17% in 2013.

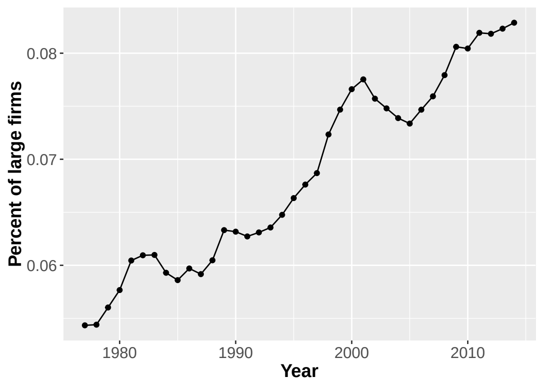


Figure 1.11: Percent of firms that are large, for the median MSA

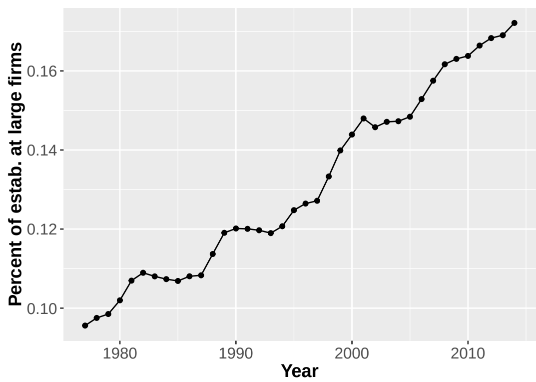


Figure 1.12: Percent of establishments operated by firms that are large, for the median MSA

The data on the firm size distribution show that large firms in the US increased in size, both in the average number of employees and in the average number of establishments. The size of large firms increased within almost all sectors, and the presence of large firms increased throughout the whole country. This evidence provided, together with the findings in Giroud (2013), Behrens and Sharunova (2015) and Hsieh and Rossi-Hansberg (2019) lend themselves to the following interpretation: Over the past 30 years firms have become more efficient at managing a spatially dispersed network of establishments. Large firms increased their presence across the US both along the extensive margin and the intensive margin. Along the extensive margin, firms entered new geographical markets. Along the intensive margin, large firms created new establishments in local markets in which they already maintained a presence.

1.2.5 The ICT revolution and its effects on large firms

The changes in the structure of large firms and their presence across the economy occurred at the same time as the widespread adoption of new information and communication technologies. These technologies have been extremely size biased, in particular in the early stages of adoption, as they are usually characterized by high sunk costs and high fixed costs at the firm level with low variable cost of spreading them across the firm. For example, enterprise software such as enterprise resource planning and customer relationship management systems can often cost several hundred million dollars to implement and requires IT personnel to maintain while it costs very little to add additional users to the system.[18]

Research by Tambe and Hitt (2012) finds that large firms receive a higher return on investments in IT than smaller firms. For the retail sector, Bagwell et al. (1997) and Holmes (2001) argue that larger firms benefit more from cost-saving technology than small firms due to the high sunk cost of investment. Furthermore, they argue that larger firms benefit more from new cost-saving technologies as their business models are more suitable for the use of particular information technologies such as barcodes and inventory management systems. Berger (2003) shows that adoption rates of new information technologies in the banking sector were much higher at larger institutions.

New information technologies also had an impact on the structure of industries. Brynjolfsson et al. (1994) show using the County Business Pattern data that industries which invested more in IT capital since the mid-1980s experienced a decrease in average establishment size. Using more recent data Saunders (2011) shows using the statistics of US businesses data that industries which in which firms invested more in information technologies experienced an increase in establishment openings of large firms relative to smaller firms and de novo establishments. Turning to specific industries, Berger (2003) shows that the adoption of information technologies in the banking sector led at the same time to fewer firms but more establishments in the industry, consistent with the findings above. For the trucking industry, Baker and Hubbard (2004) show that the adoption of onboard computers and tracking systems which significantly reduced monitoring costs led to consolidation in the industry with trucks being increasingly operated by large firms instead of single owner-operators.

The differential impact of information technology on large firms will be reflected in the model in two dimensions, reductions in establishment level fixed costs and increases in the firm-level returns to scale. Reductions in monitoring and coordination costs can be modeled as lower fixed costs at the establishment level, as less overhead is required per establishment to ensure that the employees who are performing the service or producing the product valued by the customer do so correctly. Furthermore, offshoring of back-office tasks can also be modeled as lower fixed costs at the establishment level, as the costs of performing these tasks drop when they are offshored to lower wage countries[19]. On the other hand lower communication, coordination and monitoring costs could also be

[18]See, e.g. Brynjolfsson and Hitt (2000) or National Academies of Sciences, Engineering and Medicine (2017). Note that recent innovations such as cloud computing could reverse these trends in the future as many technologies that once were available only at high fixed cost, nowadays become available through software as a service at relatively low fixed cost with companies only paying for the services that they need.

[19]Examples include offshoring of IT related overhead to India, call-centers to low-wage countries in general and also some HR related overhead to lower wage countries. For example, Lufthansa German Airlines offshored most routine HR tasks from Germany to Poland.

modeled as higher returns to scale. In firm hierarchy models such as Williamson (1967), Garicano (2000) and Garicano and Rossi-Hansberg (2006) reductions in communication costs between hierarchical layers would reduce communication frictions and result in increases in the returns to scale of firms.

1.3 An industry dynamics model with multiple industry segments

1.3.1 Consumers

Time is discrete and indexed by t. The consumer side of the economy consists of a representative household of size H_t which grows at rate g_p i.e. $H_{t+1} = H_t(1 + g_p)$. The household discounts future utility with the factor β and maximizes the population weighted discounted sum of per-period utility from per-capita consumption $c_t = C_t/H_t$:

$$U_0 = \sum_{t=0}^{\infty} \rho^t H_t \frac{c_t^{1-\theta} - 1}{1 - \theta} \tag{1.1}$$

Each household member inelastically supplies one unit of labor and earns income from wages w_t. Firms are owned by the households who receive per capita profits of π_t. The household also has access to a one period bond which earns a return r_{t+1}. Bond holdings are given by B_t. The consumption good is chosen as numeraire, so its price is normalized to 1. Together this yields the per-period budget constraint:

$$C_t + B_{t+1} = (1 + r_t)B_t + w_t H_t + \pi_t H_t$$

Normalizing to per capita terms yields:

$$c_t + (1 + g_p)b_{t+1} = (1 + r_t)b_t + w_t + \pi_t \tag{1.2}$$

Maximizing (1.1) subject to (1.2) yields the familiar Euler Equation:

$$\left(\frac{c_{t+1}}{c_t}\right)^{\theta} = (1 + r_{t+1})\rho \tag{1.3}$$

1.3.2 Incumbent firms

I consider an industry consisting of two separate segments, a mass-market (primary) segment P and a specialized segment S. The mass-market segment produces output C_P which can be thought of as very standardized services/goods in the industry. The specialized segment produces output C_S which can be thought of as specialty services which are customized to a consumer's specific needs. I assume that industry level output is a Cobb-Douglas aggregate of both types of goods, i.e. $C = C_P^{\beta} C_S^{1-\beta}$. Thus consumers spend a constant share β of their expenditures for the industry on the standardized output. Letting p_P the price for the mass-market output and p_S the price for the specialty output, the price of the industry composite is given by the Cobb-Douglas price index:

$$P_t = \left(\frac{p_{P,t}}{\beta}\right)^{\beta} \left(\frac{p_{S,t}}{1-\beta}\right)^{1-\beta}$$

I normalize this price for the composite output of the industry to one, which implies that relative prices satisfy:

$$p_{P,t} = (1-\beta)^{\frac{1-\beta}{\beta}} \beta p_{S,t}^{\frac{\beta-1}{\beta}} \tag{1.4}$$

To fix ideas about the differences between the mass-market and the specialty segment, think of beverage stores. The mass-market segment consists of stores selling all kinds of beverages like soft-drinks, common beer brands, and ordinary wine and spirits. The specialty segment consists e.g. of specialized wine and liquor stores which sell a hand-selected variety of wine and liquor and which offer personalized advice. Entry into the mass-market segment is relatively easy as it does not require to build up specialized knowledge. Furthermore, as the mass-market segment requires less specialized knowledge and is more standardized, business processes can be codified much easier, which makes the mass-market segment more suited for large firms.

Thus the specialized segment will be populated mainly by small firms which are run by owner-operators in single establishments as, e.g. the owner-operator cannot easily spread his specialized knowledge to other stores, and the provision of the service requires face to face contact between the customers and the owners. As the owner has sunk much time into building up his specialized knowledge, these firms will not be very volatile, and there will be relatively low entry and exit in this segment.

Two different sets of firms populate the mass-market segment. First, there will be many small-scale, i.e. single establishment, mom-and-pop stores which have not built up any specialized knowledge and enter and exit the industry frequently due to the low sunk cost. On the other hand, there will be a smaller number of very efficient large-scale businesses with standardized business processes which are replicated at a large number of establishments. These firms as they have sunk a substantial amount into the development of these business processes will be much more stable than the small-scale mom-and-pop stores in the primary segment.

Production of standardized goods/services

Production of the homogeneous standardized goods is performed by two different classes of firms, entrepreneurial firms, and mom-and-pop firms. Both types of firms operate with a decreasing returns to scale technology.

Entrepreneurial firms operate with higher establishment level fixed cost at higher returns to scale and can expand to multiple establishments. Mom-and-pop firms operate with lower establishment level fixed cost at lower returns to scale and can only operate a single establishment. Furthermore, I assume that mom-and-pop firms have a higher probability to exit the market each period.[20]

Entrepreneurial firms There is a mass M_t^E of entrepreneurial firms which are heterogeneous with respect to their productivity z. Firm-level productivity is assumed to be fixed over time. An entrepreneurial firm chooses the number of establishments it operates and the amount of labor to hire at each establishment. As establishments are symmetric, the same amount of labor will be hired at each establishment. An entrepreneurial firm with productivity z, operating e establishments and using l units of labor per establishment produces output according to the following decreasing returns to scale production function:

$$y(z, E) = z\left(el^{\alpha_E}\right)^{\eta_E} \tag{1.5}$$

where $\eta_E < 1$ parametrizes firm-level decreasing returns to scale and $\alpha_E < 1$ parametrizes establishment-level decreasing returns to scale. I assume that running an establishment requires c_e^E units of establishment-level overhead labor. Furthermore running an entrepreneurial firm requires c_f^E units of firm-level overhead labor.

Taking the market price p_P of the primary segment as given, the entrepreneurial firm maximizes:

$$\pi(z, E) = \max_{l,e} p_P z\left(el^{\alpha_E}\right)^{\eta_E} - w(c_e^E + l)e - wc_f \tag{1.6}$$

subject to the constraint that $e \geq 1$.[21] The resulting optimal number of establishments is given by:

$$e(z, E) = \max\left\{ z^{\frac{1}{1-\eta_E}} \left(\frac{p_P}{w}\right)^{\frac{1}{1-\eta_E}} \left(\frac{\eta_E(1-\alpha_E)}{c_e^E}\right)^{\frac{1-\alpha_E\eta_E}{1-\eta_E}} \left(\alpha_E\eta_E\right)^{\frac{\alpha_E\eta_E}{1-\eta_E}}, 1\right\} \tag{1.7}$$

[20] This assumption can be microfounded with a model of creative destruction. In such models, the exit probability depends negatively on the number of establishments at the firm level. As mom-and-pop firms have only one establishment, their exit rate will be higher in equilibrium.

[21] One could also assume that $e \in \mathbb{N}_+$. While this assumption would make some of the algebra more cumbersome, it does not change the qualitative results.

The optimal labor choice is given by:

$$l(z,E) = \min\left\{\frac{c_e^E \alpha_E}{(1-\alpha_E)}, z^{\frac{1}{1-\alpha_E\eta_E}}\left(\frac{\alpha_E\eta_E p_P}{w}\right)^{\frac{1}{1-\alpha_E\eta_E}}\right\} \tag{1.8}$$

Profits are then given by:

$$\pi(z,E) = \left[p_P z e(z,E)^{\eta_E - 1} l(z,E)^{\alpha_E\eta_E} - w(l(z,E) + c_e^E)\right] e(z,E) - wc_f^E \tag{1.9}$$

Note that the optimal unconstrained number of establishments is increasing in z. For small z the firm would want to choose $e(z) < 1$, it is however constrained to choose $e \geq 1$. Furthermore, the firm has to pay the firm level overhead costs. Together these assumptions imply that for small z firm level profits become negative, as variable profits no longer cover the fixed cost. The cutoff value below which profits are negative is given by:

$$z_E^* = \left(\frac{c_e^E + c_f^E}{1-\alpha_E\eta_E}\right)^{1-\alpha_E\eta_E}\left(\frac{1}{\alpha_E\eta_E}\right)^{\alpha_E\eta_E}\frac{w}{p_P} \tag{1.10}$$

Entrepreneurial firms exit the market with a constant exit probability δ_E. The present value of an entrepreneurial firm with productivity z is thus:

$$V_t(z,E) = \sum_{\tau=0}^{\infty}\left(\frac{1-\delta_E}{1+r}\right)^{\tau}\pi_{t+\tau}(z,E) \tag{1.11}$$

Mom-and-pop firms Mom-and-pop firms also operate a decreasing returns to scale technology using only labor at a single establishment. These firms are also heterogeneous with respect to their productivity z and produce according to the technology:

$$y(z,M) = zl^{\alpha_M} \tag{1.12}$$

with $\alpha_M < 1$ parametrizing decreasing returns to scale in production. The firm has overhead of c_e^M units of labor for its establishment and thus solves:

$$\pi(z,M) = \max_l p_P z l^{\alpha_M} - w(l + c_e^M) \tag{1.13}$$

The optimal amount of labor is then given by:

$$l(z,M) = \left(\frac{z\alpha_M p_P}{w}\right)^{\frac{1}{1-\alpha_M}} \tag{1.14}$$

and profits are given by:

$$\pi(z,M) = (1-\alpha_M)\left(p_P z\left(\frac{\alpha_M}{w}\right)^{\alpha_M}\right)^{\frac{1}{1-\alpha_M}} - wc_e^M \tag{1.15}$$

which yields the following cutoff value below which profits are negative and no production will take place:

$$z_M^* = \left(\frac{c_e^M}{1-\alpha_M}\right)^{1-\alpha_M} \left(\frac{1}{\alpha_M}\right)^{\alpha_M} \frac{w}{p_P} \tag{1.16}$$

Mom and pop firms exit the market with a constant exit probability $\delta_M > \delta_E$ such that the value of a mom and pop firm with productivity z is given by:

$$V_t(z, M) = \sum_{\tau=0}^{\infty} \left(\frac{1-\delta_M}{1+r}\right)^{\tau} \pi_{t+\tau}(z, M) \tag{1.17}$$

Production of customized goods/services

Production of customized goods or services is very similar to the mom-and-pop firms in the standardized segment of the industry. I assume that the only difference is that because the owners built up specific knowledge which is hard to replicate, the per period exit probability δ_S is lower than for the mom-and-pop stores in the standardized segment. With p_S the price of the specialized industry, the labor demand of a specialized firm with productivity z is given by:

$$l(z, S) = \left(\frac{z\alpha_M p_S}{w}\right)^{\frac{1}{1-\alpha_M}} \tag{1.18}$$

and profits are given by:

$$\pi(z, S) = (1-\alpha_M)\left(p_S z \left(\frac{\alpha_M}{w}\right)^{\alpha_M}\right)^{\frac{1}{1-\alpha_M}} - wc_e^M \tag{1.19}$$

The cutoff below which specialty firms are not profitable is given by:

$$z_S^* = \left(\frac{c_e^M}{1-\alpha_M}\right)^{1-\alpha_M} \left(\frac{1}{\alpha_M}\right)^{\alpha_M} \frac{w}{p_S} \tag{1.20}$$

The value of a specialized firm is given by:

$$V_t(z, S) = \sum_{\tau=0}^{\infty} \left(\frac{1-\delta_M}{1+r}\right)^{\tau} \pi_{t+\tau}(z, S) \tag{1.21}$$

1.3.3 Entry

I assume an entry process similar to Sedláček and Sterk (2017), who draw on Saint-Paul (2002). In order to enter, potential firm owners have to come up with an idea. Coming up with a business idea takes time and is thus associated with a cost of ϕ units of labor. However, there is a limited number of viable business ideas per capita ψ_j for each firm type $j \in \{E, M, S\}$ per period. Coming up with ideas within firm classes is undirected. Coordination failure between potential firm owners can lead to multiple potential owners coming up with the same idea. Only one of these potential owners will successfully start a business. Specifically, I assume that if there are A_j startup attempts, the number of unique business ideas M_E^j for each firm type is given by the following Cobb-Douglas matching function:

$$M_E^j = A_j^\gamma (\psi_j H_t)^{1-\gamma} \tag{1.22}$$

The probability of finding a unique idea is then given by:

$$P_E^j = M_E^j / A_j = \left(\frac{\psi_j H_t}{A_j} \right)^{1-\gamma} \tag{1.23}$$

I assume that each unique idea is associated with a productivity z, drawn from a firm type specific productivity distribution $G_j(z)$ with pdf $g_j(z)$. As the search for an idea is assumed to be undirected, all unique ideas are i.i.d. draws from this distribution. The free entry assumption then implies that the cost of coming up with an idea is equal to the expected benefit of coming up with an idea which is the probability of having a unique idea times the expected value of an idea for a firm of type j:

$$w\phi = P_j \mathbb{E}_z[V(j,z)] \tag{1.24}$$

The free entry condition implies that increases in the expected value of starting a firm have to be balanced by decreases in the probability of successfully finding a unique idea. By (1.23) a decrease in this probability is triggered by an increase in the number of startup attempts which by (1.22) increases the mass of unique ideas. Thus increases in the value of successfully starting a firm cause an increase in entry.

1.3.4 The firm distribution

Denote by $\mu_t(z,j)$ the mass of active firms of type j with productivity z. This mass changes due to exit and entry. The law of motion is thus given by

$$\mu_t(z,j) = \underbrace{(1-\delta_j)\mu_{t-1}(z,j)}_{\text{Surviving firms}} + \underbrace{M_t^j g_j(z) \mathbb{1}(z \geq z_j^*)}_{\text{Entering Firms}} \tag{1.25}$$

The per capita law of motion is obtained by dividing through with H_t and given by:

$$\frac{\mu_t(z,j)}{H_t} =: \nu_t(z,j) = \frac{1-\delta_j}{1+g_p} \nu_{t-1}(z,j) + m_t^j g_j(z) \mathbb{1}(z \geq z_j^*) \tag{1.26}$$

Where $\nu_t(z,j)$ is the per capita mass of firms of type (z,j) and $m_{E,t}^j$ is the mass of unique ideas for firms in firm class j per capita. In a stationary equilibrium with a stationary per capita firm type distribution $\nu(z_j) = \nu_t(z,j) = \nu_{t-1}(z,j)$ this allows to solve (1.26) for $\nu(z,j)$:

$$\nu(z,j) = \frac{1+g_p}{g_p+\delta_j} m^j g_j(z) \mathbb{1}(z \geq z_j^*) \tag{1.27}$$

1.3.5 Aggregation

Total supply of primary goods is given by:

$$Y_{S,t}^P = \underbrace{\int z l_t(z,E)^{\alpha_E \eta_E} e_t(z,E)^{\eta_E} d\mu_t(z,E)}_{\text{Entrepreneurial firms' supply}} + \underbrace{\int z^{\frac{1}{1-\alpha_M}} \left(\frac{\alpha_M p_P}{w}\right)^{\frac{\alpha_M}{1-\alpha_M}} d\mu_t(z,M)}_{\text{Mom-and-pop firms' supply}} \tag{1.28}$$

Total demand for primary goods is given by

$$Y_{D,t}^P = \beta C_t \tag{1.29}$$

Total supply of specialty goods is given by:

$$Y_{S,t}^S = \int z^{\frac{1}{1-\alpha_M}} \left(\frac{\alpha_M p_S}{w}\right)^{\frac{\alpha_M}{1-\alpha_M}} d\mu_t(z,S) \tag{1.30}$$

Total demand for specialty goods is given by:

$$Y_{D,t}^S = (1-\beta)C_t \tag{1.31}$$

Total profits accruing to households are given by:

$$\Pi_t = \int \pi_t(z,E) d\mu_t(z,E) + \int \pi_t(z,M) d\mu_t(z,M) + \int \pi_t(z,S) d\mu_t(z,S) \tag{1.32}$$

Total labor demand is given by:

$$\begin{aligned} L_t^D &= \int e_t(z,M)(c_e^E + l_t(z,M)) + c_f^E d\mu_t(z,E) + \int c_e^M + l_t(z,M) d\mu_t(z,M) \\ &\quad + \int c_e^M + l_t(z,S) d\mu_t(z,S) + \sum_{j \in \{E,M,S\}} \phi A_t^j \end{aligned} \tag{1.33}$$

The corresponding per capita measures $y_{S,t}^P, y_{D,t}^P, y_{S,t}^S, y_{D,t}^S, \pi_t, l_t^D$ are derived by dividing through H_t and are given in terms of other per-capita quantities (e.g. the per-capita distribution ν_t) as:

Primary supply per capita:

$$y_{S,t}^P = \int z l_t(z,E)^{\alpha_E \eta_E} e_t(z,E)^{\eta_E} d\nu_t(z,E) + \int z^{\frac{1}{1-\alpha_M}} \left(\frac{\alpha_M p_P}{w}\right)^{\frac{\alpha_M}{1-\alpha_M}} d\nu_t(z,M) \tag{1.34}$$

Primary demand per capita:

$$y^P_{S,t} = \beta c_t \tag{1.35}$$

Specialty supply per capita:

$$y^S_{S,t} = \int z^{\frac{1}{1-\alpha_M}} \left(\frac{\alpha_M p_S}{w}\right)^{\frac{\alpha_M}{1-\alpha_M}} d\nu_t(z,S) \tag{1.36}$$

Specialty demand per capita:

$$y^P_{S,t} = (1-\beta)c_t \tag{1.37}$$

Profits per capita:

$$\pi_t = \int \pi_t(z,E)d\nu_t(z,E) + \int \pi_t(z,M)d\nu_t(z,M) + \int \pi_t(z,S)d\nu_t(z,S) \tag{1.38}$$

Labor demand per capita

$$\begin{aligned} l^D_t &= \int e_t(z,M)(c^E_e + l_t(z,M)) + c^E_f d\nu_t(z,E) + \int c^M_e + l_t(z,M)d\nu_t(z,M) \\ &\quad + \int c^M_e + l_t(z,S)d\nu_t(z,S) + \sum_{j\in\{E,M,S\}} \phi a^j_t \end{aligned} \tag{1.39}$$

where a^j_t is the pre capita mass of startup attempts for firm type j.

1.3.6 General equilibrium and balanced growth

Definition 1.3.1 (Sequential equilibrium). *A sequential equilibrium is a sequence of wages w_t, prices $p_{P,t}$ and $p_{S,t}$, interest rates r_t and per capita profits π_t, per capita consumption c_t and bond holdings b_t, as well as for each firm type $j \in \{E,M,S\}$ firm level labor demand decisions $l_t(z,j)$, establishment number decisions $e_t(z,j)$, productivity cutoffs $z^*_{j,t}$ a measure of startup attempts a^j_t per capita and a per capita measure of firms over productivities and firm types $\nu_t(z,j)$ such that:*

1. *Consumption and bond-holdings c_t, b_{t+1} maximize utility* (1.1) *subject to the sequence of budget constraints* (1.2) *given $\{w_t,\pi_t,r_t\}^\infty_{t=0}$.*
2. *Given $\{w_t,r_t,p_{P,t},p_{S,t}\}^\infty_{t=0}$ firms labor demand $l_t(z,j)$ is given by* (1.8), (1.14), (1.18), *the number of establishments for entrepreneurial firms $e_t(z,E)$ satisfies* (1.7) *and for the other firms $e_t(z,j)=1$, firm level profits $\pi_t(z)$ are given by* (1.9), (1.15), (1.19) *and firms operate only if their value functions* (1.11), (1.17), (1.21) *are positive.*
3. *Given $\{w_t,r_t,p_{P,t},p_{S,t}\}^\infty_{t=0}$, the mass of entry attempts a^j_t, satisfies the free entry condition* (1.24) *in every period.*

4. The goods markets clear ($y^S_{P,t} = y^D_{P,t}$, $y^S_{S,t} = y^D_{S,t}$), the labor market clears ($l^D_t = 1$), the bond market clears ($b_{t+1} = 0$) and per capita profits π_t are consistent with (1.38).

5. Relative prices of the industry segments satisfy (1.4).

6. The distribution of firms ν_t satisfies the law of motion (1.26).

Given this definition of a sequential equilibrium, I can now derive the balanced growth path. The only source of growth in this economy is the growing population. Thus a balanced growth path is defined by normalizing all aggregate quantities by the size of the population. Along the balanced growth path, prices and firm-level policies will be time-invariant as will be all normalized quantities, which yields the following definition of a balanced growth path:

Definition 1.3.2 (Balanced growth path)**.** *A balanced growth path consists of a wage w, prices p_P and p_S, an interest rate r, per capita consumption c, bond holdings b and per capita profits π as well as for each firm type $j \in \{E, M, S\}$, firm level labor demand decisions $l(z, j)$, establishment number decisions $e(z, j)$ and productivity cutoffs z^*_j, a measure of startup attempts per capita a_j and a per capita measure of firms over productivities and firm types $\nu(z, j)$, which are consistent with a sequential equilibrium with:*

1. $(c_t, b_{t+1}) = (c, b)$ *for all* t

2. $(l_t(z, j), e_t(z, j), z^*_{j,t}) = (l(z), e(z, j), z^*_j)$ *for all* t

3. $(w_t, \pi_t, p_{P,t}, p_{S,t}, r_t) = (w, \pi, p_P, p_S, r)$ *for all* t

4. $(a^j_t, \nu_t(z, j)) = (a^j, \nu(z, j))$ *for all* t

1.4 Quantitative analysis

I calibrate the model to match several characteristics of US firm dynamics such as exit rates and the firm size distribution in the period 1987-1989, the first years for which data on firms until age 10 is available. The model period is assumed to be one year.

1.4.1 Numerical solution and calibration

Productivity distributions I assume that the productivity distribution of specialized firms and mom-and-pop firms follows the same truncated Pareto distribution with shape parameter b_M, location parameter 1 and upper truncation $x_{M,max}$. Furthermore, I assume that entrepreneurial firms draw their productivity from a Pareto distribution with shape parameter b_E and location parameter $x_{E,min}$.

Parameters Some parameters of the model are calibrated externally. I assume that the elasticity of intertemporal substitution is equal to 1, yielding log-utility. The discount factor ρ is assumed to be equal to 0.96 yielding an interest rate of roughly 4%. Following Sedlacek and Sterk (2017), I assume that the elasticity parameter γ of the entry cost function is 0.3. I normalize the amount of labor required to generate a business idea to 1. The remaining 17 parameters that have to be calibrated internally are:

1. Firm level parameters
   (a) Returns to scale parameters $\alpha_E, \eta_E, \alpha_M$
   (b) Overhead labor requirements c_f^E, c_e^E, c_e^M
   (c) Exit rates $\delta_E, \delta_M, \delta_S$
2. Productivity parameters
   (a) Shape b_E and location $x_{E,min}$ of the E type productivity distribution
   (b) Shape b_M and upper truncation $x_{M,max}$ of the M and S type distribution
3. Entry parameters
   (a) Mass of business opportunities ψ_j for each type
4. Specialty vs. primary share β

I calibrate the model by minimizing the sum of squared distances between various target statistics and the model implied counterparts. The targets describe firm and industry dynamics in the late 1980s and are calculated from the business dynamics statistics. Firm size and establishment size statistics are targeted to calibrate the returns to scale and overhead labor requirements. Firm and employment shares are targeted to calibrate the productivity distribution parameters and the mass of business opportunities. Exit rates are targeted to calibrate the exit rates and also in order to calibrate the relative mass of specialty vs. mom-and-pop firms which are identified by the difference in exit rates of young vs. old small firms. Thus exit rates also help to identify the expenditure share on specialty goods.

1. Firm level parameters
   (a) Returns to scale parameters and overhead labor requirements: workers per firm and establishments per firm for small and large firms
   (b) Exit rates: exit rates for large firms δ_E, exit rates for small young firms δ_M, exit rates for old small firms δ_S
2. Productivity parameters

(a) Firm size distribution / share of firms with more than X employees

3. Entry parameters

   (a) Entry rate

   (b) Shares of firms at different ages

4. Specialty vs. primary share

   (a) Relative entry and exit rate for young and old small firms

Moment	Model	Data
Startup rate	11.90 %	11.95 %
Average firmsize	20.86	20.88
Startup size	7.27	6.39
Conditional firm size		
<20emp	3.92	5.00
<100emp	7.16	8.49
>1000emp	4363	4401
Conditional Establishment size		
>1000emp	62.05	61.88
Conditional firm shares		
<50emp	95.92 %	95.83 %
>1000emp	0.177 %	0.193 %
<5yr	49.56 %	47.68 %
>10yr	31.98 %	32.88 %
Conditional employment share		
>1000emp	37.02 %	40.85 %
Conditional exit rates		
<=5yr	13.13 %	13.33 %
>10yr	6.47 %	6.32 %
<50emp, <=5yr	13.25 %	13.38 %
<50emp, >10yr	7.08 %	6.8 %
>100emp, >10yr	0.45 %	0.29 %

Table 1.6: Model fit

Discussion of calibration results Table 1.6 shows the data moments along with the model implied moments. Table 1.7 displays the resulting parameter estimates. Overall the match between the model and the data moments is quite good. The moments with the largest discrepancy between model and data are the startup size and the conditional firm sizes for small firms. The startup size is too high, reflecting the fact that individual firms in this model do not have a firm life cycle. Due to the abstraction from adjustment frictions, firms do not start small and grow over time but instead remain at the same

Parameter	Variable	Value
Primary segment expenditure share	β	0.903
Returns to scale		
E-firms estab. level	α_E	0.668
E-firms firm level	η_E	0.675
M,S-firms	α_M	0.63
Overhead		
E-firms firm level	c_f^E	40.1
E-firms estab level	c_e^E	20.6
M-firms estab level	c_e^M	0.97
S-firms estab level	c_e^S	0.92
Exit probabilities		
E-firms	δ_E	0.45%
M-firms	δ_M	17.95 %
S-firms	δ_S	5.5 %
Business opportunities		
E-firms	ψ_E	0.15
M-firms	ψ_M	90.27
S-firms	ψ_S	10.00
Productivity distribution parameters		
E-firm shape parameter	b_E	3.955
M,S-firm shape parameter	b_M	5.35
E-firm scale parameter	$x_{min,E}$	5.05
M,S-firm maximum productivity level	$x_{max,M}$	16.8

Table 1.7: Internally calibrated parameters

constant size over their lifetime. The firm size of small firms is too low, possibly reflecting the fact that the Pareto distribution which was assumed not only for entrepreneurial but also for mom-and-pop and specialty firms is not a perfect description of the underlying productivity distribution of small firms. The exit rates for different groups of firms are matched quite well. These exit rates inform the parameters governing the exit rates of the different firm types. The exit probabilities imply that small mass-market firms exit much more frequently than small specialty firms. Multi-establishment mass-market firms have the lowest exit rates. In the next subsection, I will relate these differences in exit rates to differences in entry rates in the steady state.

1.4.2 How changes in the parameters affect the firm creation rate

The firm creation rate is defined as the mass of successful entrants divided by the total number of firms in the economy. Let m the mass of firms per capita, m_j the mass of firms of type $j \in \{E, M, S\}$ per capita, m^e the mass of entrants per capita and m_j^e the mass of entrants of type j per capita. For each class of firm, the firm creation rate can be derived by rearranging (1.27) as:

$$fcr_j = \frac{m_j^e}{m_j} = \frac{g_p + \delta_j}{1 + g_p}$$

The aggregate firm creation rate can then be written as the weighted sum of firm creation rates of the individual firm classes:

$$fcr = \frac{m^e}{m} = \frac{\sum_j m_j^e}{m} = \sum_j \frac{m_j}{m} \frac{m_j^e}{m_j} = \sum_j \frac{m_j}{m} fcr_j$$

This formulation illustrates the two channels through which the aggregate firm creation rate can change: (1) through changes in the firm class specific firm creation rates fcr_j and (2) through changes in the firm class composition m_j/m. Changes in the population growth rate affect the aggregate firm creation rate both through a change in the firm class specific firm creation rates and changes in the composition. Changes in returns to scale or firm class specific cost parameters only affect the firm creation rate through changes in the composition of firms. The effects of parameter changes can be decomposed into three distinct channels: (a) a *partial equilibrium* channel, holding prices and wages fixed, (b) a *labor market* channel, where in addition to the partial equilibrium the wage adjusts while the output prices remain fixed and (c) an *output market* channel, where in addition to the partial equilibrium the output prices adjust while the wage remains fixed. I illustrate the importance of these channels separately for a change in population growth, entrepreneurial firms' returns to scale and entrepreneurial firms' establishment level fixed cost. Quantitatively I change each parameter in line with the recalibration in section 1.4.4.

Population growth Changes in population growth affect both the firm class specific firm creation rates and the composition of firms in equilibrium. In Table 1.8 I illustrate the channels through which a change in population growth from 1.8% to 1.25% affects the firm creation rate.

	Baseline	Partial Eq.	Labor Market	Output Market	Total
p_S	0.8239	0.8239	0.8239	0.8486	0.8486
p_P	0.7176	0.7176	0.7176	0.7153	0.7153
w	1.5404	1.5404	1.5954	1.5404	1.5954
fcr_E	2.21 %	1.68%	1.68%	1.68%	1.68%
fcr_M	19.40%	18.96%	18.96%	18.96%	18.96%
fcr_S	6.78%	6.27%	6.27%	6.27%	6.27%
m_E	0.0016	0.0022	0.0019	0.0021	0.0017
m_M	0.0172	0.0176	0.0144	0.0172	0.0131
m_S	0.0220	0.0238	0.0196	0.0291	0.0228
m_E/m	4.01 %	4.96%	5.20%	4.39%	4.62 %
m_M/m	42.07%	40.37%	40.26%	35.56%	34.79%
m_S/m	53.91%	54.66%	54.53%	60.05%	60.59%
fcr	11.90%	11.17 %	11.14%	10.58%	10.47%

Table 1.8: Channels through which a change in population growth (g_p) affects the firm creation rate.

The change in population growth directly reduces the firm class specific firm creation rates by about 0.5 percentage points. Furthermore, the drop in population growth affects the per-capita dilution rate of firms with lower exit rates more, inducing a shift in the firm composition towards specialized and entrepreneurial firms. This compositional effect reduces the firm creation rate by another 0.2 percentage points.

Absent general equilibrium effects that induce a further decrease in the firm creation rate, the decrease in population growth induces an increase in labor demand per capita. This effect can be seen by the increase in the mass of firms per capita. In order to clear the labor market, the wage increases by 3.5% which reduces the mass of firms per capita. However, the decrease is almost uniform across firm types, leaving the composition of firms and the aggregate firm creation rate almost unchanged.

The compositional effect highlighted above is also behind the additional decrease in firm creation rates through output market competition. Within the mass market, there is a shift towards entrepreneurial firms which increases the average efficiency of firms on the mass-market. Thus the relative price of mass-market goods falls inducing a shift of activity towards the specialized market. As the average entry rate of mass-market firms is higher than that of specialized firms, the aggregate firm creation rate falls by 0.6 percentage points as a result of this compositional shift.

Thus overall about 50% of the decline in startup rates following a decline in population growth is due to direct effects occurring in partial equilibrium, 45% is due to price effects resulting from changes in output market competition, and the rest is due to labor market effects and non-linearities.

Firm level returns to scale Changes in firm-level returns to scale leave the firm class specific firm creation rates unaffected an only affect the aggregate firm creation rate through changes in the firm composition. In Table 1.9 I document the channels through which a change in the firm-level returns to scale of entrepreneurial firms from 0.675 to 0.6866 affects the firm creation rate.

	Baseline	Partial Eq.	Labor Market	Output Market	Total
p_S	0.8239	0.8239	0.8239	0.8774	0.8774
p_P	0.7176	0.7176	0.7176	0.7128	0.7128
w	1.5404	1.5404	1.6393	1.5404	1.6393
m_E	0.0016	0.0021	0.0015	0.0020	0.0014
m_M	0.0172	0.0172	0.0106	0.0163	0.0101
m_S	0.0220	0.0220	0.0137	0.0357	0.0221
m_E/m	4.01 %	5.08%	5.73%	3.74%	4.23 %
m_M/m	42.07%	41.60%	41.21%	30.19%	30.01%
m_S/m	53.91%	53.32%	53.06%	66.07%	65.76%
fcr	11.90%	11.79 %	11.72%	10.42%	10.37%

Table 1.9: Channels through which a change in E-firms' returns to scale (η_E) affects the firm creation rate.

In partial equilibrium, the increase in returns to scale positively affects the profits of entrepreneurial firms. As profits increase, more entrepreneurial firms enter inducing an increase in the mass of entrepreneurial firms by about 1/3. As entrepreneurial firms only account for a small share of all firms in the economy, the share of entrepreneurial firms increases only by one percentage point. Accordingly, the direct impact on firm creation rates is quite low, leading to a 0.1 percentage point decrease.

The increase in entrepreneurial firms' returns to scale induces an increase in labor demand. In order for the labor market to clear, wages increase by 6.2%. As the elasticity of small firms' profits with respect to wages is slightly higher, the increase in wages leads to an additional shift toward entrepreneurial firms. Thus firm creation rates drop by an additional 0.07 percentage points.

The main channel through which an increase in large firms' returns to scale acts is the output market channel. The increase in entrepreneurial firms' returns to scale and the resulting shift towards entrepreneurial firms leads to an increase in output in the mass-market. In order for the market to clear the relative price of output in the mass-market decreases or vice-versa the relative price of specialty goods increases. The increase in the price of specialty output induces entry of specialty firms. The composition of firms shifts towards specialty firms which leads to a decrease in the aggregate entry rate. Overall this channel leads to a decrease in the aggregate entry rate by about 1.4 percentage points.

Overall slightly more than 90% of the decline in the aggregate firm creation rate is explained by the output market channel. Increased competition from entrepreneurial firms on the mass market crowds out entry by small mom-and-pop firm and the composition shifts towards specialty firms with relatively lower entry and exit rates. The rest is explained both by direct effects and the labor market channel.

Establishment level fixed cost Analogously to changes in firm-level returns to scale, changes in establishment level fixed cost leave the firm class specific firm creation rates unaffected and only affect the aggregate firm creation rate through changes in the firm composition. In Table 1.10 I document the channels through which a decrease in the establishment level fixed costs of entrepreneurial firms from 20.6 to 17.76 affects the firm creation rate. The relative contributions of the different channels and their interpretation are quite similar to the changes in entrepreneurial firms' returns to scale, while the magnitudes are weaker.

	Baseline	Partial Eq.	Labor Market	Output Market	Total
p_S	0.8239	0.8239	0.8239	0.8456	0.8456
p_P	0.7176	0.7176	0.7176	0.7156	0.7156
w	1.5404	1.5404	1.5802	1.5404	1.5802
m_E	0.0016	0.0019	0.0016	0.0019	0.0016
m_M	0.0172	0.0172	0.0141	0.0168	0.0138
m_S	0.0220	0.0220	0.0181	0.0269	0.0221
m_E/m	4.01 %	4.67%	4.91%	4.15%	4.36 %
m_M/m	42.07%	41.78%	41.67%	36.89%	36.79%
m_S/m	53.91%	53.55%	53.42%	58.96%	58.84%
fcr	11.90%	11.83 %	11.81%	11.24%	11.22%

Table 1.10: Channels through which a change in E-firms' establishment level fixed costs (c_e^E) affects the firm creation rate.

The decrease in entrepreneurial firms' establishment level fixed cost increases the value of starting an entrepreneurial firm. Entry for this firm class increases, increasing the share of entrepreneurial firms in partial equilibrium. The shifting composition results in a decrease in aggregate firm creation rates by 0.07 percentage points. Labor demand of entrepreneurial firms expands slightly and thus in order to clear the labor market the wage increases by around 2%. This increase results in a small additional shift towards entrepreneurial firms, reducing the entry rate by 0.02 percentage points.

As above, the main channel through which a decrease in entrepreneurial firms' establishment level fixed cost acts is the output market channel. As production by entrepreneurial firms expands, the relative price of mass-market goods decreases, resulting in less entry by small mass-market firms. The corresponding increase in the price of specialty goods leads to increasing entry of specialty firms. Overall the composition shifts away from small mass-market firms towards specialty firms, resulting in a decrease in the aggregate firm creation rate by 0.6 percentage points. Of the total decrease in the firm creation rate by 0.7 percentage points, around 85% is due to the output market channel, and 15% is due to direct effects and the labor market channel.

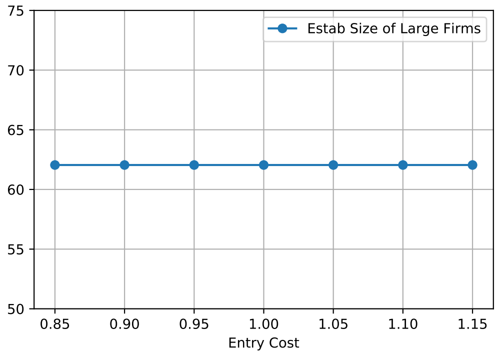
1.4.3 Quantitative analysis of potential sources of the startup decline

Firm entry can decline due to different underlying sources each of which has implications for other dimensions of industry dynamics and the firm size distribution. In this subsection I will use these additional dimensions to shed more light on the relative importance of changes in entry cost, establishment-level overhead costs of large firms, firm-level returns to scale and population growth in explaining the decline in firm entry.

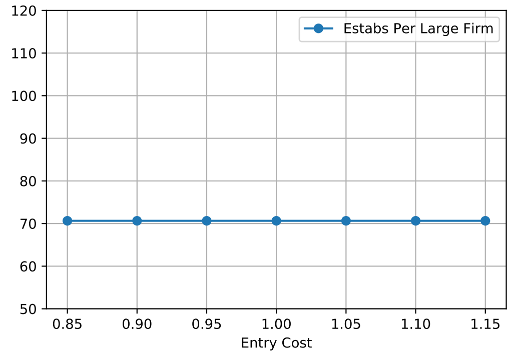
For each of these potential underlying sources, I calculate the new balanced growth path equilibrium while holding all other parameters at their equilibrium values calibrated to the 1987-89 economy. For entry costs (Figure 1.13) and establishment level overhead costs (Figure 1.14) I vary the values between 85% and 115% of their calibrated values. For large firms' firm-level returns to scale (Figure 1.15) I vary the values between 95% and 105% of the calibrated values. Finally population growth (Figure 1.16) is varied between 0% and 4% per year.

The figures show four key statistics that have changed significantly since the late 1980s. Panel (a) displays the average establishment size of large firms (with more than 1000 employees) which has dropped from 62 in the late 1980s to 54 in the early 2000s. Panel (b) displays the average number of establishments of large firms which has increased from around 70 to 93. Panel (c) displays the average firm size of all firms which has increased from 20.9 to 22.7. Finally, panel (d) displays the entry rate which has dropped from ca. 12% to 10.4%.

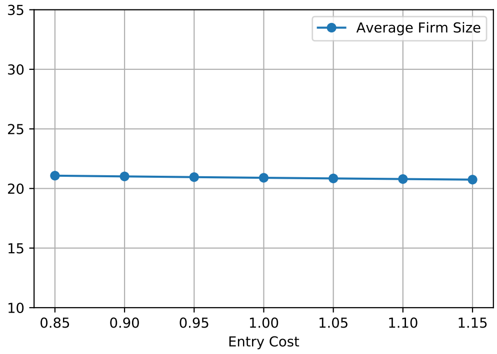
Entry cost (ϕ) Changes in entry cost affect the economy through general equilibrium effects. In order for the free entry condition to hold, an increase in entry costs ϕ have to be counterbalanced by either an increase in the probability of successfully finding a



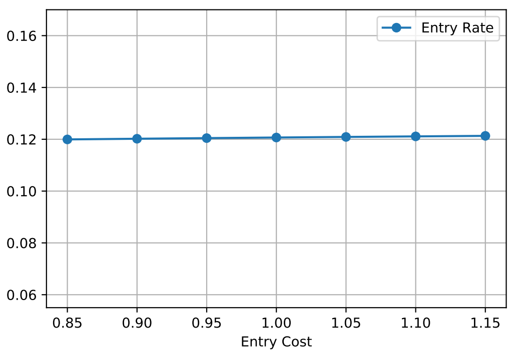
(a) establishment size of large firms



(b) no. of establishments of large firms



(c) average firm size



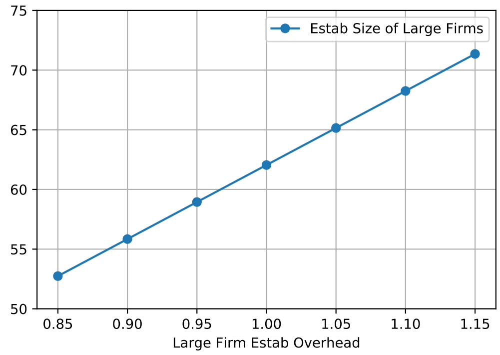
(d) entry rate

Figure 1.13: Implications of changes in entry costs for firm dynamics

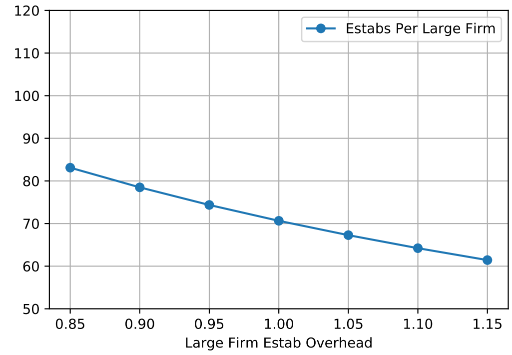
unique business idea or an increase in the value of a unique business idea. Although both margins are active, the effects of changes in entry costs are negligible compared to the effects of changes in the other parameters. The characteristics of large firms do not change at all, while the average firm size and the entry rate only change a little. Increases in entry costs lead to a shift away from entrepreneurial firms towards mom-and-pop and specialized firms. The compositional shift induces a decrease in the average firm size and a slight increase in the entry rate.

Large firms' establishment-level overhead (c_e^E) In the BDS data I showed that the average establishment size of large firms has decreased from ca. 62 to 53 between the late 1980s and the mid-2000s. In the model, the establishment size is tightly linked to the establishment level overhead costs.

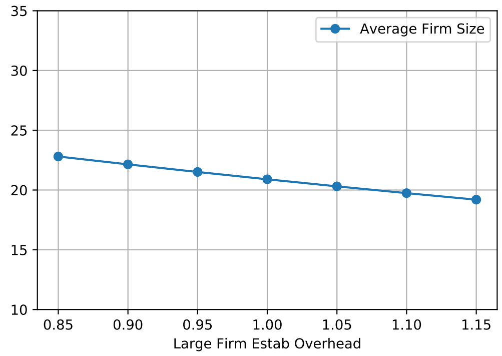
For fixed establishment-level decreasing returns to scale α_E, the optimal establishment size is linear in the overhead labor. Thus decreases in the overhead labor requirement directly lead to a decrease in the optimal establishment size of large firms. Furthermore, a decrease in the establishment level overhead costs leads to an increase in the number of establishments per firm. Intuitively, the cost of increasing production through a larger



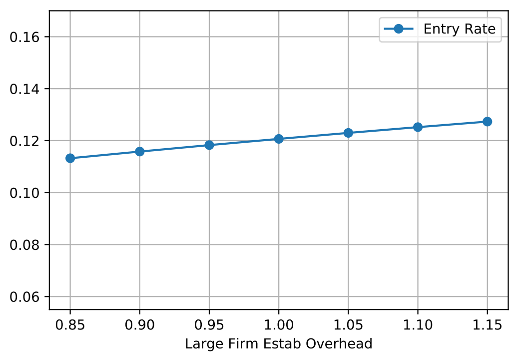
(a) establishment size



(b) no. of establishments of large firms



(c) firm size



(d) entry rate

Figure 1.14: Implications of changes in E-firms establishment-level overhead for firm dynamics

number of establishments decreases relative to the cost of increasing production through a larger number of workers per establishment. Thus in order to minimize costs, firms reduce the number of workers per establishment and increase the number of establishments per firm.

The reduction of establishment-level overhead costs reduces the costs of entrepreneurial firms relative to mom-and-pop and specialized firms. The reduction in costs implies an increase in the relative profitability of entrepreneurial type firms relative to the other firms. Due to free entry, an increase in the relative profitability induces an increase in the number of entry attempts for entrepreneurial firms which induces a shift of the composition of active firms in the economy. This shift is exacerbated through the increased competition from entrepreneurial type firms. The inflow of entrepreneurial firms leads to an increase in competition in the primary segment. Thus the value of mom-and-pop firms decreases relative to the value of specialized firms leading to an additional shift towards among the small firms towards the specialized firms. The shift towards entrepreneurial type firms leads to an increase in the average firm size. Both the shift towards less volatile entrepreneurial firms and specialized firms induces a decrease in entry rates.

Large firms' firm-level returns to scale (η_E) In the data, the size of large firms in terms of employment has also increased since the late 1980s from around 4400 to 5000 employees per firm. In the model, for a fixed distribution of firm-level productivities, the size of large firms is closely linked to the firm level returns to scale parameter of entrepreneurial firms η_E.

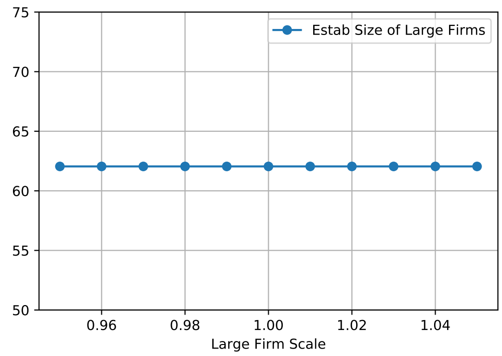
An increase in the returns to scale leaves the optimal establishment size unaffected as the relative costs of increasing production through an increase in the number of workers per establishment vs. increasing the number of establishments remains unaffected. The optimal size of entrepreneurial firms increases when the firm level returns to scale increase, as the marginal benefit of increasing the size of the firm increases. The increase in the optimal size of firms leads to an increase in the average size of large firms through an increase in the number of establishment these firms operate.

Similar to the reduction in establishment-level overhead costs, higher firm-level returns to scale for entrepreneurial firms induce an increase in the profitability of entrepreneurial firms relative to mom-and-pop and specialized firms. Furthermore, the increased production of entrepreneurial firms leads to more competition in the primary segment and thus an additional reduction in the profitability of mom-and-pop firms. Overall the shift towards entrepreneurial type firms and the increase in the size of these firms leads to a substantial increase in the average firm size when returns to scale increase. The compositional shift away from mom-and-pop firms towards the less volatile entrepreneurial and specialized firms leads to a decrease in the entry rate.

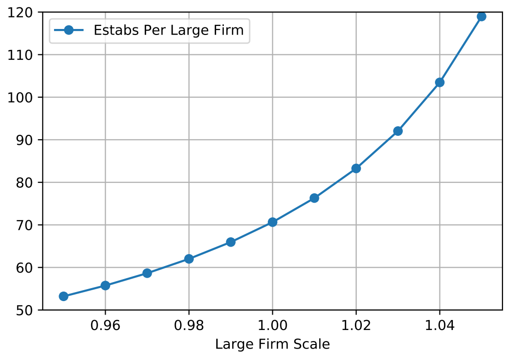
Population growth rate (g_p) Karahan et al. (2018) have shown that a significant fraction of the decline in startup rates since the 1980s can be explained by changes in the population growth rate. On the balanced growth path of my model the population growth also directly affects the entry rate.

The characteristics of large firms remain unaffected by changes in the population growth rate. Neither the optimal establishment size nor the optimal number of establishments per large firm change when the population growth rate changes.

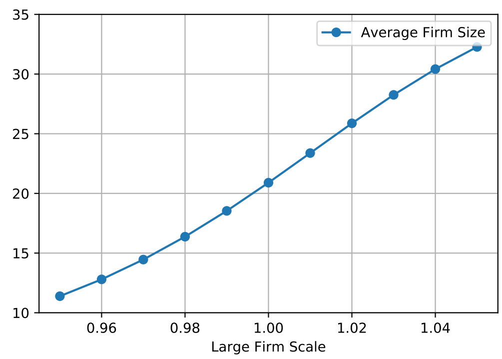
Average firm size and the entry rate, on the other hand, are strongly affected by changes in the population growth rate. Both changes are due to a composition effect. The equilibrium per capita mass of entrepreneurial firms reacts much stronger to changes in population growth rates than of mom-and-pop or specialized firms. The reason is that the elasticity of the mass of firms of a particular type with respect to the population growth rate is negatively related to a type's idiosyncratic exit probability. Thus firm types with lower exit probability react stronger to changes in the population growth rate. The higher elasticity is due to the fact that the share of exits from the per-capita firm distribution that is due to population growth is much higher for firms with low idiosyncratic exit rates.



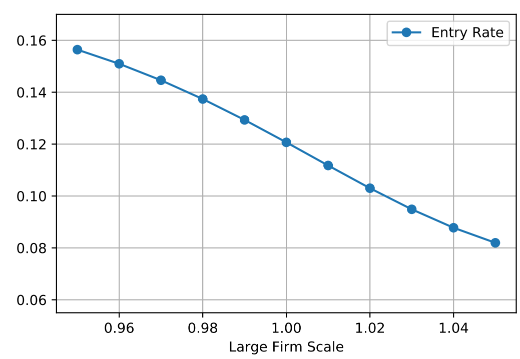
(a) establishment size



(b) no. of establishments of large firms

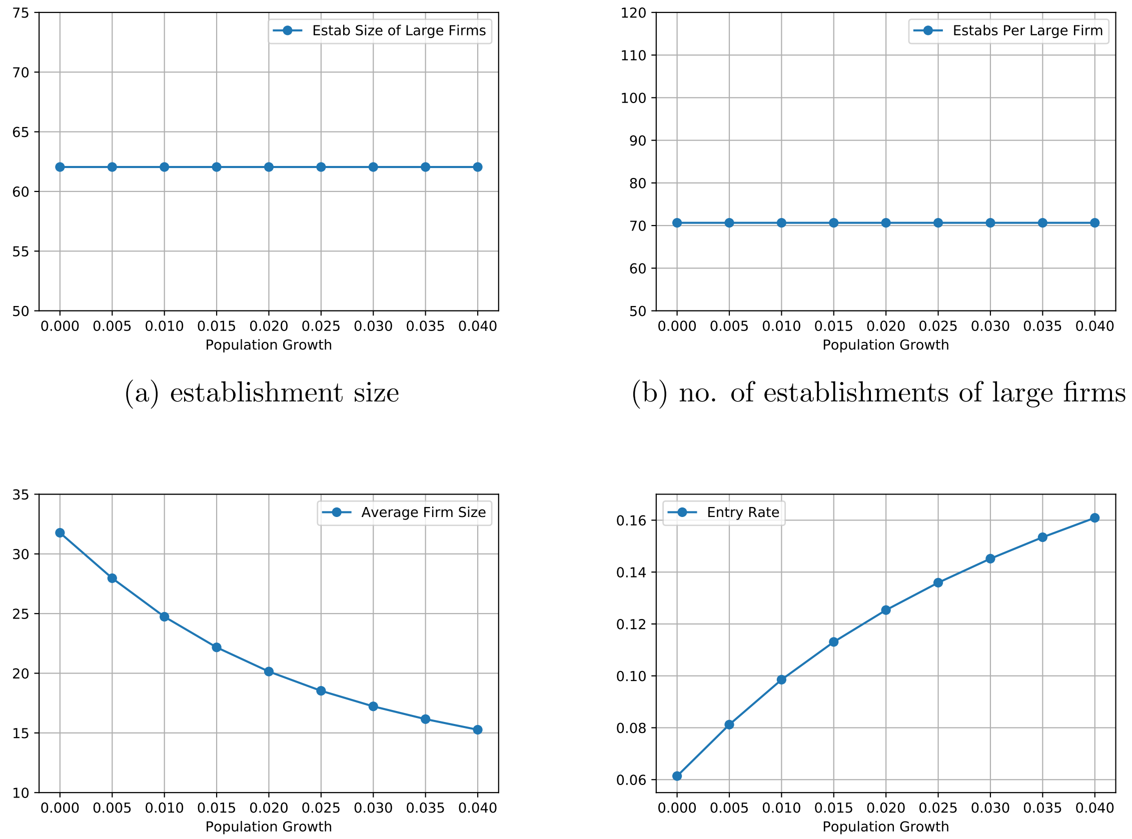


(c) firm size



(d) entry rate

Figure 1.15: Implications of changes in E-firms firm-level returns to scale for firm dynamics



(a) establishment size (b) no. of establishments of large firms

(c) firm size (d) entry rate

Figure 1.16: Implications of changes in population growth for firm dynamics

This compositional shift toward entrepreneurial and specialized firms leads to an increase in the average firm size and a decrease in entry rates.

1.4.4 A decomposition of the underlying sources of the startup decline

The previous section has shown that changes in the establishment level overhead costs of entrepreneurial firms, the returns to scale of entrepreneurial firms and the population growth rate all lead to quantitatively important declines in the firm entry rate. However, all three factors have different implications for the size of establishments, the number of establishments per firm and the number of employees per firm.

In this section, I first investigate the implications of changes in each parameter separately, when it matches its key identifying statistic. Specifically, I change the establishment level overhead costs such that the average establishment size for large firms is equal to 53.5, the average value in 2005-07, holding all other parameters fixed. I also change the firm level returns to scale such that the number of employees at large firms is equal to 5008, the average value in 2005-07, holding all other parameters fixed. Finally, I change the population growth rate to 1.25, the average value in 2005-07, holding all other parameters fixed.

As a second exercise, I recalibrate the model, holding all parameters except for the entrepreneurial firms' overhead costs, the entrepreneurial firms' firm-level returns to scale, the population growth rate and the mass of business opportunities fixed. I then decompose the change in the startup rate into contributions of changes in the overhead costs, the firm level returns to scale and the population growth rate, sequentially changing each parameter from its calibrated 1987-89 value to its recalibrated 2005-2007 value.

Matching key statistics A first approach to understand the contribution of changes in parameters to explain the decline in entry rates is to set the parameters such that the key statistic that identifies the parameter is precisely at its level in the 2005-07 period, holding all other parameters at their calibrated 1987-89 values. The results of these counterfactuals are displayed in Table 1.11. The bold number in each column is the key statistic that is matched by the change in the parameter.

The establishment-level overhead costs for entrepreneurial firms are identified by the average establishment size of large firms. In order to match the decline in the average establishment size from 61.88 to 53.49, the establishment level overhead costs have to decline by 13.8 % from 20.6 to 17.76. The implied entry rate is 11.22 %, accounting for about 45 % of the actual decline. The counterfactual also performs well along other dimensions, almost exactly matching the average firm size and the firm shares conditional on size. Along most other dimensions, the changes are qualitatively consistent with the observed changes but don't quantitatively match up. In particular, the conditional firm size of large firms is invariant to changes in the overhead costs of establishments.

The firm-level returns to scale of large firms are identified by the average firm size of large firms conditional on the shape parameter of the productivity distribution function. In order to match the increase in the firm size of large firms from 4403 to 5008, the firm level returns to scale have to increase by 1.72 % from 0.675 to 0.6866. The implied entry rate is 10.37 %, accounting for all of the actual decline. The counterfactual also performs well along other dimensions, almost exactly matching the conditional employment share of large firms and the conditional firm shares. Along the other dimensions, the implied changes are qualitatively consistent with the observed changes but don't match up quantitatively. In particular, the establishment size of large establishments is invariant to changes in the returns to scale of large firms.

Finally, we can investigate the implied changes of a decrease in the population growth rate to 1.25 %, its value in 2005-2007. The implied entry rate is 10.47 % accounting for all of the actual decline in the entry rate. The counterfactual economy almost exactly matches the firm shares conditional on firm size. Along other dimensions, the counterfactual performs well qualitatively but does not quantitatively match the observed changes. However, both the establishment size of large firms and the average employment of large firms remain unaffected by changes in the population growth rate.

Overall I find that changing each parameter separately can explain a significant part of the decline in startup rates. However, the implications for some key statistics such as the size of large firms or the establishment size of large firms are counterfactual and would require changes in all three parameters. So in the next section, I recalibrate all parameters jointly.

Moment	Data	$c_e^E = 17.76$	$\eta_E = 0.6866$	$g_p = 0.0125$
Population growth rate	1.25 %	1.8 %	1.8 %	**1.25** %
Startup rate	10.44 %	11.22 %	10.37 %	10.47 %
Average firmsize	22.68	22.69	25.35	23.38
Startup size	6.10	7.81	8.66	7.24
Conditional firm size				
<20emp	4.82	3.90	3.89	3.90
<100emp	8.42	4.11	7.18	7.61
>1000emp	5008	4363	**5008**	4363
Conditional Establishment size				
>1000emp	53.49	**53.49**	62.05	62.05
Conditional firm shares				
<50emp	95.53 %	95.46 %	95.69 %	95.31 %
>1000emp	0.200 %	0.198 %	0.225 %	0.204 %
<5yr	39.84 %	47.5 %	44.91 %	44.89 %
>10yr	42.38 %	34.04 %	36.49 %	36.99 %
Conditional employment share				
>1000emp	44.03 %	38.08 %	44.52 %	38.04 %
Conditional exit rates				
<=5yr	13.08 %	12.44 %	11.45 %	12.36 %
>10yr	5.44 %	6.14 %	5.83 %	6.01 %
<50emp, <=5yr	13.13 %	12.58 %	11.58 %	12.48 %
<50emp, >10yr	5.82 %	6.74 %	6.34 %	6.61 %
>100emp, >10yr	0.28 %	0.45 %	0.45 %	0.45 %
Welfare change				
		2.78 %	6.33 %	7.06%

Table 1.11: Counterfactual model fit

Recalibration and decomposition The joint changes in the firm size and the establishment size of large firms point towards joint changes in establishment level overhead costs and the returns to scale of large firms. However, jointly changing the returns to scale, establishment-level overhead costs and the population growth rate to the values found in the previous section has strong counterfactual implications for most model moments. These results from the joint change are reported in Table 1.12. Changing overhead costs, returns to scale of large firms and the population growth rate in order to match the observed establishment size, firm size and population growth rate in the mid-2000s would imply a very strong shift away from mom-and-pop firms towards entrepreneurial and specialty firms. This shift results in a reduction of the startup rate by 3.4 percentage points to 8.54% per year compared to an observed reduction by 1.5 percentage points. The implied average firm size is much too high at 28.99 employees per firm compared to an observed size of 22.68. Similarly, the implied startup size is much too high. Implied exit rates for young firms are much too low, reflecting the shift away from the more volatile mom-and-pop firms towards more stable specialty and entrepreneurial firms.

The fact that the shift towards entrepreneurial and specialty firms is too strong could be due to the model assumption that we are in a steady state, while in reality, the economy is on a transition path towards a new steady state. Along the transition, the economy will exhibit more mom-and-pop firms than in steady state, lowering the average firm size relative to the steady state bringing the model closer to the observed data. Another explanation is that other parameters might have changed. For example, Bloom et al. (2017) argue that new ideas are getting harder to find. In the model, new business ideas would be harder to find if the available mass of business ideas decreases. In particular, if the mass of available business ideas for entrepreneurial businesses which are closer to the concept of an idea in Bloom et al. (2017) would decrease, the shift towards these businesses would be smaller. To take this possibility into account, I recalibrate the model allowing for changes in the population growth rate, the returns to scale of entrepreneurial firms, the establishment and firm level overhead costs of entrepreneurial firms and the mass of business opportunities for all firms. The resulting parameter values are displayed in Table 1.13. The parameters which were changed already in the previous section remain stable in the new calibration, reflecting the fact that they are tightly linked to the key statistics targeted above. Consistent with the narrative proposed above, the mass of business opportunities for entrepreneurial firm drops significantly. This shift changes the composition of newly started businesses towards small mass-market and specialized firms which have higher entry and exit rates. As a result, the overall shift towards large mass-market firms is dampened compared to the scenario in which the mass of business opportunities is held fixed.

The joint recalibration allows me to decompose the total change in the firm creation rate into the contributions of the different channels considered in the literature. This decomposition is illustrated in Figure 1.17. The firm creation rate in the model calibrated to the 1987-89 period was 11.90%. Starting from this calibration, I first adjust the startup costs to the recalibrated values. As discussed above the startup costs increase primarily for the entrepreneurial firms, inducing a compositional shift towards small mass-market and specialty firms. This compositional shift leads to an increase in the firm creation rate

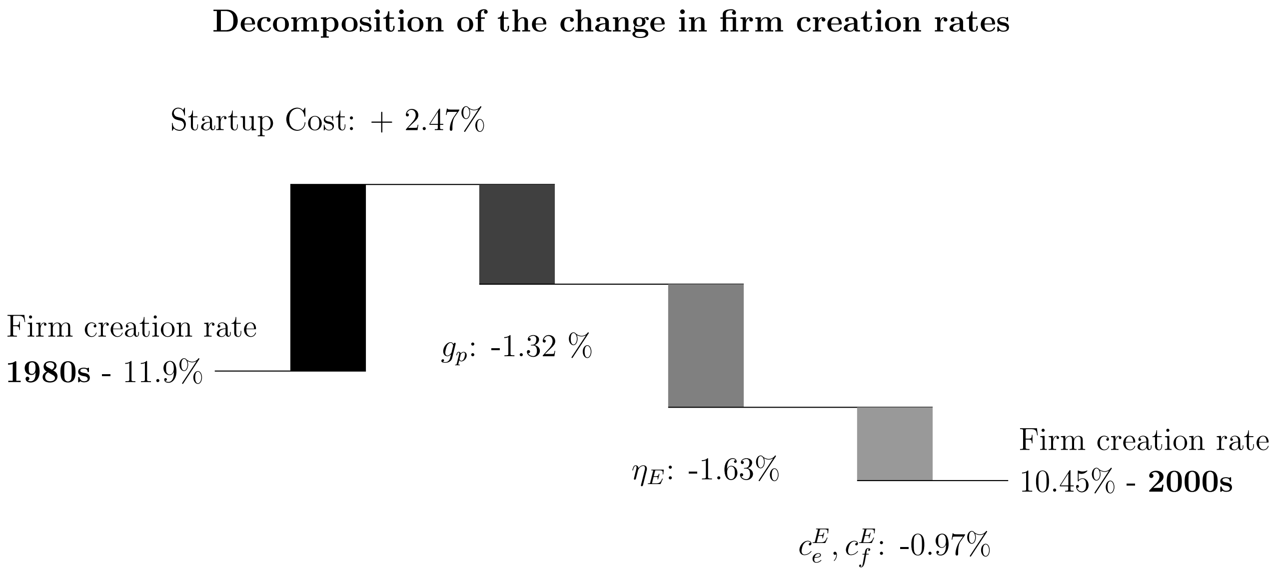


Figure 1.17: Decomposition of the change in the firm creation rate between the calibration to 1987-89 and the recalibration to 2005-2007. The calibrated change in the startup costs which showed an increase in startup costs particularly for large firms would have led to an increase in the startup rate by 2.47%. The observed change in population growth (g_p) decreased the startup rate by 1.32%. The increase in large firms returns to scale (η_E) led to an additional decrease by 1.63% and the decrease in large firms' establishment and firm level fixed cost (c_e^E, c_f^E) caused an additional decrease by 0.97%. In total the firm creation rate decreased by 1.45% from 11.9% in 1987-89 to 10.45% in 2005-07.

by 2.47%. Additionally considering the observed decrease in population growth (g_p) leads to a decrease in labor supply growth and puts upward pressure on wages. These effects cause a decrease in the firm creation rate by 1.32%. When I also change the returns to scale for entrepreneurial firms (η_E) to their recalibrated value, the firm creation rate further decreases by 1.64%. The change in establishment and firm-level fixed cost (c_e^E, c_f^E) to the recalibrated value induces an additional decrease by 0.97%.

The improvements in the large firms' technologies reduce the entry rate through two effects. On the one hand large firms' profitability increases which induces an increase in the entrant composition toward large firms. On the other hand, the increase in labor demand and output produced by large firms puts upward pressure on wages and downward pressure on prices in the mass-market segment. These effects reduce the relative profitability of small mass-market firms and induce a shift in the entrant composition away from small mass-market firms towards small niche-market firms. Overall the shifts in the firm composition away from small mass-market firms which have high entry and exit rates towards large mass-market and small niche-market firms with lower entry and exit rates significantly reduce the aggregate entry and exit rate.

The decomposition suggests that relative to the firm creation rate which we would have observed if only the mass of business opportunities for entrepreneurial firms had dropped, 35% of the decline in startup rates is due to lower population growth, another 40% is due to an increase in large firms' returns to scale and 25% is due to decreases in the establishment level fixed costs of large firms. Importantly, while the decline in the population growth rate can explain a large part of the decrease in firm creation rates, the changes in other firm characteristics suggest an essential role for the other channels discussed in the literature, which are dismissed in Karahan et al. (2018).

Moment	Data	Joint Change	Recalibration
Population growth rate	1.25 %	**1.25** %	1.25 %
Startup rate	10.44 %	8.54 %	10.45 %
Average firmsize	22.68	28.99	23.50
Startup size	6.10	9.00	7.26
Conditional firm size			
<20emp	4.82	3.87	3.90
<100emp	8.42	4.06	4.10
>1000emp	5008	**5008**	5059
Conditional Establishment size			
>1000emp	53.49	**53.49**	53.42
Conditional firm shares			
<50emp	95.53 %	95.08 %	95.46 %
>1000emp	0.200 %	0.266 %	0.197 %
<5yr	39.84 %	38.86 %	44.8 %
>10yr	42.38 %	43.00 %	37.00 %
Conditional employment share			
>1000emp	44.03 %	45.99 %	42.45 %
Conditional exit rates			
<=5yr	13.08 %	9.86 %	12.32 %
>10yr	5.44 %	5.35 %	6.02 %
<50emp, <=5yr	13.13 %	9.97 %	12.43 %
<50emp, >10yr	5.82 %	5.83 %	6.59 %
>100emp, >10yr	0.28 %	0.45 %	0.45 %
Welfare Change			
		17.3%	7.3%

Table 1.12: Recalibrated model fit

Welfare implications The question with which I started the analysis was whether the observed decline in startup rates is worrisome or not. With the model calibrated to the late 1980s and the mid 2000s economy at hand, we can analyze the welfare consequences of the change in startup rates by comparing the aggregate consumption levels of each economy. While the improvements in large firms' technologies should have led to an increase in aggregate consumption, the reduction in business opportunities which increased the effective cost of starting a business should have led to a decrease in aggregate consumption. The welfare change relative to the economy calibrated to the early 1980s is

Parameter	Baseline	Joint Change	Recalibration
Jointly changed			
g_p	0.018	0.0125	0.0125
η_E	0.675	0.6866	0.6870
c_e^E	20.6	17.76	17.74
Additionally recalibrated			
c_f^E	40.1	40.1	35.5
ψ_E	0.15	0.15	0.077
ψ_M	90.27	90.27	87.2
ψ_S	10.00	10.00	10.41

Table 1.13: Parameter changes

shown in the last line of Table 1.12. If there had been no change in the availability of business opportunities and only population growth, the returns to scale of large firms and the establishment level fixed cost of large firms had changed as in the column denoted by "Joint Change", agents in the economy would have enjoyed 17.3% higher consumption levels of the composite good. However, once the decreasing availability of entrepreneurial business opportunities is taken into account as in the column denoted by "Recalibration", consumption levels only increase by 7.3%. Nonetheless, the results from the recalibration suggest that the observed decline in firm creation rates has not been associated with increasing inefficiency and lower levels of consumption.

1.5 The increasing presence of large firms and the decline in startup rates across metropolitan areas in the US

The improved ability of large firms to managing a spatially dispersed set of establishments and the resulting increased penetration of MSAs by large firms could plausibly reduce the entry of new firms. As large firms expand, competition increases both in local product markets and local labor markets. As a result, firm profits conditional on productivity decline. As the returns to starting a firm of a given level of productivity decline, the least productive firms will no longer be started and the firm startup rate should decline.

A measure of the degree of competition by large firms is the percentage of establishments in an MSA that belong to large firms. A change in this measure captures both an expansion along the intensive margin of large firms that already are present in the MSA as well as along the extensive margin of large firms that open their first establishment in an MSA. To test the hypothesis developed above, I can test whether MSAs that experienced a more substantial increase in the fraction of establishments belonging to large firms also saw a more sizable decrease in the startup rate. I compare changes between the late 1980s (averages over 1987-1989) and the mid-2000s (averages over 2004-2006) just before the great recession. Both periods correspond to similar phases of the business cycle to get rid of business cycle effects on the startup rate. In Figure 1.18 I plot the correlation across MSAs between the change in the fraction of establishments operated by large firms and the change in the firm startup rate. The unconditional correlation is statistically significantly negative. A 1 percentage point increase in the fraction of establishments operated by large firms is associated with a 0.2 percentage point decrease in the startup rate.

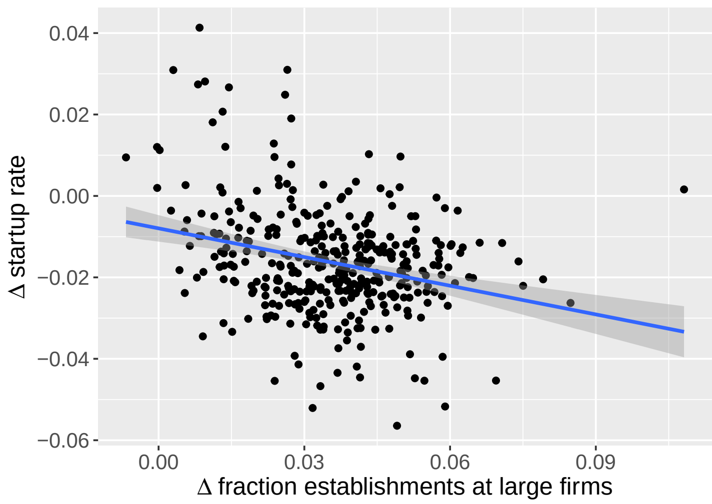


Figure 1.18: Correlation between the change in the fraction of establishments operated by large firms and the change in startup rates between 1987-89 and 2004-06 across MSAs

Hathaway and Litan (2014c) find that in their analysis, the correlation between their measure of a change in business consolidation and the change in startup rates is much lower after conditioning on additional explanatory variables. Additional variables they include are the startup rate at the beginning of the sample, in order to control for mean reversion, and changes in the population growth rate. Conditioning on changes in the population growth rate also shows that the effect of the expansion of large firms on firm startup rates is orthogonal to the theory of Karahan et al. (2018) who propose a decline in the population growth rate as the source of the decline in firm startup rates. In Table 1.14 I present results from a differences in differences regression of changes in the startup rate on the initial startup rate, changes in the population growth rate and changes in the fraction of establishments operated by large firms. The conditional correlation between changes in the fraction of establishments operated by large firms and changes in the firm startup rate remains statistically significantly negative and is quantitatively close to the unconditional coefficient.

	Dependent variable:
	Δ firm formation rate
firm formation rate 1987-89	0.068*
	(0.040)
Δ frac. estab. at large firms	−0.225***
	(0.041)
Δ population growth rate	0.098***
	(0.015)
Constant	−0.016***
	(0.005)
Observations	364
R^2	0.190
Adjusted R^2	0.183
Residual Std. Error	0.012 (df = 360)
F Statistic	28.189*** (df = 3; 360)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1.14: Regression of long run differences (1987-89) to (2004-06) in the firm formation rate on long run differences in the fraction of establishments at large firms and the population growth rates across MSAs.

1.6 Conclusion

In this paper, I study the decline in startup rates and link it to the simultaneous increase in the size and geographic dispersion of large firms in the US economy over the past 30 years. I argue that an important part of the decline in firm creation rates can be explained as an efficient response to size biased technological change that benefited large firms relative to small firms.

The empirical contribution of this paper is to document three stylized facts on changes in the characteristics of incumbent firms that occurred contemporaneously to the decline in firm formation rates. First, I show that the 15% increase in the average number of employees per firm since 1980 was caused both by a shift towards larger firms and an increase in the size of larger firms. Second, I analyze the increasing size of larger firms and show that large firms expanded mainly by increasing the number of establishments operated by the average while reducing the average number of employees per establishment. Finally, I document that the growth of large firms by expanding the number of establishments caused the share of establishments operated by large firms to increase by 60% since the early 1980s. These findings suggest that it has become more profitable for multi-establishment firms to operate more but smaller establishments plausibly due to recent innovations in logistics, information and communications technology. The expansion of large firms due to technological improvements could affect the firm formation rate by crowding out smaller firms which represent the majority of startups.

As a second contribution, I incorporate this insight into a model of industry dynamics to quantitatively evaluate the impact of technological improvements specific to large firms on firm formation rates. I find that reductions in establishment-level fixed cost and returns to scale of large firms that match the observed change in the structure of large firms can explain about 2/3 of the decline in firm formation rates. The decline in firm formation rates has no adverse welfare effects: Consumption per capita increases by 7.3% in the economy recalibrated to the early 2000s relative to the late 1980s calibration. Finally, using cross-sectional data, I provide empirical support for the prediction that expansion by large firms crowds out new firm formation by showing that metropolitan areas in which the fraction of establishments operated by large firms increased more, the firm creation rate decreased more.

Overall these findings suggest that a significant part of the decline in startup rates can be explained as an efficient response to size-biased technological change in contrast to being a response to inefficient increases in regulatory burden. Thus lower startup rates are not necessarily a bad sign, as suggested by parts of the literature, media, and policymakers.

Chapter 2

On the Cyclicality of Entrepreneur and Startup Characteristics

2.1 Introduction

Recessions are times of distress for the economy. As authors since Schumpeter (1939) argue, they nonetheless play an important role in the process of creative destruction, as low productivity firms are cleansed out of the economy at a higher rate, making room for new firms to enter. New firms however, are very heterogeneous.[1] In particular, it is not clear whether new firms entering during recessions are of higher or lower quality than new firms entering during expansions. As firm characteristics are highly persistent and a substantial part of firm heterogeneity is already determined at startup,[2] cohorts of low quality firms entering in recessions could generate persistent hysteresis effects and substantially dampen the cleansing effects of recessions.

Recent research on firm creation in the United States indeed suggests that firms entering in recessions are of lower quality in the sense that they start smaller and persistently stay smaller than firms started during expansions.[3] Various mechanisms that result in procyclical relative incentives to start high-growth vs. low-growth firms, such as countercyclical financial frictions or procyclical variations in demand, have been proposed to explain the procyclicality of startup quality. We propose a contrasting mechanism. Based on the insight that entrepreneurs differ substantially in their ability to start high-growth firms, we show, using Swedish micro-data, that part of the variation in the share of new

[1] See e.g. Hurst and Pugsley (2011), Decker et al. (2016), Guzman and Stern (2016b), Haltiwanger et al. (2016), Catalini et al. (2019).

[2] Pugsley et al. (2018) show that 40% of firm heterogeneity at age 20 is due to ex-ante heterogeneity. Sedláček and Sterk (2017) show that variation in cohort level employment at age five is to more than 1/3 explained by variation in startup size.

[3] Sedláček and Sterk (2017) show a positive correlation between total startup employment and GDP growth as well as employment growth. Moreira (2016) documents the procyclicality of the size of new establishments using US establishment level microdata.

high-growth firms is driven by the procyclical variation in the supply of entrepreneurial quality.

A priori this is far from obvious. While one might naturally expect that the number of firms created out of unemployment varies countercyclically and thus influences the procyclical variation in the average quality of entrepreneurs, the corresponding variation is far from obvious for high quality entrepreneurs that create firms by moving out of dependent employment. It is rather easy to generate conflicting hypotheses on the incentives of such entrepreneurs to create a new firm in an expansion vs. a recession. In the end, the relative weight of the forces influencing that transfer is an empirical matter.

At any rate, we establish that the procyclicality of the number and size of new firms is a stylized fact across multiple countries. Using Swedish microdata at the firm level, we confirm prior results based on U.S. data. At the regional level, both the number and the size of new firms is lower during times of high than during times of low unemployment. While the differences in average firm size in Sweden are less persistent than in the US, aggregate employment of cohorts started during times of high unemployment is still significantly lower four years after firm creation. These findings suggest that firm quality varies procyclically.

Continuing to use Swedish microdata that allow us to match firm owners to firms which in turn we can follow over their life cycle, we document that entrepreneur characteristics significantly affect firm characteristics. We draw on the notion of necessity vs. opportunity entrepreneurship which suggests that better outside options at the time of firm creation raise the firm quality threshold above which entrepreneurship becomes attractive and segment potential entrepreneurs according to their labor force status just prior to firm startup. Beyond most prior studies who differentiate only between previously unemployed and employed entrepreneurs, our data allow us to further differentiate both the previously employed by their previous income in dependent employment and the previously unemployed by unemployment duration.[4] Consistent with the intuition that entrepreneurs with higher outside options should start better firms we find sizable within group heterogeneity in terms of firm characteristics. Among previously employed workers there is a monotone relationship between prior income and startup size at birth, as well as firm size at age five. By contrast, the sizes at birth and at age five of firms started by low wage founders are not significantly different from those of founders who previously registered as unemployed or were out of the labor force. The survival rates are significantly higher for all firms started by founders who previously were employed. All this suggests that some unemployed founders use entrepreneurship as an interim status

[4] Choi (2017) only includes previously employed entrepreneurs and estimates the relationship between prior wages and firm size. Camarero Garcia and Murmann (2019) only includes previously unemployed entrepreneurs and shows differences in average firm size of entrepreneurs who experienced above and below median unemployment duration.

towards transiting back into dependent employment.[5]

The substantial differences in the size and survival rate of firms started by different founders naturally lead to asking how much the employment created by cohorts of new firms is affected by variation in the characteristics of entrepreneurs. We show that cohort employment is persistent both in the aggregate as well as for different founder types. More precisely, the correlation between cohort employment at birth and at age 4 is above 0.5 for all firms as well as for five out of the eight entrepreneur segments we analyze. We also document that the average firm size of cohorts is quite persistent with a correlation of more than 0.3, and it is particularly persistent for entrepreneurs starting firms out of high-income employment.

The variation in the employment impact of cohorts can be caused by variation in the number of firms as well as by variation in the average firm size. For all firms together around 60% of variation in employment at birth is due to the variation in the number of new firms and about 40% is due to the variation in average firm size. The contribution of average firm size increases over the life cycle so that the share of variation in total employment at age four that can be attributed to the variation in the average size of firms increases to about 70%. When we disaggregate by entrepreneur segments, the variation in the number of firms accounts for a substantially larger part of the variation in employment for firms started by entrepreneurs out of unemployment or out of the lower income groups. However a significant share of variation, in particular for entrepreneurs coming out of well paid employment, is still due to variation in average firm size. This suggests that while entrepreneur composition might play an important role for the variation in average firm size, even within entrepreneur segments there is substantial variation in firm quality.

We further decompose the variation in average firm size into variation within groups and variation in the composition of entrepreneurs. This decomposition shows that only 13.5% of the variation in average firm size is due to the variation across groups, and 86.5% of the variation is due to within group changes.[6] Finally we investigate whether the composition of entrepreneurs and the firm characteristics within entrepreneur segments change systematically with the macroeconomic conditions at firm startup. We show that employment creation by startups out of longer term unemployment increases significantly in times of high unemployment, while employment creation out of employment does not change or decreases significantly, depending on the income quintile. While increased employment creation out of longer term unemployment is mostly due to a significantly larger number of firms, the observed decrease in employment creation out of high-income

[5] That entrepreneurship among employees coming out of unemployment or from lower income deciles may be transitional is also supported by the fact that entrepreneurs from these segments who close their firms early have higher incomes than those who continue operating their firms.

[6] Observations are at the county by year level. When the observations are weighted by the average county employment over time, the share of within group variation drops to 74% and the share of across group variation increases to 23.5%.

employment is largely driven by a significant decrease in average startup size. So there are two overlapping effects that affect startup size over the cycle and the ensuing longer term hysteresis effect. In times of high unemployment, on the one hand there is a shift towards lower quality entrepreneurs who on average start smaller firms, and on the other hand high quality entrepreneurs start firms that are on average smaller.

We also discuss extensions to the empirical analyses. First, we propose different approaches to improve the identification of founders in general as well as high quality founders. Second, we suggest further analyses on the cyclicality of firm creation by different founder types making more intensive use of the rich set of covariates available in our microdata. Finally we delineate additional explanations for the variation in firm quality and entrepreneur composition along with several empirical specifications to test these hypotheses.

The remainder of the paper is structured as follows. In Section 2.2 we discuss the related literature and place our contributions. In Section 2.3 we describe the data and the definitions used in the empirical analysis. Section 2.4 contains evidence on the cyclicality of firm creation and startup characteristics for the US and Sweden. In Section 2.5 we provide cross sectional evidence on the relationship between founder characteristics and firm characteristics in Sweden with a focus on the prior labor force status of the founder. In Section 2.6 we analyze the sources of variation of the employment impact of firm cohorts and the cyclicality of firm characteristics conditional on founder type. In Section 2.7 we discuss various mechanisms explaining variation in founder quality over the cycle as well as other possible extensions to the analysis, and conclude with Section 2.8.

2.2 Literature review

By the perspective offered in this paper, entrepreneurial characteristics critically influence the characteristics of newly founded firms. Therefore, we juxtapose variations in entrepreneurial and firm characteristics both cross-sectionally and during the business cycle. We distinguish between cross-sectional variation and variation over the business cycle in view of structural, i.e. long term vs. business cycle, i.e. short term oriented policies intended to influence firm birth and longevity. We start with an overview of the literature on cross sectional differences in new firm and entrepreneurial characteristics, followed by that on variations over the business cycle.

2.2.1 Cross sectional variation in new firm characteristics

There is a large literature documenting heterogeneity across firms within countries and sectors (Foster et al., 2008; Hsieh and Klenow, 2009, 2014). In his survey, Syverson (2011) documents large and persistent differences in firms' productivity. While firms exhibit idiosyncratic dynamics, a substantial part of the differences between firms are driven by heterogeneity in startup characteristics, and they remain quite persistent. Pugsley et al. (2018) show that even after 20 years, ex-ante heterogeneity still accounts for about 40% of the cross sectional employment dispersion. Furthermore, enterprises rarely change their incorporation status (Levine and Rubinstein, 2017) and only a small set of high-growth firms is responsible for most of employment growth (Decker et al., 2016; Haltiwanger et al., 2016; Guzman and Stern, 2016a). Guzman and Stern (2016b) show that characteristics founders decide on at the time of firm creation such as the legal form of an enterprise, the firm name and trademark or patent applications are related to the firm's quality as indicated by the probability of an IPO, or a high-value acquisition within the first six years. Using these observable characteristics they construct an entrepreneurial quality index which varies procyclically at the aggregate level.

2.2.2 Cross sectional variation in entrepreneurs' characteristics

An important reason for differences between firms are differences in the quality of management and ownership. New firms are often managed by their founders. There is an extensive literature documenting how entrepreneurs' characteristics affect the performance of new firms. An important distinction, made for example by Schoar (2010), is between subsistence and opportunity entrepreneurs. The former start a business out of lack of a good alternative, while the latter start a business because they want to pursue a good business idea. We construct a proxy measure for such differences in motivation based on an entrepreneur's outside option when starting the firm and study, and how they vary across the business cycle. Hurst and Pugsley (2011) show for the U.S. that most entrepreneurs don't start a business out of growth motives but rather out of life-style motives, for example because they want to be their own boss. Levine and Rubinstein (2017) document that growth oriented entrepreneurs whom they identify with entrepreneurs starting incorporated enterprises, are more highly educated,[7] smarter as measured by learning aptitude tests and engaged in more illicit activities in their youth. Closest to our cross sectional analysis is a recent paper by Choi (2017) that considers the outside options of entrepreneurs as an important determinant of young firms' success. He builds on the notion of Vereshchagina and Hopenhayn (2009) and Dillon and Stanton (2017) that an entrepreneur's outside option in case of failure affects his incentives to enter entrepreneurship and, once entered, his risk-taking incentives. Using prior earnings as a proxy for the post-entrepreneurship outside option, he shows that, consistent with our results for previously employed entrepreneurs, entrepreneurs with higher prior earnings

[7] Blanchflower and Oswald (1998) and Hombert et al. (2014) also show that more highly educated entrepreneurs start larger businesses.

start firms which are larger and which exit at a slightly lower rate. Complementing the cross sectional evidence in Choi (2017) we also include firms started by entrepreneurs who were previously unemployed or out of the labor force that account for roughly one third of all newly started firms. We show that these firms are of similar size as firms started by low wage founders but exit at a significantly higher rate. In addition while Choi (2017) focusses on the role of outside options in case of failure as an insurance mechanism that differs in the cross section and affects risk-taking incentives, we analyze how the outside option to entrepreneurship at the time of starting a firm affects entry into entrepreneurship and the characteristics of new firms started by different entrepreneurs. Furthermore we look at the business cycle variation in outside options.

2.2.3 Business cycle variation of firm characteristics

Our paper relates to an extensive literature going back to Schumpeter (1939), whose authors investigate whether recessions are cleansing in the sense that the average productivity of active firms improves. Intuitively in recessions the profitability of all firms decreases. The least productive firms become unprofitable and exit, and the production factors they had used will be reallocated to more productive firms. Caballero and Hammour (1994) formalize this intuition in an an industry dynamics model with a vintage structure in which newer vintages of production units are more productive. Using data on job creation in US manufacturing, they show that job destruction rates increase substantially during recessions while job creation rates are much less procyclical. Interpreted through the lense of their model this implies that recessions cleanse out the least productive older firms. Ouyang (2009) proposes a model in which firms learn about their productivity over time. In this world, negative demand shocks have both a cleansing and a sullying effect. The sullying effect is caused by the premature exit of firms who have high productivity but are still in the process of learning their true type. The interruption of the learning process lowers average productivity and quantitatively dominates the typical cleansing effect of recessions by which a larger share of firms exits that are known to be of lower productivity. While newly created firms in these models are ex-ante identical, Barlevy (2002) highlights that recessions might affect the composition of newly created firms or jobs, reduce the average productivity of new firms and thus counteract the cleansing effect among existing firms. Rampini (2004) argues that the fraction of risky high return projects decreases in recessions as entrepreneurs need to bear more risk in recessions due to countercylical agency cost. Thus risk-averse entrepreneurs will choose less risky projects in recessions which leads to a deterioriation of the productivity distribution. Kehrig (2015) documents countercyclical dispersion of the productivity distribution in manufacturing which is driven by declining productivity of unproductive firms during recessions. In contrast to prior work he shows that these results are not only driven by entering and exiting firms but also by incumbents who become less productive. He proposes an explanation based on less efficient use of overhead in recessions which results on the one hand in lower measured average productivity but on the other hand in an increase in the average underlying technical efficiency. Countercyclical variation in the average technical efficiency of entrants and incumbents is typical of this class of models in which entrepreneurs cannot choose endogenously between different firm types and the only fundamental difference between firms is their exogenously determined technical efficiency as e.g. in Clementi and Palazzo (2016).

2.2.4 Business cycle variation of startup characteristics

Countercyclical average productivity of entrants however is at odds with recent research by Sedláček and Sterk (2017) who study the time series properties of startup employment using the US Business Dynamics Statistics. Empirically they establish three stylized facts about startup employment: it is volatile and procyclical, it exhibits high positive autocorrelation, and fluctuations in cohort level employment, where cohorts are specified by year of firm creation, are to a large extent due to fluctuations in average firm size per cohort. The importance of average firm size for explaining total cohort employment increases as cohorts age. Using a structural model they estimate fluctuations in the composition of startup cohorts between high growth and low growth firms. They argue that demand shocks are the quantitatively dominant explanation for fluctuations in the growth potential of startups: Positive demand shocks reduce the cost of customer accumulation making it relatively more attractive to start high growth firms whose share among startups increases. These fluctuations in the initial composition of firm cohorts are the main determinant of a cohort's long term employment impact. They show that the average size of five year old firms is positively correlated with the advertising to GDP ratio at birth, providing additional empirical evidence for the proposed mechanism.

Moreira (2016) studies how business cycle conditions at firm birth affect the composition and growth of firm cohorts. Using establishment level microdata from the US Longitudinal Business Database, she documents that the average establishment size at birth varies procyclically and that the resulting average size differences between expansionary and recessionary cohorts do not dissipate over time. She explores various explanations for the cyclicality of average size at entry and the persistence of size differences over time. Size at entry is more procyclical for new establishments of multi-establishment firms as well as in industries with higher average startup costs and a larger median establishment size. Interestingly, the differences in average firm size across cohorts seem not to be driven by differences in labor revenue productivity. To the contrary, Moreira (2016) shows that revenue per worker varies countercyclically with respect to conditions at birth which suggests stronger selection in recessionary periods. These differences seem to persist over time and are not muted by stronger post-entry selection of expansionary cohorts. This finding is slightly at odds with her model which posits an important role of stronger post-entry selection of expansionary cohorts as a result of weaker selection at entry for explaining the high persistence of average size differences over time.

Lee and Mukoyama (2015) study patterns of establishment level entry and exit over the business cycle for the US manufacturing sector using the Annual Survey of Manufactures. Consistent with Sedláček and Sterk (2017) they find that aggregate employment creation by new plants varies procyclically. In contrast to the findings for the all sectors in Sedláček and Sterk (2017) and Moreira (2016) they show however that the size of new manufacturing establishments varies countercyclically with recessionary establishments being larger. These size differences coincide with mild countercyclical fluctuations in the productivity of new plants.

Adelino et al. (2017) study how employment in the non-tradable sector adjusts to local demand shocks, focusing on differences in the response of firms at different points on their life cycle. They show that net employment creation by new firms is much more procyclical than of incumbent firms. They use a Bartik (1991) instrumental variable strategy to generate exogenous variation in local income. Entry of new firms is responsible for 90% of net job creation in response to these positive demand shocks, while incumbent firms account for less than 10%. They also show that jobs created by new firms in response to positive demand shocks are longer lived than jobs created in times of low demand. As their analysis is based on the Quarterly Workforce Indicators, they look at total job creation by firm age however, and cannot differentiate between the changes in average firm size and changes in the number of firms, which we look at.

2.2.5 Business cycle variation of entrepreneur characteristics

Bernstein et al. (2018) study how the composition of entrepreneurs in the non-tradable sector in Brazil changes in response to local demand shocks. Using exogenous variation in global crop prices and heterogeneity in historical crop production across municipalities they construct exogenous variation in municipality level income, which proxies for local demand. They show that firm creation increases in response to positive demand shocks, but only in the non-tradable sector. These firms are of similar quality in terms of survival and employment growth as firms created in normal times. Their main contribution shows that the entrepreneurs starting firms in response to these shocks are younger and more educated than the average entrepreneur. This results in a small change in the composition of entrepreneurs with the share of young entrepreneurs increasing by at most 2.9 percent. Among the young potential entrepreneurs it is particularly the group of more highly educated and those with non-routine cognitive jobs who start a firm in response to positive demand shocks. Older potential entrepreneurs do not differ in their responsiveness across education groups.

A different strand of literature uses data from large representative surveys to investigate the cyclicality of entry into entrepreneurship. Fairlie (2013) uses CPS data to show that business creation reacts countercyclically to labor market conditions. High unemployment rates increase the probability that individuals start businesses. This effect is especially pronounced for individuals who are unemployed prior to entry into entrepreneurship. These findings include all employer and nonemployer businesses and stand in contrast to our findings of procyclical firm creation for employer firms in the US and for all firms in Sweden.[8] Taken together with our results for the US, the quality of firms changes countercyclically as the quality of employer firms decreases and the share of nonemployer firms increases in times of high unemployment.

Starting from the observation that the decision of previously unemployed workers to start a firm reacts more strongly to high unemployment rates, Fairlie and Fossen (2019) propose to differentiate between opportunity and necessity entrepreneurs according to the entrepreneurs labor market status prior to starting the firm. Their differentiation between opportunity and necessity entrepreneurs is thus similar to our differentiation albeit significantly coarser. The dichotomy between these two groups masks significant heterogeneity within each group of entrepreneurs. They confirm the results from Fairlie (2013) using a longer time span for the US and an additional survey based data set from Germany. Firm creation by previously unemployed founders, whom they label necessity entrepreneurs, increases in times of high national and local unemployment while firm creation by previously non-unemployed founders, whom they label opportunity entrepreneurs, decreases. In addition they show that the characteristics of firms created by opportunity entrepreneurs differ from those of firms created by necessity entrepreneurs. Opportunity entrepreneurs are more likely to start incorporated businesses and employer businesses. Interestingly with only about 15% of new firms, the share of employer businesses is quite low even for opportunity entrepreneurs in their sample.

Fossen (2019) shows that most of the variation in entry rates into self-employment during the time of the great recession is driven by changes in the stock of unemployed workers. Both unemployed and employed workers' propensity to start a business does not change over the cycle, it is just the higher propensity of unemployed workers relative to employed workers and their higher share in recessions that is responsible for the countercyclicality of entry into self-employment.

Svaleryd (2015) investigates the relationship between the local business cycle and self-

[8] Congregado et al. (2012) find that the rate of self-employed entrepreneurs owning non-employer businesses moves countercyclically while the rate for entrepreneurs owning employer businesses moves procyclically. While their findings concern stocks of businesses and not flows, they are consistent with the different findings regarding the cyclicality of firm creation of employer and non-employer businesses between our paper and Fairlie (2013)

employment rates in Sweden. She finds that higher local vacancy rates or lower local unemployment rates are related to higher self-employment rates, particularly for more educated persons. These findings are consistent with ours as education is positively related to an individual's earnings.

Albert and Caggese (2019) investigate how the interplay between fluctuations in GDP and financial conditions affect the composition of startup cohorts. Using survey data from the Global Entrepreneurship Monitor they show that during times of financial distress the share of new firms with high growth potential declines. This effect is amplified in times of low GDP growth. To identify high growth firms they rely on ex-ante expectations of entrepreneurs about their firm's growth potential. Those expectations however cannot be verified ex-post as the GEM is structured as a repeated cross section and firms thus cannot be followed over time. In contrast our data allows us to identify high-growth potential entrepreneurs by using observed firm performance of different entrepreneur types.

2.3 Data

2.3.1 United States

Firm data

For the US we use the publicly available Business Dynamics Statistics (BDS) data product of the US Census. The BDS is based on the US Census' Longitudinal Business Database (LBD) which is a yearly panel containing establishment level information on the employment of nearly all private US establishments with at least one payroll employee in the week of March 12th each year. The Census aggregates individual establishments into firms, where a firm is a collection of all establishments under common ownership. The public use BDS data contains yearly data on, e.g. the number of firms, the number of establishments and employment, aggregated along different strata as, e.g. sectors, MSAs, firm size groups or firm age groups. The data run from 1976 to 2014.[9]

New firms In this data set, firm age is assigned according to the age of the oldest establishment in a firm. Establishments age naturally one year at a time and are assigned age 0 in the year when they first report positive payroll employment. We denote age 0 firms as startups; thus a startup is a firm which is responsible for genuinely new business activity at a new location. M&A activity or legal reorganization of firms thus will not be classified as a startup.

[9] We use data starting in 1979, due to data problems in 1977 and 1978 as noted by Moscarini and Postel-Vinay (2012). Furthermore, when using MSA data, we start in 1981 due to data problems for Texas MSAs in 1979 and 1980.

Cyclical indicators

We use the local area unemployment statistics (LAUS) from the BLS for data on MSA level unemployment rates. The LAUS data run from 1990 to 2017. As MSA level GDP data are only available from 2000 to 2017 and for consistency with the analysis for Sweden and the analysis in Sedláček and Sterk (2017), we use MSA level employment growth instead of MSA level GDP growth as a demand side cyclical indicator. We calculate MSA level employment growth from the BDS data.

2.3.2 Sweden

Firm data

We use data from Statistics Sweden on firms and workers from the LISA database. LISA is a linked employer-employee panel that covers all registered inhabitants in Sweden above the age of 16 in the years 1990 to 2011. Since LISA is administrative data based on government records we capture the population of individuals that live in Sweden and work in the firms of interest. LISA also contains accounting information from the firm level FEK database, which allows us to observe a multitude of company level characteristics as well as information on the operative CEO of all firms in Sweden.

New firm We consider both limited liability companies and sole proprietorships. Statistics Sweden differentiates between firms and establishments. A firm is defined as a new firm if it consists entirely of new establishments. An establishment is defined as a new establishment in year t if it has employees registered in LISA in year t but not in any prior year.

Entrepreneurs We use the methodology from Statistics Sweden which is detailed in Andersson and Andersson (2012) to assign an owner to each firm in Sweden. The methodology uses data on an individual's age and labor income relative to other employees in the firm, as well as an "entrepreneur status" assigned by Statistics Sweden based on asset ownership in and income from the firm. Statistics Sweden changed the calculation of the entrepreneur status which has the highest weight in the ownership definition in 2004 which produces a break in the time series. In section 2.7.1 we discuss other possible ways to define entrepreneurs using additional individual and firm level characteristics in order to overcome the time series break in the definition of the entrepreneur status.

Cyclical indicators

We primarily look at data on labor market conditions at the time of startup which makes our results on cyclicality comparable with the results using US data where we also look at labor market conditions at startup. Labor market conditions also provide a good proxy for the outside options on the labor market at the time at startup and due to persistence in labor market conditions also proxy for the outside option in case of failure. Furthermore we can generate these variables from within the LISA dataset which includes data on unemployment from which we can calculate the unemployment rate as well as on employment from which we can calculate employment growth.

2.4 The cyclicality of startup characteristics

2.4.1 Evidence from the U.S.

In a recent paper, Sedláček and Sterk (2017) use aggregate data for the U.S. to show that job creation by startups positively correlated with aggregate employment growth as well as GDP growth. Furthermore they argue that the economic conditions during the startup phase have a persistent impact on total cohort employment by affecting the quality of new firms. They show that the average size of a cohort that starts small never catches up. In this section we investigate the relationship between economic conditions at birth and the employment impact of cohorts at a geographically more disaggregated level. We use cross sectional variation at the level of US metropolitan statistical areas (MSAs), which are groups of counties that are tightly economically integrated.

Regressions We document the relationship between economic conditions during the startup phase and various characteristics of firm cohorts. Two measures describe the economic conditions at birth. First, we use the total employment growth rate in the metro area which proxies for growth in final demand during the startup phase. The idea is that most startups initially serve local demand which is positively related to local employment.

Second, we use the unemployment rate during the startup phase which proxies for the supply of entrepreneurs, namely the outside options that potential entrepreneurs had in the labor market during the startup phase. On the one hand, when unemployment is high, there are more unemployed potential entrepreneurs, which have bad outside options and start bad firms on average. On the other hand potential entrepreneurs that are employed, who would on average start better firms, may be reluctant to give up their job as they face a higher probability of firm failure which would result in unemployment. Overall, we expect the quality of firm cohorts to be positively correlated with demand conditions during startup and negatively correlated with the unemployment rate.

We start by looking at the total employment of firm cohorts by age including and following their startup year, by considering the following regression:

$$Y_{t+a,a,m} = \beta_1 \frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}} + \beta_2 u_{t,m} + \gamma_m + \delta_t + \epsilon_{t,m} \tag{2.1}$$

where $Y_{t+a,a,m}$ is outcome variable for firms of age a which were born in year t in MSA m. $\frac{Emp_{t,m}-Emp_{t-1,m}}{Emp_{t-1,m}}$ is the employment growth rate from years $t-1$ to t in MSA m. $u_{t,m}$ is the unemployment rate in year t in MSA m, γ_m is an MSA fixed effect and δ_t is a time fixed effect.

Consider first $\log(Emp)_{t,0,m}$, the log of cohort employment at age 0 as the outcome variable. In column (1) and (4) in Table 2.1 we use only the unemployment rate as a measure of economic conditions at birth while in column (2) and (5) we use only the employment growth rate. Column (3) and (6) contain specifications involving both measures. In columns (1) to (3) we control for MSA fixed effects and in columns (4) to (6) additionally for time fixed effects. The preferred specification is (6) where we use both the unemployment and the employment growth rate and control for both MSA and time fixed effects. We see that a one percentage point increase in the local employment growth rate is associated with a 0.98 percent increase in the employment of a startup cohort. Furthermore, if the unemployment rate is one percentage point higher, the employment of a new cohort of firms is 1.86 percent lower.[10]

Dep. var: $\log(Emp)_{t,0,m}$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		1.6951***	1.1830***		1.1000***	0.9819***
		(0.0849)	(0.0793)		(0.0877)	(0.0882)
MSA unemp. rate	-5.6303***		-5.2792***	-2.1882***		-1.8627***
	(0.1385)		(0.1387)	(0.2010)		(0.2016)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.17	0.05	0.19	0.01	0.02	0.03
No. observations	7992	7992	7992	7992	7992	7992

Table 2.1: U.S.: Total cohort employment of startups and economic conditions at birth

Consider next a decomposition by the two margins along which cohort employment can vary: the extensive margin, i.e. the number of firms per cohort and the intensive

[10] To make these measures slightly more comparable note that the standard deviation of MSA level employment growth is 3.8 percentage points and the standard deviation of the unemployment rate is around half of that at 1.9 percentage points. Thus a one standard deviation increase in MSA employment growth is associated with an increase in cohort employment at birth of ca. 3.7 percent. A one standard deviation increase in the unemployment rate is associated with a decrease in cohort employment at birth of ca. 3.5 percent.

margin, i.e. the employment per firm in the cohort. Table 2.2 shows how economic conditions in year t affect the number of firms that are started in year t. The outcome variable is $\log(Firms_{t,0,m})$, the log of the number of startups in year t. Similar to total employment, the MSA employment growth rate is positively and the unemployment rate is negatively correlated with the number of startups. A one percentage point increase in the employment growth rate is associated with a 0.18% increase in the number of new firms. A one percentage point higher unemployment rate at birth is associated with a 1.48% lower number of startups. Thus around 75% of the impact of unemployment rate and 20% of the impact of employment growth at birth on total cohort employment is due to the variation in the number of firms.

Dep. var: $\log(Firms_{t,0,m})$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		1.3041***	0.8765***		0.2794***	0.1851***
		(0.0580)	(0.0517)		(0.0509)	(0.0509)
MSA unemp. rate	-4.6682***		-4.4081***	-1.5488***		-1.4874***
	(0.0907)		(0.0904)	(0.1152)		(0.1164)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.25	0.06	0.28	0.02	0.00	0.02
No. observations	7992	7992	7992	7992	7992	7992

Table 2.2: U.S.: Number of startups and economic conditions at birth

In Table 2.3 we look at the intensive margin, i.e. the average size of firms at birth. The outcome variable is $\log(\frac{Emp_{t,0,m}}{Firms_{t,0,m}})$, the log of the size of firms started in year t at age 0. Here we see that the remaining impact of economic conditions at birth on total cohort employment at birth comes through the intensive margin.[11] Thus roughly 25% of the impact of the unemployment rate at birth and 80% of the impact of employment growth at birht on initial employment of startups comes through their impact on firm size. A one percentage point increase in MSA level employment growth is associated with a 0.79% increase in average firm size and a one percentage point increase in the unemployment rate is associated with a 0.38% decrease in average firm size of newly started firms.

Let us now turn to longer term measures of startup success. First we consider the relationship between economic conditions at birth and the long term employment impact of a startup cohort. In Table 2.4 we consider $\log(Emp)_{t+5,5,m}$ as outcome variable, the total employment of five year old firms in year $t+5$ or in other words the total employment at age five of firms that were started in year t. Here we see that the economic conditions at birth have a long lasting impact on the total employment generated by a cohort of firms. A one percentage point higher total employment growth in the MSA in the year

[11]Summing the coefficients from Table 2.2 and Table 2.3 yields the coefficients from Table 2.1.

Dep. var: $\log(\frac{Emp_{t,0,m}}{Firms_{t,0,m}})$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		0.3884***	0.3042***		0.8183***	0.7945***
		(0.0645)	(0.0652)		(0.0752)	(0.0760)
MSA unemp. rate	-0.9586***		-0.8683***	-0.6389***		-0.3756**
	(0.1125)		(0.1140)	(0.1731)		(0.1738)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.00	0.01	0.00	0.01	0.02
No. observations	7992	7992	7992	7992	7992	7992

Table 2.3: U.S.: Startup size at birth and economic conditions at birth

in which the cohort entered is associated with 0.88% more employment generated by the cohort five years later. A one percentage point higher unemployment rate in the year in which the cohort entered is associated with 0.64% less employment generated by the cohort five years later.

Dep. var: $\log(Emp)_{t+5,5,m}$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		1.0978***	1.0465***		0.9278***	0.8853***
		(0.1181)	(0.1180)		(0.1243)	(0.1259)
MSA unemp. rate	-1.7918***		-1.6426***	-0.9888***		-0.6415**
	(0.2530)		(0.2521)	(0.3033)		(0.3061)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.01	0.02	0.00	0.01	0.01
No. observations	6323	6323	6323	6323	6323	6323

Table 2.4: U.S.: Cohort employment after five years and economic conditions at birth

In Table 2.5 we consider $\log(\frac{Emp_{t+5,5,m}}{Firms_{t+5,5,m}})$ as outcome variable, the average firm size at age five of firms started in year t. That average size is also affected by the economic conditions in the startup year. A one percentage point higher employment growth in the MSA during the startup year is associated with a 0.54% higher average firm size at age five. This is around 60% of the total impact on employment after five years. A one percentage point higher unemployment rate is not associated with a significantly different average firm size at age five.

Finally, in Table 2.6 we consider $\frac{Firms_{t+5,5,m}}{Firms_{t,0,m}}$, the five year survival rate of firms started in year t as the outcome variable. Vardishvili (2018) has recently investigated the impact of economic conditions at birth on survival rates at the aggregate level. She finds that *establishments* (not necessarily corresponding to firms) created during recessionary periods

Dep. var: $\log(\frac{Emp_{t+5,5,m}}{Firms_{t+5,5,m}})$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		0.6045***	0.5867***		0.5303***	0.5460***
		(0.1062)	(0.1064)		(0.1135)	(0.1151)
MSA unemp. rate	-0.6532***		-0.5695**	0.0232		0.2375
	(0.2274)		(0.2274)	(0.2766)		(0.2798)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.00	0.01	0.01	0.00	0.00	0.00
No. observations	6323	6323	6323	6323	6323	6323

Table 2.5: U.S.: Average size after five years and economic conditions at birth

have higher survival rates than those created during expansionary periods. Considering *firms* and looking at a more disaggregate level, we find ambiguous results. In particular the total employment growth during the year of birth at the MSA level is slightly positively correlated with the five year survival rate of a startup cohort. In line with Vardishvili (2018), the unemployment rate in the startup year is significantly negatively correlated with the five year survival rate. A one percentage point higher unemployment rate in the year of birth is associated with a 0.21 percentage point higher survival rate until age five.

Dep. var: $\frac{Firms_{t+5,5,m}}{Firms_{t,0,m}}$	(1)	(2)	(3)	(4)	(5)	(6)
MSA empl. growth		0.0154	0.0141		0.0321**	0.0462***
		(0.0166)	(0.0167)		(0.0158)	(0.0160)
MSA unemp. rate	-0.0446		-0.0426	0.1942***		0.2123***
	(0.0355)		(0.0356)	(0.0384)		(0.0389)
MSA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.00	0.00	0.00	0.25	0.25	0.25
No. observations	6323	6323	6323	6323	6323	6323

Table 2.6: U.S.: Five year firm survival rates and economic conditions at birth

2.4.2 Evidence from Sweden

For Sweden we run regressions at the county (NUTS 2 region) level. There are a total of 20 counties in Sweden over the sample period. In 1997 and 1998 there were two reorganizations where several counties merged.[12] We aggregate the data for these counties

[12]In 1997 Kristianstad County and Malmöhus County merged into Skane County. In 1998 Älvsborg County, Gothenburg and Bohus County and Skaraborg County merged into Västra Götaland County.

in the years prior to the merger in order to represent the current county division. Akin to the U.S. case, we run regressions of the following type:

$$Y_{t+a,a,m} = \beta_1 \frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}} + \beta_2 u_{t,m} + \gamma_m + \delta_t + \epsilon_{t,m} \tag{2.2}$$

where $Y_{t+a,a,m}$ is outcome variable for firms of age a which were born in year t in county m. $\frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}}$ the employment growth rate from year $t-1$ to t in county m. $u_{t,m}$ is the unemployment rate in year t in county m, γ_m is a county fixed effect and δ_t is a year fixed effect to flexibly control for a time trend and a time series break in 2004.

We consider similar dependent variables as in the U.S. case and run a total of six different regression specifications. County fixed effects are included in all specifications, time fixed effects only in specifications (4) - (6). In specification (1) and (4) we only consider county level employment growth as an additional independent variable. In specifications (2) and (5) we only consider the county level unemployment rate, and in specifications (3) and (6) we consider both the county level employment growth rate and the county level unemployment rate. As there is a visible time series break in 2004 due to a change in the several variable definitions in the underlying LISA microdata, our preferred specifications include the year fixed effects to mitigate distortions introduced by this time series break.

In Table 2.7 we consider total cohort employment at birth which is positively but insignificantly correlated with county level employment growth in the birth year (column (4)), but significantly negatively correlated with the county level unemployment rate in the birth year (column (5)). Taking both employment growth and the unemployment rate into account (column (6)), the semi-elasticity of total birth year employment with respect to the unemployment rate is -2.86. Thus a one percentage point increase in the unemployment rate is associated with a 2.86 percent decrease in the total employment of newly started firms. This is about half the elasticity for the U.S. and points to a less cyclical employment impact of new firms in Sweden. The effect of employment growth is statistically insignificant.

As described in subsection 2.4.1, total employment at birth can change along two margins, a change in the total number of new firms or a change in their average size. In Table 2.8 we look at the cyclicality of the number of firms. The number of firms does is not statistically significantly affected by the county employment growth rate. It is negatively correlated with the unemployment rate at birth, however, with a semi-elasticity of -1.37 which is significant at the five percent level. Thus a one percentage point increase in the unemployment rate is associated with a decrease in the number of new firms of 1.37 percent. Thus about half of the procyclicality of total cohort employment at birth is generated by a procyclical number of new firms.

In Table 2.9 we look at the cyclicality of firm size at birth, the second margin which

Dep. var: $\log(Emp)_{t,0,m}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.6698		0.4551	0.3636		0.1442
	(0.4339)		(0.4376)	(0.4200)		(0.4213)
County unemp. rate		-0.9944***	-0.9250**		-2.9211***	-2.8624***
		(0.3484)	(0.3546)		(0.9546)	(0.9713)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.03	0.03	0.00	0.03	0.03
No. observations	300	300	300	300	300	300

Table 2.7: Sweden: Total cohort employment and economic conditions at birth

Dep. var: $\log(Firms_{t,0,m})$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	-0.2706		-0.6298	0.1056		0.0004
	(0.4604)		(0.4563)	(0.2745)		(0.2772)
County unemp. rate		-1.4518***	-1.5478***		-1.3727**	-1.3725**
		(0.3638)	(0.3699)		(0.6280)	(0.6391)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.00	0.05	0.06	0.00	0.02	0.02
No. observations	300	300	300	300	300	300

Table 2.8: Sweden: Number of startups and economic conditions at birth

determines total cohort employment. Once we account for year fixed effects, there is no significant correlation between county level employment growth and initial firm size. The correlation between initial firm size and the local unemployment rate is significantly negative. The semi-elasticity is of similar magnitude as the semi-elasticity of the number of firms. Thus both the extensive and the intensive margin contribute about half to the cyclicality of total cohort employment at birth.

An important finding from Sedláček and Sterk (2017) is that the employment impact of cohorts is quite persistent. Thus a natural question that arises is whether the cyclicality of firm characteristics at entry also persists into later years. Taking into account year fixed effects, the results in Table 2.10 show a negative correlation between the birth year unemployment rate and total cohort employment after four years, which is significant at the 10% level.[13] This indicates that the cyclicality of employment at birth indeed persists into later years.

[13]The substantially larger elasticity compared to the employment impact at birth is partially explained by the fact that we cannot compute four year employment for cohorts after 2007. If we calculate the elasticity of employment at birth using this restricted sample, we find an elasticity of roughly 4.

Dep. var: $\log(\frac{Emp_{t,0,m}}{Firms_{t,0,m}})$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.9404***		1.0849***	0.2580		0.1437
	(0.2446)		(0.2455)	(0.3223)		(0.3259)
County unemp. rate		0.4574**	0.6229***		-1.5484**	-1.4899*
		(0.2014)	(0.1989)		(0.7385)	(0.7513)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.05	0.02	0.08	0.00	0.01	0.02
No. observations	300	300	300	300	300	300

Table 2.9: Sweden: Size of startups and economic conditions at birth

Dep. var: $\log(Emp)_{t+4,4,m}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	1.0082		2.1491***	0.0857		-0.1886
	(0.7396)		(0.7488)	(0.9484)		(0.9549)
County unemp. rate		2.1167***	2.6343***		-4.4787*	-4.5531*
		(0.5365)	(0.5584)		(2.4093)	(2.4434)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.06	0.09	0.00	0.01	0.01
No. observations	240	240	240	240	240	240

Table 2.10: Sweden: Cohort employment after four years and economic conditions at birth

Interestingly the firms started in times of high unemployment don't seem to catch up in terms of firm size to those created in better times. the average size of startups after four years is also negatively but statistically insignificantly correlated with the local unemployment rate at birth as shown in Table 2.11.

Also in terms of the number of firms, cohorts started in times of high unemployment don't catch up. Table 2.12 shows that the four year survival rate of these firms is not higher than that of firms started in better times. Thus the number of firms is still significantly lower after four years.

To summarize, when it comes to the contribution of unemployment on the employment generated by startups, the conclusions that we can draw from the Swedish data are qualitatively similar to those drawn from the U.S. data. However, the magnitudes differ substantially. While the effect from unemployment is negative in both countries, the effect on both the number of startups and startup size is stronger in Sweden. In the U.S., the contribution of the number of startups to the employment generated by startups is with roughly 75 % substantially higher than in Sweden with roughly 50%. While in the U.S.

Dep. var: $\log(\frac{Emp_{t+4,4,m}}{Firms_{t+4,4,m}})$	(1)	(2)	(3)			
County empl. growth	0.1177		1.1228	0.1985		0.0797
	(0.6847)		(0.6962)	(0.9214)		(0.9331)
County unemp. rate		2.0505***	2.3210***		-2.0046	-1.9732
		(0.4931)	(0.5192)		(2.3542)	(2.3877)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.00	0.07	0.08	0.00	0.00	0.00
No. observations	240	240	240	240	240	240

Table 2.11: Sweden: Average firm size after four years and economic conditions at birth

Dep. var: $\frac{Firms_{t+4,4,m}}{Firms_{t,0,m}}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.1625		0.2289**	-0.0606		-0.0726
	(0.1023)		(0.1075)	(0.1095)		(0.1110)
County unemp. rate		0.0983	0.1535*		-0.1702	-0.1989
		(0.0764)	(0.0802)		(0.2802)	(0.2839)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.01	0.03	0.00	0.00	0.00
No. observations	240	240	240	240	240	240

Table 2.12: Sweden: Four year firm survival rates and economic conditions at birth

the effects persist over the ensuing years, they become insignificant in Sweden. Finally, in contrast to the U.S., we see no significant positive demand side effect on the basis of employment growth.

2.5 Founder quality and startup characteristics in Sweden

Below we show that the employment impact of new startups is highly correlated with the characteristics of the founders. In particular, founders with different outside options at the time of starting their firm start significantly different firms. We divide the Swedish population according to their labor force status, which proxies for their outside option, into three major groups, namely

1. people who were working and did not register unemployment in that year

2. people who were unemployed at some point in that year

3. people who were out of the labor force in that year

We further subdivide group 1 into deciles by last year's labor income (segments 1 to 10); group 2 into four segments involving the length of unemployment in the previous year of firm founders classified as previously unemployed, namely 0-90; 91-180, 181-270 and 271-360 days (segments 11 to 14); and group 3 into four segments reflecting those founders' level of education that were classified as previously out of labor force, namely basic, high school, vocational training, and university education (segments 15 to 18). This gives us a total of 18 segments of the Swedish population of founders, all classified by their status in the year before founding the firm.

In order to understand whether the firms created by individuals in these segments differ significantly, we look at several statistics regarding the employment performance of newly started firms. Sedláček and Sterk (2017) argue that the initial employment is an important indicator of a firm's growth potential, and the average initial size of a cohort of new firms an important indicator of the cohort's long term employment impact.

In Figure 2.1 we show the average size at birth for firms started by the different founder segments. There is a clear positive relationship between the size of newly started firms and the prior income of previously employed entrepreneurs. For example a newly started firm of an entrepreneur who was previously employed and had an income in the highest decile employs around 3.5 employees, roughly three times as many as the average firm started by an entrepreneur who had an income in the lowest decile. Firms started by entrepreneurs who were previously employed but received only low incomes employ a similar number of employees at startup as firms started by entrepreneurs coming out of unemployment or from out of the labor force. Thus from a revealed firm performance perspective these entrepreneurs conform more to the notion of subsistence or necessity entrepreneurs than to the notion of opportunity entrepreneurs.

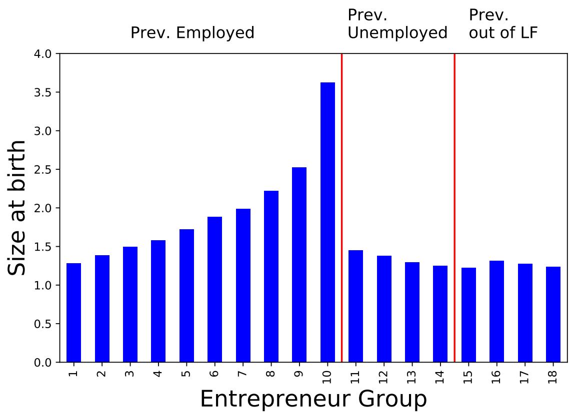


Figure 2.1: Firm size at birth

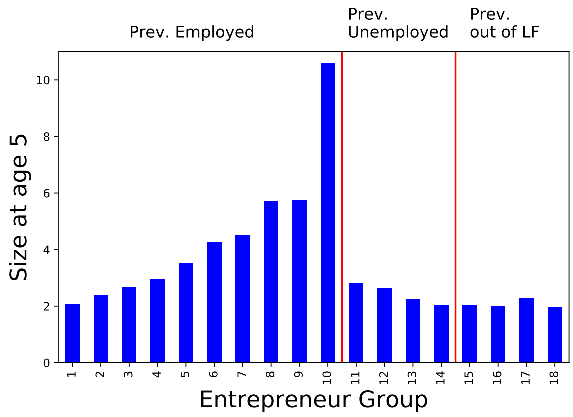


Figure 2.2: Firm size after five years

Firms started by different entrepreneurs not only differ in their initial size but also in their longer term performance in terms of size, employment growth and survival. In

Figure 2.2 we plot for each entrepreneur type the average size of surviving firms after five years. The relationship between entrepreneur type and size at birth persists also in the longer term. In terms of absolute and relative performance, differences are exacerbated in comparison to the differences at birth. On average firms started by founders from the highest income decile are more than three times as large at age five as firms started either by founders from the lowest four income deciles or from founders who come out of unemployment or from out of the labor force.

The overall employment impact of firms started by founders from high income deciles could be mitigated if these firms were high risk firms and survived with a relatively low probability. This is suggested for example by Choi (2017). We plot the survival rate until age five by each entrepreneur type in Figure 2.3. In contrast to the higher risk hypothesis, we do not find a negative relation between an entrepreneur's outside option and the survival probability of the newly started firm. If anything, there is a slightly positive relationship. Overall, among the previously employed entrepreneurs the differences in average survival rates are relatively low at about 5 percentage points or 15% compared to the much larger differences in terms of average size at age five. Firms started by previously unemployed entrepreneurs have slightly lower survival rates, suggesting that they partially use temporary entrepreneurship to transition from unemployment to employment. These differences are nonetheless relatively small such that the overall employment impact is thus mostly driven by the size differences of firms started by different entrepreneur segments.

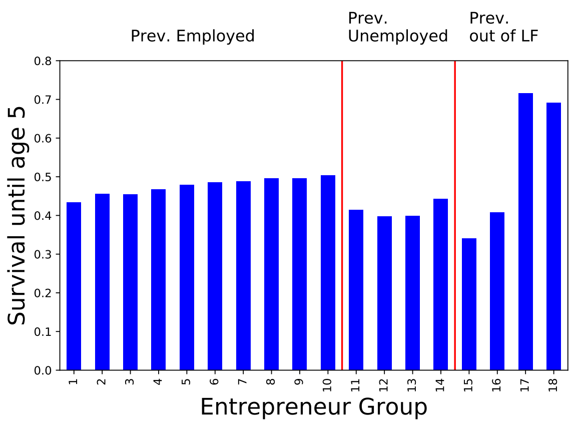


Figure 2.3: Firm survival rate after five years

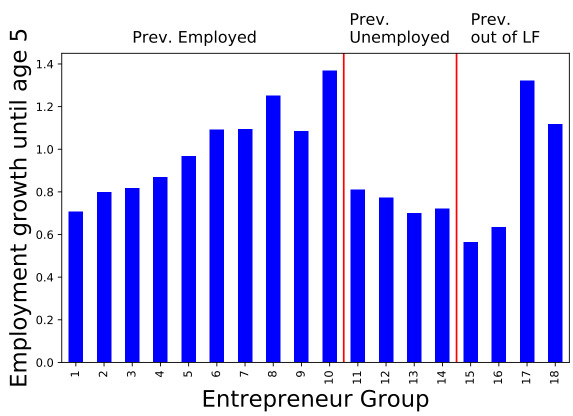


Figure 2.4: Employment growth after five years

We look at the overall employment impact in Figure 2.4. This figure shows for firms started from the different entrepreneur groups and segments, the number of jobs that firms offer at age five relative to the number of jobs that existed at age 0 in all firms started from the respective entrepreneur groups and segments. This measure can be interpreted as the five year employment impact of new firms, accounting for both for firm growth and firm exit. For example, the measure says that for every job in firms started by entrepreneurs from the highest income decile, there will be roughly 1.4 jobs after five years. For every job in firms started by entrepreneurs from the lowest income decile there will only be 0.7 jobs after five years. Thus we observe huge differences in the employment impact across firms started by different entrepreneur segments, which is only natural as this measure combines both the employment growth conditional on survival as measured by the growth in average firm size as well as the survival rates, which are both higher for the higher quality entrepreneurs. Taken together with the similar job creation at startup by firms started by low income entrepreneurs and entrepreneurs out of unemployment, the similar employment growth reinforces the notion that these entrepreneur types are quite similar.

In terms of the share of startups created by each entrepreneur segment where the sum of startups across all three groups constitutes the base, we document in Figure 2.5 that the lowest and highest quality entrepreneurs among the employed are particularly numerous. This U-shaped pattern of firm creation among the previously employed is similar to Poschke (2013). In addition we show that entrepreneurs who were unemployed or out of the labor force constitute an important part of total firm creation as they start roughly 40% of all new firms.[14]

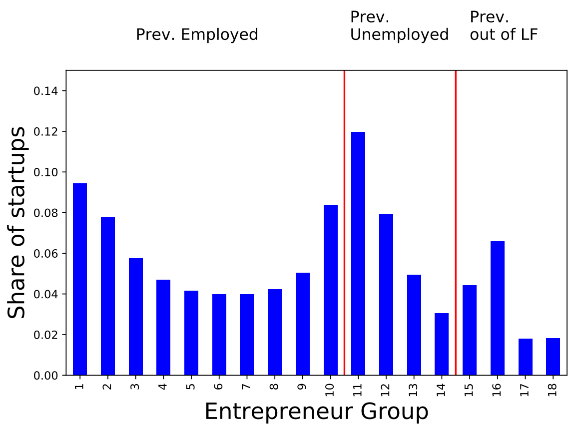


Figure 2.5: Share of startups created by each entrepreneur segment

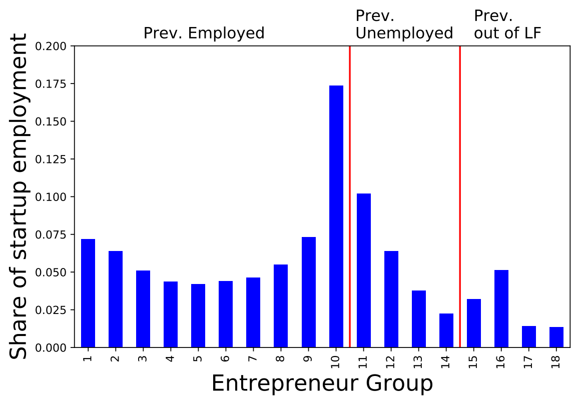


Figure 2.6: Share of employment in startups created by each entrepreneur segment

In terms of the startup employment share, the U-shaped pattern of the firm share combines with the larger firm size for higher quality entrepreneurs to an even more pronounced U-shaped pattern of the employment share. At startup, by far the largest share of employment is created by entrepreneurs coming out of the highest income group. Combined with the much stronger growth of these firms, aggregate medium term employment creation by entrepreneurs from the highest income deciles is substantially higher than that of other entrepreneurs.

[14]In this respect our sample covers a substantial segment of entrepreneurs that other studies such as Choi (2017) that focus only on previously employed entrepreneurs miss.

Another interesting aspect regarding the quality of startups and the relevance of outside options to the entrepreneurs starting firms are the income trajectories of the entrepreneurs around the time of starting a firm. In Figure 2.7 we plot the income trajectories around the time of startup for the different entrepreneur groups. Except for the entrepreneurs from the highest income deciles, the entrepreneurs who leave the firm they started, e.g. by closing it down, have a higher income trajectory than those who stay at their firm. This finding is particularly pronounced for the entrepreneurs with the worst outside options at the time of startup i.e. those who were unemployed and those who among the employed had the lowest income. These patterns suggest that those entrepreneurs with worse outside options either prefer entrepreneurship out of idiosyncratic non-pecuniary reasons, as suggested by Hurst and Pugsley (2011, 2016) or they use entrepreneurship as a transitory option to transition to a new job. The second explanation is probably very relevant for entrepreneurs coming out of unemployment and is consistent with the lower survival probabilities of these entrepreneurs' firms. These entrepreneurs keep searching for a job while they run their own business and quit their business once they find a more attractive dependent employment option. Thus the entrepreneurs who remain are negatively selected and have a lower income than those who close down their firm. Entrepreneurs who start a business out of a good, high paying job on the other hand only start businesses when they see a valuable opportunity to do so. Of those entrepreneurs, the ones who close down their firm do so mostly because their business idea was not successful. Thus the entrepreneurs who remain are positively selected and have a higher income than those who close down their firms. The income difference is due to higher capital income of continuing entrepreneurs. In fact the entrepreneurs who continue operating their firm do have lower labor incomes than those who close down but have significantly higher non-labor incomes.

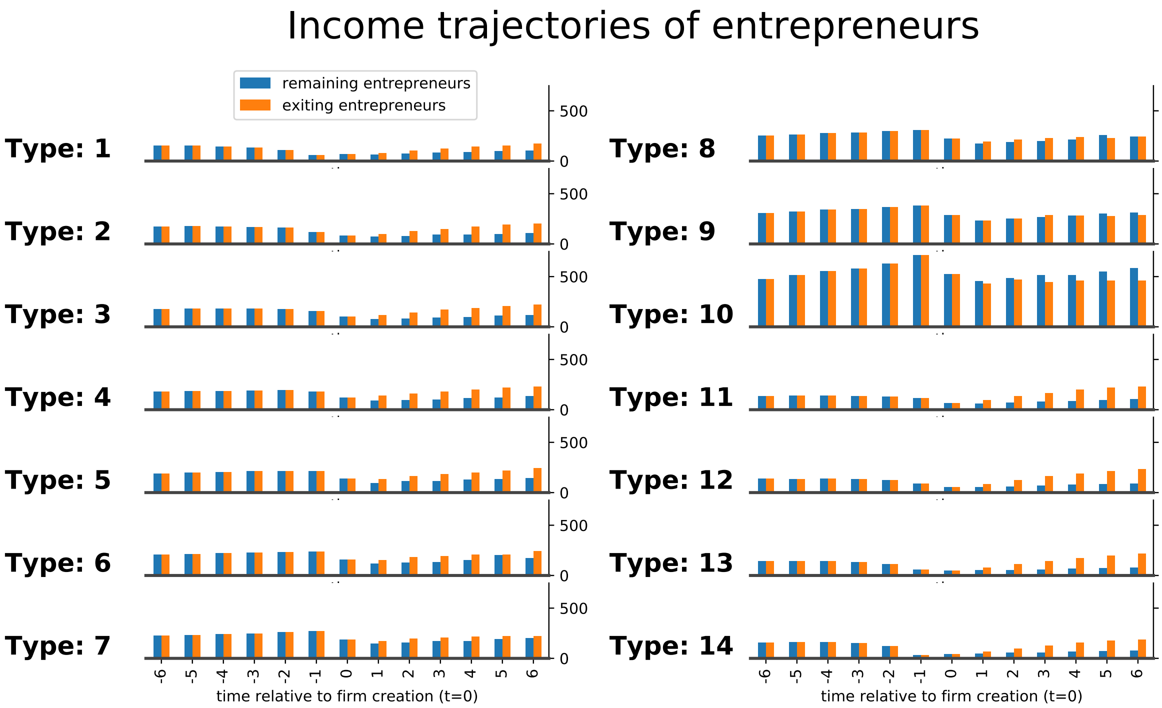


Figure 2.7: Mean income in 1.000 SEK of entrepreneurs around the time of starting a firm differentiated by entrepreneur type

2.6 Start-up founder quality and the employment impact of cohorts

In light of the substantial differences in firm performance across different founder types, a plausible explanation of the variation of the quality of startup cohorts by year of birth as documented by Sedláček and Sterk (2017) could be variation in the composition of founders. While the previous section discussed cross sectional differences between firms started by different entrepreneur types, this section discusses time variation in the employment impact of firm cohorts on aggregate and by entrepreneur type. We follow Sedláček and Sterk (2017) and decompose variation in the employment impact of cohorts over time into variation in the number of firms and variation in average firm size and show that both margins are relevant. We then seek to understand the sources behind the variation in average firm size, by investigating whether they are driven by variation in the initial composition of firm cohorts by entrepreneur segment. Finally we analyze the cyclicality of these margins separately by entrepreneur type, adding to the aggregate evidence from section 2.4.2. As we use county-years as the observational unit, we aggregate some of the entrepreneur segments to guarantee that all cells are filled. From left to right, we display the different entrepreneur types. Segments one to five are entrepreneurs who were employed during the last year earned a wage in the respective quintile. For example segment five earned an income in the fifth quintile i.e. the top 20% of the previous year's income distribution. Segments six and seven were registered as unemployed for either less (group six) or more (group seven) than 180 days in the last year. Persons from segment eight were not in the labor force last year. For comparison we also display the total cohort measures in the last column.

2.6.1 Persistence of cohort employment

One of the main findings of Sedláček and Sterk (2017) for the U.S. is that deviations of cohort employment from the mean are quite persistent, contrary to such deviations of aggregate employment. In particular the correlation between log cohort employment at birth and log cohort employment after four years is around 0.75, while the correlation between log aggregate employment at time t and time $t+4$ is around -0.6. In the Swedish data we calculate the correlation between log cohort employment at birth and at age 4 after demeaning each variable at the county level. The relatively high resulting correlation of more than 0.5 is shown in Figure 2.8. Differentiating between different types of entrepreneurs shows that the correlation is particularly high for the group involving previously unemployed, followed by that previously out of the labor force. Except for a drop from the third to the fourth quintile, the correlation increases by income level, indicating that for firms created by low income entrepreneurs, total employment is less persistent than for those created by high income entrepreneurs.

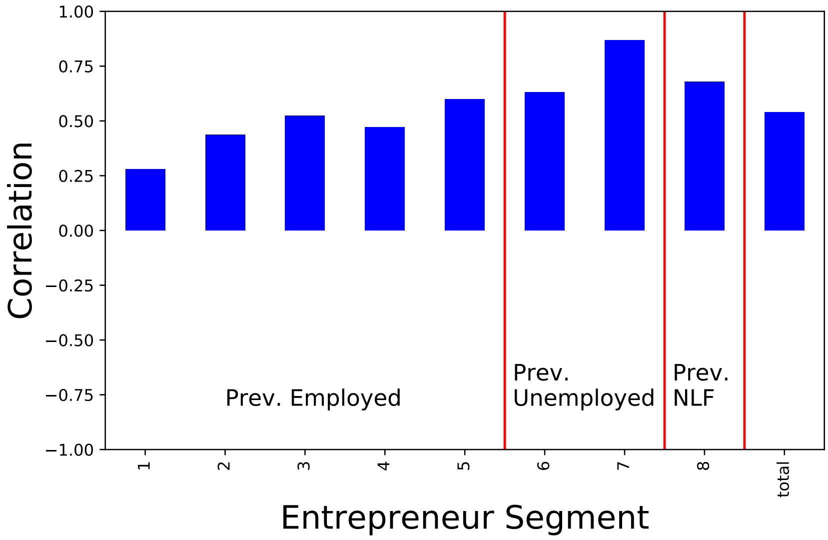


Figure 2.8: Correlation between cohort employment at birth and after four years.

Changes in total cohort employment can be caused by changes in the number of firms or changes in average firm size. While both margins exhibit persistence, the degree of persistence in the number of firms both overall as well as for each entrepreneur segment is substantially higher. As we show in Figure 2.9, the correlation between the number of firms at startup and the number of firms after four years is above 0.7 for all segments.

This finding reflects that convergence between cohorts is exclusively caused by higher exit rates of cohorts with a high initial number of firms. Furthermore the decision to exit is irreversible, so that temporary shocks should have a more limited impact on exit decisions than e.g. on hiring and firing decisions.

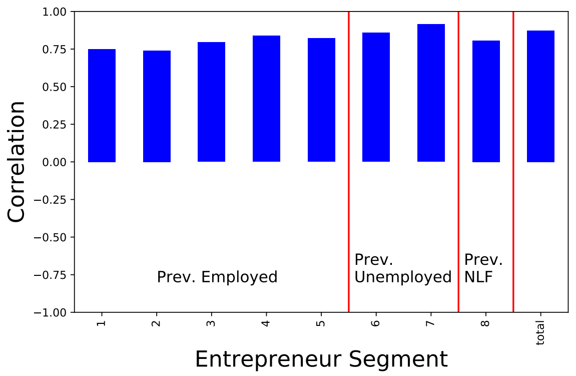


Figure 2.9: Correlation between a cohort's number of firms at birth and after four years.

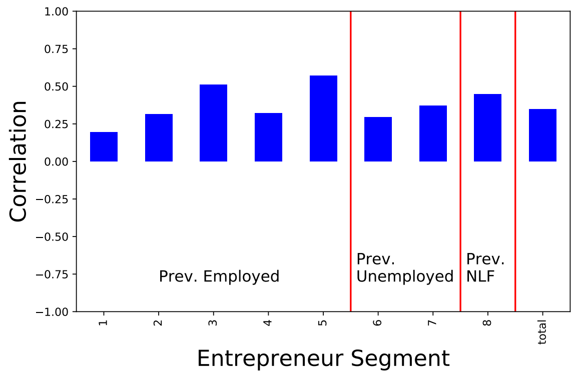


Figure 2.10: Correlation between a cohort's firm size at birth and after four years.

Average firm size however can change both due to cohorts with initially many firms shedding more jobs or due to cohorts with initially few firms creating more jobs. While hiring and firing is associated with adjustment costs, both decisions are not irreversible and should react more to transitory shocks. Consistent with this, we find in Figure 2.10 that the persistence of average firm size, the other margin of total cohort employment, is lower. However, overall persistence is still substantial with a correlation coefficient of more than 0.3 and particularly so for entrepreneurs coming from the highest income quintile, for whom the correlation between initial firm size and firm size at age four is more than 0.5. These results suggest, consistent with the literature that firm quality is persistent. That it is particularly persistent for high-quality entrepreneurs, suggests that among these entrepreneurs permanent characteristics play a more relevant role than transitory shocks which might affect average firm size of lower-quality entrepreneurs.

2.6.2 Sources of variation of the employment impact of cohorts

Do cohorts that create a lot of employment do so because they contain a larger number of firms, or because the firms born in these cohorts are larger. To answer this question we can decompose the variance of cohort employment into a part attributable to variation in the number of firms and another part attributable to variation in average firm size. In particular we can use the following relationship between the total employment $E_{c,a,r}$ of a cohort c at age a in region r, the number of firms $F_{c,a,r}$ of cohort c at age a and the average firm size $e_{c,a,r}$ of cohort c at age a:

$$\log(E_{c,a,r}) = \log(F_{c,a,r}) + \log\left(\frac{E_{c,a,r}}{F_{c,a,r}}\right) = \log(F_{c,a,r}) + \log(e_{c,a,r}) \tag{2.3}$$

This allows us to decompose the variation of $\log(E_{c,a,r})$ into two components:

$$var\left(\hat{E}_{c,a,r}\right) = cov\left(\hat{E}_{c,a,r}, \hat{F}_{c,a,r}\right) + cov\left(\hat{E}_{c,a,r}, \hat{e}_{c,a,r}\right) \tag{2.4}$$

where hats denote log deviations from the region by age mean, which we further demean at the annual level to mitigate distortions introduced by the time series break in 2004.[15] The first component is the part of the variation in total cohort employment that is due to variation in the number of firms. The second component is the one due to variation in average firm size.

In Figure 2.11 we plot the fraction of the total variation in cohort employment at age 0 that is due to variation in the number of firms. We do so for each different type of entrepreneur as well as for all entrepreneur types together (the bar labeled total). For example, around 80% of the variation in employment for cohorts of firms started by entrepreneurs from group 1, i.e. entrepreneurs who were employed and earned an income in the lowest income quintile in the previous year. On the other hand only 30% of the variation in employment created by cohorts of firms started by entrepreneurs from group 5, i.e. those who were employed in the previous year and had an income in the highest income quintile, is due to variation in the number of firms. On aggregate, as can be seen from the last bar, around 60% of the variation in employment created by new firms is due to variation in the number of firms. This compares to 40% in the US, when doing a similar decomposition at the state level.[16]

The remaining part of the variation in total cohort employment that is attributable to variation in firm size is plotted in Figure 2.12. The variation in employment for cohorts of firms started by entrepreneurs who did not come out of employment or who come out of low paid employment is only to a small extent driven by variation in the size of firms started. The fraction is well below 30% and for most groups below 20%. The fraction of variance that is explained by variation in firm size is much higher for cohorts of firms started by entrepreneurs starting up out of employment. In particular for entrepreneurs who were in the highest income quintile, about 70% of the variation in employment is due to variation in the size of firms. This finding is consistent with Sedláček and Sterk (2017) who argue that in the US data variation in average firm size is mostly driven by variation in the upper tail of the distribution. Firms with high growth potential, started primarily

[15] This procedure is akin to using county and year fixed effects.

[16] For the US due to the long time series of 35 years, we look at log deviations from a HP filtered state trend instead of log deviations from the mean.

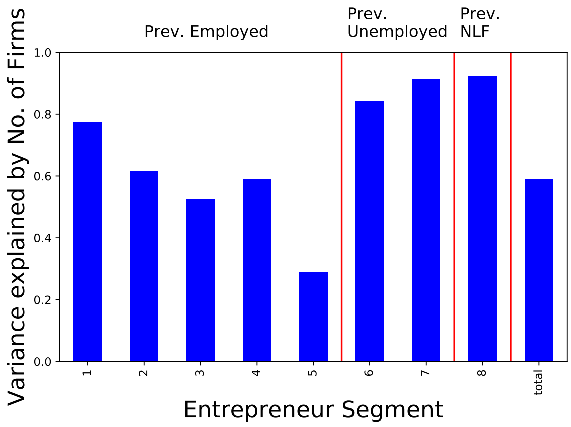


Figure 2.11: Fraction of variance of cohort employment at birth explained by variation in the number of firms

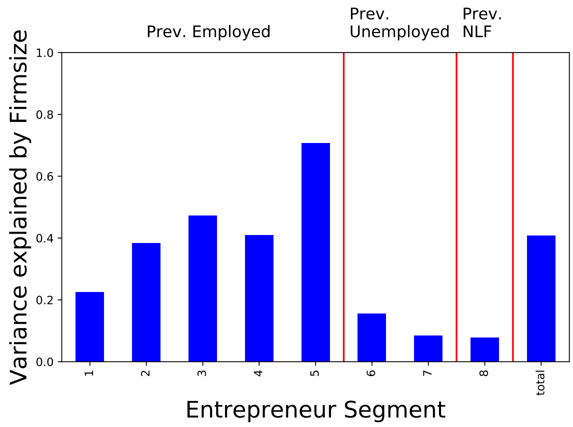


Figure 2.12: Fraction of variance of cohort employment at birth explained by variation in firm size

by high-quality entrepreneurs, might be more sensitive to transitory shocks than firms with low growth potential which are started by the low-quality entrepreneurs.

Similar to Sedláček and Sterk (2017) we find that variation in average firm size is responsible for a larger part of the variation in total cohort employment for more mature firms. In particular, we decompose the cohort level employment variation after four years, similar to the decomposition at birth. In Figure 2.13 we see that only around 30% of the variance in total cohort employment after four years is explained by variation in the number of firms. The remaining 70% are explained by variation in average firm size as shown in Figure 2.14. The pattern among entrepreneur types is similar to the pattern for variation of employment at birth but slightly less pronounced. While variation for the lower quality entrepreneur types, especially for those who come out of longer term unemployment is mostly due to variation in the number of firms, for the higher quality entrepreneur types most of the variation is due to variation in firm size.

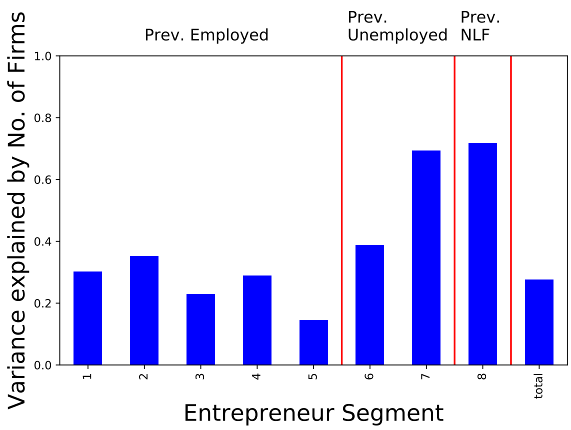


Figure 2.13: Fraction of variance of cohort employment at age 4 explained by variation in number of firms

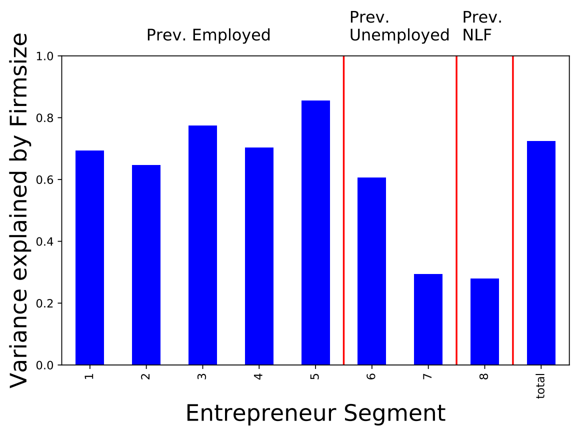


Figure 2.14: Fraction of variance of cohort employment at age 4 explained by variation in firm size

2.6.3 Sources of variation of the average firm size

Given the findings from Section 2.5 that initial firm size varies substantially with the type of entrepreneur starting the firm, shifts in the composition of founders over the business cycle might play an important role for explaining variation in average firm size. For example when the share of entrepreneurs coming from higher income deciles is higher, we expect an increase in the average firm size. However, variation in firm size might also be due to variation within segments. This seems also quite plausible as in Figure 2.12 we already documented that a substantial part of the variation in the employment impact of firms started by a specific entrepreneur segment is due to variation in the average size of firms started by this segment. Letting g index the different entrepreneur segments and $s_{c,a,r,g}$ the share of firms that were started by entrepreneurs of segment g in cohort c at age a in region r, we can rewrite the average size of firms as follows:

$$e_{c,a,r} = \frac{E_{c,a,r}}{F_{c,a,r}} = \frac{\sum_g E_{c,a,r,g}}{F_{c,a,r}} = \sum_g \frac{E_{c,a,r,g}}{F_{c,a,r,g}} \frac{F_{c,a,r,g}}{F_{c,a,r}} = \sum_g e_{c,a,r,g} s_{c,a,r,g} \tag{2.5}$$

This allows us to do a simple shift-share decomposition of deviations from the region-age specific mean[17]

$$\begin{aligned} e_{c,a,r} - \bar{e}_{a,r} = \sum_g \Bigg[& \underbrace{(e_{c,a,r,g} - \bar{e}_{a,r,g})\bar{s}_{a,r,g}}_{(1)} + \underbrace{(\bar{e}_{a,r,g} - \bar{e}_{a,r})(s_{c,a,r,g} - \bar{s}_{a,r,g})}_{(2)} \\ & + \underbrace{(e_{c,a,r,g} - \bar{e}_{a,r,g})(s_{c,a,r,g} - \bar{s}_{a,r,g})}_{(3)} \Bigg] \end{aligned} \tag{2.6}$$

In this decomposition (1) measures deviation due deviations of current average firm size from the mean average firm size within group, (2) measures deviation due to deviation in the share of firms by each group and (3) measures a covariance term. Denoting deviations from the group, age, region specific mean by hats, we decompose the variance decomposition according to:

$$var(\hat{e}_{c,a,r}) = \sum_g \Bigg[\underbrace{cov_{r,c}(\hat{e}_{c,a,r}, \bar{s}_{a,r,g}\hat{e}_{c,a,r,g})}_{(A)} + \underbrace{cov_{r,c}(\hat{e}_{c,a,r}, \hat{e}_{a,r,g}\hat{s}_{c,a,r,g})}_{(B)}$$

[17] For the derivation see Appendix B.4

$$+\underbrace{cov_{r,c}(\hat{e}_{c,a,r}, \hat{e}_{c,a,r,g}\hat{s}_{c,a,r,g})}_{(C)}\Bigg] \quad (2.7)$$

In this variance decomposition, (A) measures the part of the variation in average cohort firm size that is due to variation of firm size within entrepreneur type groups, (B) measures the part that is due to variation in the entrepreneur type shares and (C) measures a covariance term that is usually small. Furthermore this decomposition allows us to show how much of the variance in average firm size is driven by correlations between average firm size and the individual entrepreneur segments. The variance decomposition shows that 86.5% of the variation in initial firm size is driven by variation within segments while 13.5% is due to variation in the share of segments. The covariance term is negligible. When regions are weighted by their average employment, the share of variation that is due to within segment variation in the average firm size decreases to 74% and the share that is due to variation in the composition of entrepreneurs increases to 23.5%. The entrepreneur segment that is responsible for most of the variation in either case are the entrepreneurs who come from the highest income quintile. Within that segment there is a substantial amount of variation in average firm size over time. This is consistent with the results from Sedláček and Sterk (2017) that variation in the upper part of the startup size distribution is responsible for most of the variation in average firm size.

2.6.4 The cyclicality of the employment impact of new firms

In the previous sections we uncovered substantial heterogeneity in the sources of variation of the employment impact of startup cohorts, depending on the entrepreneur type. Whereas total employment of firms started by high quality entrepreneurs varies mostly due to variation in the average size of these firms, for low quality entrepreneurs most of the variation stems from variation in the number of new startups. Furthermore variation in average startup size is to the most part due to variation in startup size of firms created by high income entrepreneurs but also to roughly one quarter due to variation in the composition of entrepreneurs. An important question is whether this variation is cyclical and if so how it contributes to the cyclicality of total startup employment which we analyzed in Section 2.4.2. In particular, considering the huge differences in firm characteristics between the firms started by the different entrepreneur types, variation in the shares of firms started by each entrepreneur type over the business cycle could result in sizable variation in the total employment impact of recessionary vs. expansionary cohorts.

To analyze the cyclicality by entrepreneur type, we run the regression specification for column (6) from Section 2.4.2 separately for the different entrepreneur segments:

$$Y_{t+a,a,m,e} = \beta_1 \frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}} + \beta_2 u_{t,m} + \gamma_m + \delta_t + \epsilon_{t,m} \tag{2.8}$$

where $Y_{t+a,a,m,e}$ is outcome variable for firms of age a which were started in year t in county m by entrepreneurs of type e. $\frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}}$ the employment growth rate from year $t-1$ to t in county m. $u_{t,m}$ is the unemployment rate in year t in county m, γ_m is a county fixed effect and δ_t is a year fixed effect to flexibly control for a time trend and a time series break in 2004.

In the following figures, the blue column displays the coefficient β_1 i.e. the estimate of the elasticity with respect to county level employment growth in the year of birth. The orange column displays the coefficient β_2 i.e. the estimate of the elasticity with respect to the local unemployment rate in the year of birth. The thin black bar shows the 90 percent confidence intervals for each point estimate.

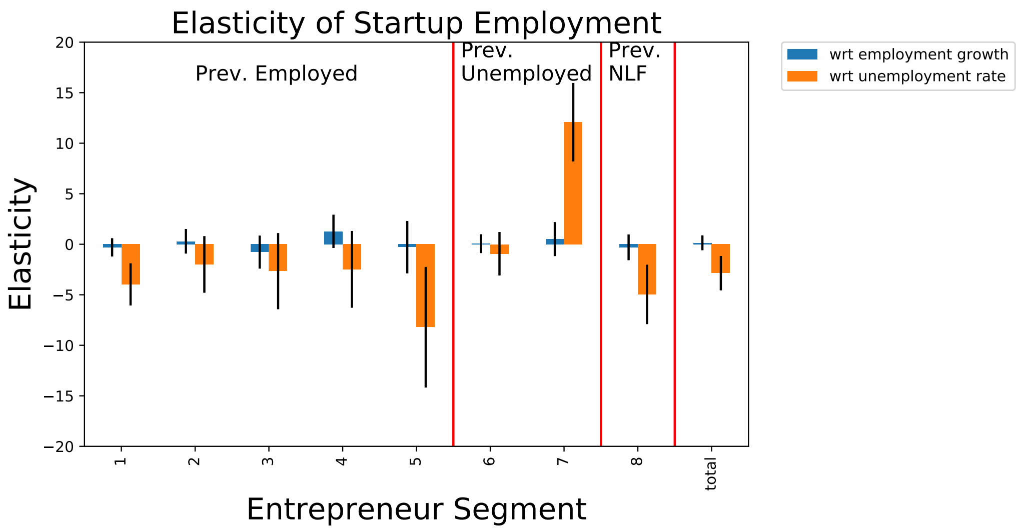


Figure 2.15: Partial elasticity of total cohort employment at birth with respect to employment growth and the unemployment rate.

We first consider the cyclicality of total employment at birth $\log(Emp_{t,0,m,e})$ in Figure 2.15. Underlying the procyclicality of employment at birth of the total cohort is considerable heterogeneity across entrepreneur groups. The employment of firms started by entrepreneurs coming out of employment is procyclical i.e. it decreases as the unemployment rate increases. On the other hand employment of firms started by entrepreneurs coming out of unemployment increases in times of high unemployment for the long term unemployed or remains almost the same for the short term unemployed.

The heterogeneity of the cyclicality in terms of total employment is also reflected in the cyclicality of the number of firms started by each entrepreneur group $\log(Firms_{t,0,m,e})$, which is documented in Figure 2.16. While the number of firms started by entrepreneurs who come out of employment decreases when the unemployment rate is high, the op-

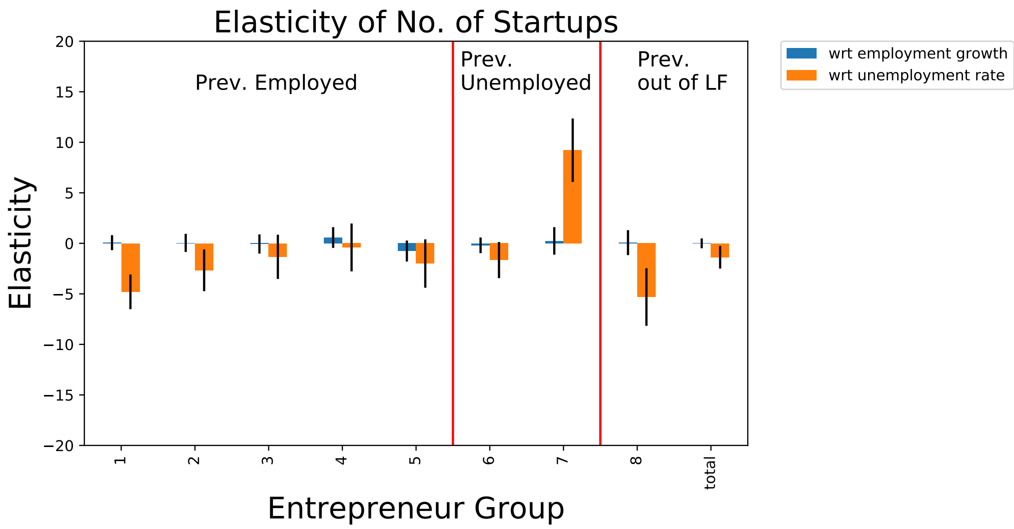


Figure 2.16: Partial elasticity of the number of firms at birth with respect to employment growth and the unemployment rate.

posite is true for entrepreneurs who come out of unemployment. So in times of high unemployment, there is a clear shift in the composition of newly started firms towards firms started by previously unemployed entrepreneurs. In light of the differences in terms of firm size between the different entrepreneur types, this compositional shift contributes to the procyclicality of average startup size.

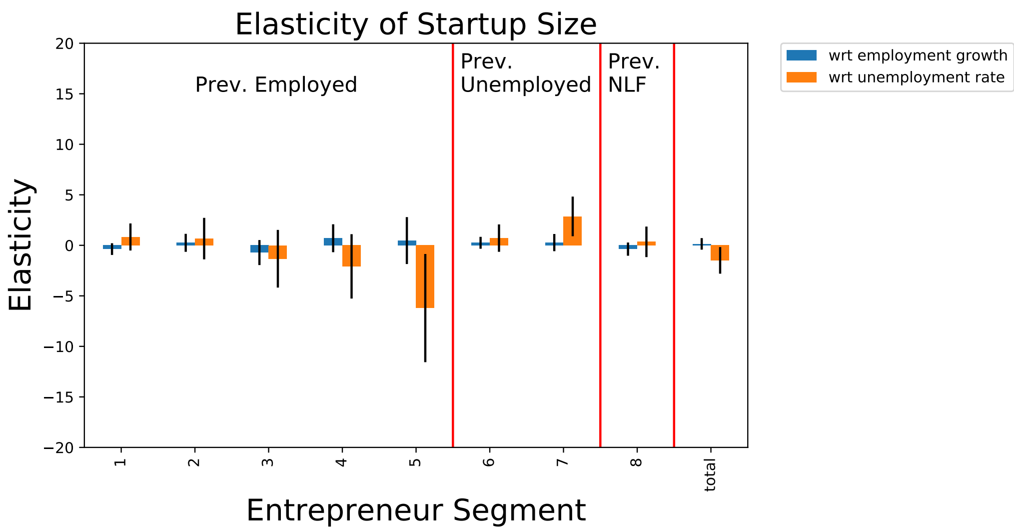


Figure 2.17: Partial elasticity of the average firm size at birth with respect to employment growth and the unemployment rate.

However, average firm size at birth, $\log\left(\frac{Emp_{t,0,m,e}}{Firms_{t,0,m,e}}\right)$, also exhibits cyclical variation within the different firm groups. In particular, the average size of firm started by high quality entrepreneurs declines substantially when the unemployment rate is high. This result is consistent with the findings from Sedláček and Sterk (2017) that it is primarily the procyclicality of the size of large firms which drives the procyclicality of average firm size at birth. Interestingly the average size of firms started by low quality entrepreneurs is countercyclical, where low quality entrepreneurs are considered those starting from low wage employment, plus those starting from unemployment. One explanation might be that the pool of unemployed improves slightly in bad times when unemployment is high, as not only the lowest quality workers remain unemployed for a substantial amount of time anymore. This effect is not economically significant however, as the baseline average size of these firms is very low as shown in Figure 2.1.

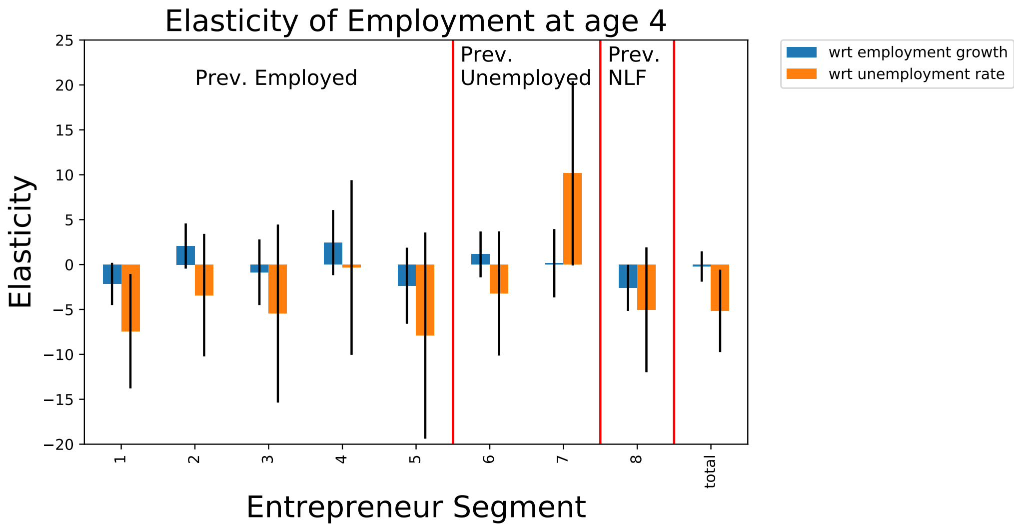


Figure 2.18: Partial elasticity of total cohort employment at age 4 with respect to employment growth and the unemployment rate at birth.

Results on the persistence of these cyclical effects by entrepreneur type mirror those of the short run cyclicality. The elasticities of total employment of firms at age 4 with respect to the unemployment rate at firm birth are similar to those of age 0 employment. The elasticities are negative for entrepreneurs coming out of employment, while they are positive for the lowest quality entrepreneurs. However as firms are faced with additional shocks during their growth phase, these coefficients are much less precisely estimated. Except for the coefficients on employment at age 4 of all firms and of firms started by entrepreneurs from the lowest income quintile, none is significant at conventional levels.

Similarly the estimates of the elasticity of the size at age 4 and the survival rate until age 4 with respect to business cycle conditions at startup are imprecisely estimated. None is significantly different from zero.

2.6.5 The cyclicality of exit and survival

We hypothesize that firms started by entrepreneurs out of unemployment are mostly started due to lack of good alternatives i.e. a sort of necessity entrepreneurship. This should be particularly true for firms started in times of high unemployment. However

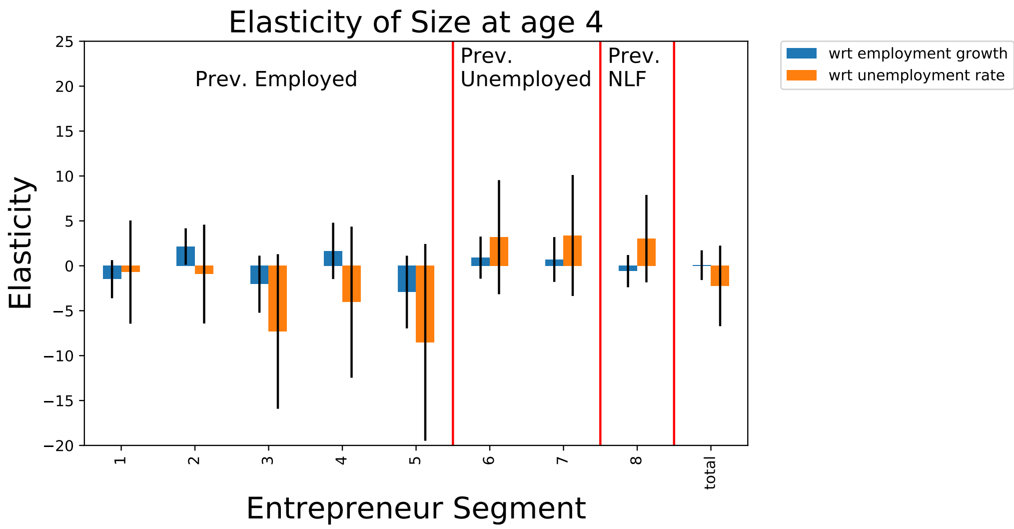


Figure 2.19: Partial elasticity of the average firm size at age 4 with respect to employment growth and the unemployment rate at birth.

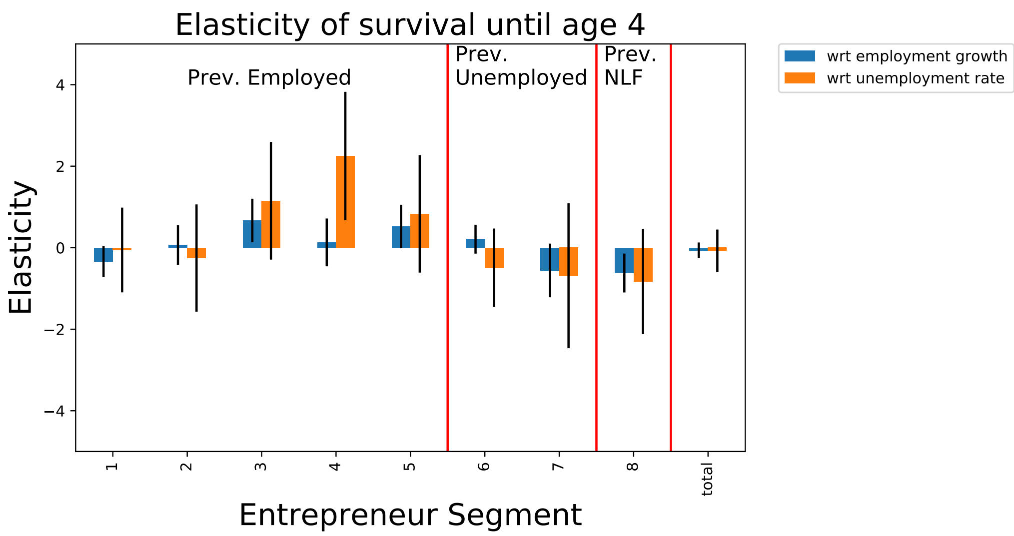


Figure 2.20: Partial elasticity of the survival rate until age 4 with respect to employment growth and the unemployment rate at birth.

once the labor market conditions improve this should imply that these entrepreneurs close down their firms and move back towards dependent employment. The flipside of this hypothesis is that people who start firms out of dependent employment do so because they spot a good business opportunity. For their exit decision it should be less relevant how the labor market evolves as it mostly depends on the success of their own business and not so much on the outside options available in the market for dependent employment. We can test these hypotheses using the following regression of exit rates on business cycle / labor market conditions:

$$exit_{t,a,m,e} = \beta_1 \frac{Emp_{t,m} - Emp_{t-1,m}}{Emp_{t-1,m}} + \beta_2 u_{t,m} + \gamma_m + \delta_t + \xi_a + \epsilon_{t,m} \tag{2.9}$$

where we regress the exit rates of cohorts in year t and region m at age a on the current employment growth rate and the current unemployment rate in year t in region m. We also control for region, year and firm age fixed effects. As above we perform the regressions separately for each entrepreneur type e and for the total cohort. The results of these regressions are shown in Figure 2.21 where the blue column represents β_1 and the orange column β_2. There is no clear pattern. Furthermore all the coefficients are insignificant except β_2 for all firms which is significant at the 10 percent level.

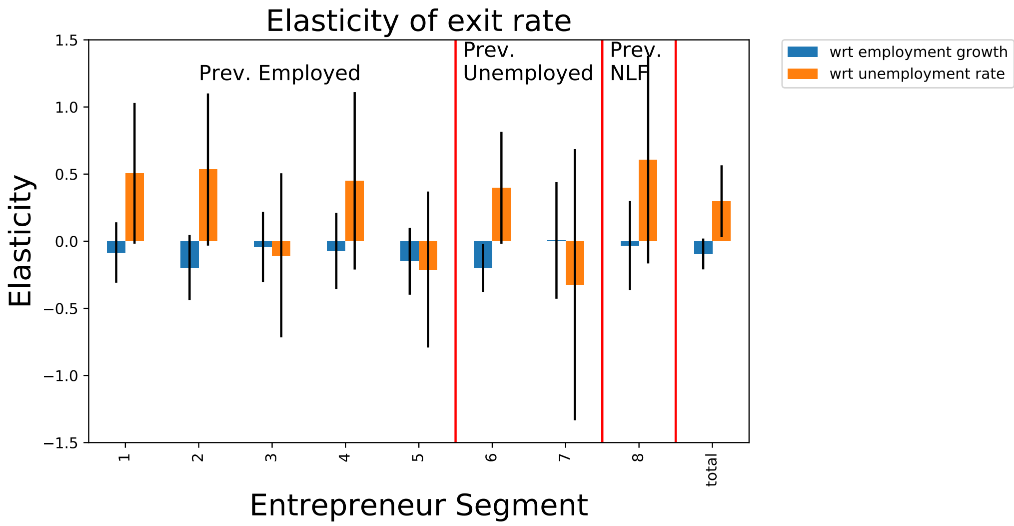


Figure 2.21: Elasticity of the exit rate with respect to current employment growth and the current unemployment rate.

We can perform a similar exercise by replacing the exit rates in the regression above by survival rates since birth. The results of this exercise are shown in Figure 2.22. Here we see some slightly clearer evidence for a negative relationship between the current unemployment rate and survival rates of firms from entrepreneurs initially coming out of employment, while there is no negative relationship for firms started by entrepreneurs coming out of long-term unemployment. While the small positive elasticities with respect to employment growth are not statistically significant, the negative elasticities with respect to the current unemployment rate are significant for groups 1, 4, 6 and total. Similar effects result when we also control for the conditions at birth as shown in Figures B.7 and B.8.

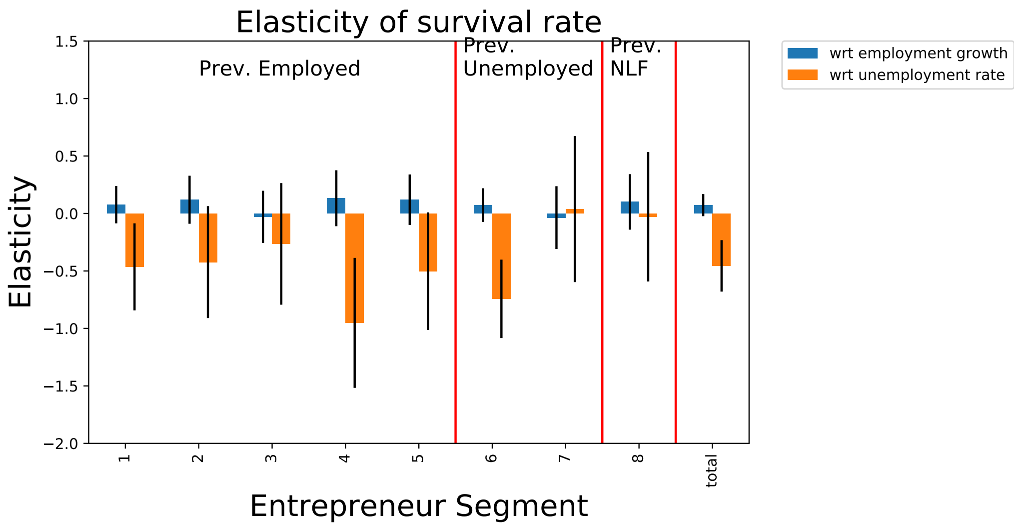


Figure 2.22: Elasticity of the survival rate since birth with respect to current employment growth and the current unemployment rate.

2.7 Extensions to the empirical analysis

2.7.1 Identifying founders and founder quality using microdata

Identifying founders pre 2004 Statistics Sweden identifies the operative CEO of a firm, whom we classify as the firm's founder, starting in 2004. Prior to 2004 we have to mimic the methodology of Statistics Sweden, which is problematic as the definition of the "entrepreneur status" variable, which is the main predictor in the Statistics Sweden methodology, changed in 2004. The fact that we have additional information on individual characteristics, on firm characteristics and on the characteristics of coworkers would allow us to pursue a data-driven approach that does not rely on the "entrepreneur status" variable. In particular we can fit an econometric model on the post 2004 data that assigns to each individual, among all individuals employed at new firms, a probability of being the founder of a new firm. We can then use the estimated model to assign founder probabilities to each individual employed at a new firm in the pre 2004 data. For each new firm we then designate the individual with the highest founder probability as the founder of the firm.

This prediction problem is a simple binary classification problem which can be approached using a probability model e.g. a linear probability model or a logistic regression. Thus we run a model on the post 2004 data, for which Statistics Sweden consistently classifies the operative CEO, using

$$Y_{ift} = \begin{cases} 0 \text{ if not entrepreneur} \\ 1 \text{ if entrepreneur} \end{cases}$$

as the outcome variable. As independent variables we use entrepreneur specific characteristics X_{it}, such as education, prior labor force status, prior income, income rank within the firm and age, individually as well as interaction terms with firm specific characteristics Z_{ft}, such as firm size or sector.

Interactions with firm characteristics account for the fact that some individual characteristics such as having the highest income among all employees in a firm might be a good predictor of being the firm owner in some firm types but not in others. A more flexible way of approaching this binary classification problem, which accounts for the heterogeneity across different segments of the population, would be to use regression or classification trees which segment the population and run regression models within segments, respectively classify each segment.[18] Segmentation in this approach is purely data driven but might for example segment the population by firm size. For example simply running a regression on the overall data possibly puts a very high weight on having the highest income in a firm for determining ownership status as by construction all workers in single worker firms are entrepreneurs and have the highest income within that firm. Segmenting the observations by firm size might show that for larger firms, having the highest income among employees is not a particularly strong predictor of being the firm owner.

Identifying founder quality We documented significant heterogeneity in the performance of firms, in terms of average employment and survival, between firms started by entrepreneurs with different prior labor force status. The analysis of variation in firm performance over time, however, showed significant remaining variation within these groups. So while the prior labor force status of an entrepreneur seems to be an important determinant of firm performance, other entrepreneurial characteristics will as well be important determinants of firm performance. As we observe firm performance we can use the joint distribution of observed firm performance and entrepreneurial characteristics to identify the characteristics of entrepreneurs starting high-quality firms.

This is a classification problem with a structure similar to the identification of founders among individuals working in a new firm. However, firm quality has to be defined first. As initial firm size as well as incorporation status are highly correlated with subsequent firm size and survival, a simple way of determining firm quality, which only uses characteristics at birth, would segment firms according to quantiles of the firm size distribution at birth and the incorporation status. Alternatively and similarly to Pugsley et al. (2018), one could use data on the first five years of a firm's life and define high quality firms using the definition of "gazelles" as firms with an annual growth rate of at least 20% in the first five years, and an employment level that exceeds 10 employees at some point of their lifetime.

[18]For an exposition of regression and classification trees as well as random forests see Hastie et al. (2009) As we are only interested in predictive performance and not in interpretability the preferred method would be to use a random forest which combines the predictions of multiple trees.

Conditional on an operational definition of firm quality, one of the classification methods described above can be used to identify the characteristics of high quality founders. In order to perform the aggregate analyses in Sections 2.5 and 2.6 we require a segmentation of the population which is automatically provided by a classification tree. Probability models also allow for a segmentation of the population for example by segmenting the population according to the deciles of the distribution of the estimated probability of starting a high quality firm.

Outside option vs. other founder characteristics In a similar spirit, regression analyses at the level of individual observations using entrepreneur and firm characteristics more intensively can be used as predictors of firm quality, to inform us on the relative importance of the outside option as proxied by prior labor force status. Regressing a firm characteristic variable $e_{f,t}$ of firm f in year t, in particular firm size, on dummies for the "outside option" segment, year, region, industry, and the education of the entrepreneur would give us the effects of the outside option conditional on other relevant observables. Let f denote the firm, t the year, $r(f)$ the region in which firm f is active, $i(f)$ the founder of firm f. Then regression specifications would look like the following:

$$e_{f,t} = \alpha_t + \beta_{r(f)} + \sum_s \beta_s segment_{i(f)} + \sum_{ind} \beta_{ind} industry_f + \sum_{ed} \beta_{ed} educ_{i(f)} + \epsilon_{f,t} \quad (2.10)$$

The relevant effects of the outside option, conditional on other entrepreneur characteristics, are measured by the coefficients β_s. Alternatively we can start from the segmentation of the population into 8 groups by prior labor force status and could further segment each group either by educational status (high school and less vs. post-secondary and more) and/or by industry (services vs. manufacturing). This approach would be similar to the approach delineated in the prior section but with a segmentation determined by prior knowledge about which entrepreneur and firm characteristics are important for firm quality instead of a purely data-driven segmentation of the population.

2.7.2 Investigating cyclicality using microdata

Moreira (2016) has recently investigated the cyclicality of firm characteristics using firm and establishment level microdata. While we so far use data aggregated at the county by entrepreneur segment level, we can also use the individual firm level observations and add entrepreneur characteristics to the specifications proposed in the prior literature. This substantially increases the number of observations and allows us to control for additional firm and entrepreneur characteristics. The basic regression without entrepreneur characteristics is given by:

$$\ln e_{f,ctr} = \alpha_r + \gamma_t + \delta_a + \beta u_{c,r} + \xi^{firm} X^{firm}_{f,t} + \epsilon_{f,ctr} \tag{2.11}$$

which relates log firm size $\ln e_{f,ctr}$ of firm f from cohort c in year t and region r to a region fixed effect, a year fixed effect, an age ($a = t - c$) fixed effect and additional firm characteristics $X^{firm}_{f,t}$ as well as the unemployment rate $u_{c,r}$ in region r at entry. Other time varying regional characteristics, such as the industry mix reflecting local supply conditions, could be included as well. At any rate, the (semi)-elasticity of size with respect to the unemployment rate[19] is measured by the coefficient β. Additional firm characteristics to be included are for example incorporation status and industry sector which could also systematically vary over the business cycle.

Entrepreneur characteristics as controls A straightforward way to investigate whether systematic shifts of the entrepreneur composition over the cycle affect the cyclicality of firm size is to include characteristics $X^{ent}_{i(f),t}$ of entrepreneur $i(f)$ who runs firm f in the regression specification:

$$\ln e_{f,ctr} = \alpha_r + \gamma_t + \delta_a + \beta u_{c,r} + \xi^{firm} X^{firm}_{f,t} + \xi^{ent} X^{ent}_{i(f),t} + \epsilon_{f,ctr} \tag{2.12}$$

Entrepreneur characteristics to be included could be variables such as prior labor force status, prior income, labor market experience, gender or education. Dummy variables indicating the entrepreneur's population segment, as defined in section 2.5, could also be included. If entrepreneurial characteristics that are relevant for firm quality vary systematically over the cycle, the elasticity of firm size with respect to the unemployment rate β should differ significantly between the specification including and excluding entrepreneur characteristics, respectively. For example, if entrepreneur types that systematically start larger firms, such as entrepreneurs with high previous income, have higher entry rates in times of low unemployment the estimate $\hat{\beta}$ would be downward biased when entrepreneur characteristics are not included.

In addition, interaction terms between the cycle indicator and entrepreneur characteristics could be included to answer the question whether different entrepreneur types exhibit different degrees of cyclicality of firm characteristics. Differences in cyclicality across entrepreneur groups are suggested by the aggregate analysis in Section 2.6. At any rate, the additional specification suggested here accounts for hetereogeneity in the composition of entrepreneurs over the cycle and for heterogeneity in the entrepreneur type

[19] Other business cycle indicators which are discussed in Appendix B.5 could also be used in this analysis.

specific cyclicality of firm size:

$$\ln e_{f,ctr} = \alpha_r + \gamma_t + \delta_a + \beta u_{c,r} + \beta^{ent} u_{c,r} X^{ent}_{i(f),t} + \xi^{firm} X^{firm}_{f,t} + \xi^{ent} X^{ent}_{i(f),t} + \epsilon_{f,ctr} \quad (2.13)$$

In this specification the unemployment rate at birth is interacted with entrepreneur characteristics and the vector β^{ent} measures differences in the cyclicality across different entrepreneur types.

Within-segment regressions A different approach to investigate heterogeneity in the cyclicality across entrepreneur types is to follow the aggregate analysis in Section 2.6 and perform the regressions specified in equations (2.11) and (2.12) separately by entrepreneur segment. This approach accounts for the possibility that the relationship between entrepreneur, respectively firm characteristics characteristics and firm size might differ by entrepreneur segment. For example, high quality entrepreneurs might have a particular advantage in starting firms in specific sectors which would be reflected by differences in sector specific fixed effects across entrepreneur segments. If this were the case, variation in the industry composition over the cycle could bias the estimated cyclicality of startup characteristics in specification (2.13). Estimating the regression specification separately on each entrepreneur segment yields segment specific effects of firm characteristics and avoids this bias.

2.7.3 The role of demand shocks

While we use regional labor market conditions at startup as the main business cycle indicator, the importance of demand shocks for startup decisions was stressed in the recent literature. Demand shocks arguably affect the composition of startups by affecting the relative value of starting growth oriented vs. non growth oriented firms. In order to operate profitably, growth oriented firms need to scale up which becomes more difficult in times of low demand. Consequently the relative value of starting a growth oriented firm decreases in such times. As growth oriented firms are mostly started by high quality entrepreneurs, demand shocks might also affect the composition of entrepreneurs. This is also the idea in Bernstein et al. (2018) who show that shifts in local demand as caused by variation in the value of local agricultural products slightly affect the composition of entrepreneurs. Sweden's high export share between 40% and 50% since 1997 allows to use shifts in world demand as exogenous demand shocks. Measures of world import demand have been used to construct firm level shocks using product level production data for example in Hummels et al. (2014) and Friedrich (2015).

As we only observe firms' industry but not their firm specific products, we instead follow Tåg et al. (2016) and construct world import demand for each industry in Sweden.

We construct the world import demand WID_{jt} for goods produced by industry j in year t using import demand from country c for product k in year t:

$$WID_{jt} = \sum_k \sum_c s_{jck} WID_{ckt}$$

where s_{jck} is the share of Swedish exports of good k by industry j to country c in total production value of Swedish industry j. As is common in the literature we fix the export shares at the value of a base period. We calculate s_{jck} for the years 1995 to 1999 and take the average value over these years.

World import demand shocks by industry can be used together with the county level industry structure to construct county by year specific demand shocks. Letting $W_{rj,95-99}$ the average wage bill of industry j in county r between 1995 and 1999 and $wid_{j,t} = WID_{j,t}/WID_{j,95-99}$ world import demand for goods produced by industry j in year t relative to average world import demand for goods from that sector between 1995 and 1999, the local demand shock can be constructed as

$$D_{r,t} = \sum_j (wid_{j,t} - 1) W_{rj,95-99}. \tag{2.14}$$

This shock proxies the change in local wage income relative to the average between 1995 and 1999 due to changes in world import demand and thus acts as an exogenous shifter of local demand. However, this measure of demand variation does only account for variation in demand for internationally traded goods and services. The production of these goods and services also uses intermediates that are not traded internationally. In order to take into account indirect demand effects of increases in final demand for goods from industry j' on intermediate input suppliers we can specify the demand shock using information on the input output structure of the Swedish Economy. Let $i_{j,j'}$ the share of intermediates from industry j in final demand of industry j' accounting for all linkages along the supply chain. Using this information we can construct the sector specific demand shock accounting for supply chain linkages as

$$\widetilde{WID}_{j,t} = \sum_{j'} i_{j,j'} WID_{j',t}.$$

This shock can either be used to construct local demand shocks as in equation (2.14) or as an industry specific demand shifter with which we can run analyses at the industry by year level instead of the region by year level.

2.7.4 Capital structure and financial constraints

Starting a new firm can require sizable capital investments, the amount of which varies by firm type. Bernanke and Gertler (1989) as well as Carlstrom and Fuerst (1997) have proposed varying financial constraints due to varying agency costs over the business cycle as an important determinant of entrepreneurs' investment decisions. Growth oriented firms might require relatively higher capital investment than non-growth oriented firms. Varying financial constraints over the business cycle thus could lead to substantial variation in the composition of startups over the cycle. Furthermore, if high quality entrepreneurs are more likely to start firms which require external finance, varying financial conditions could also affect the composition of entrepreneurs. Counteracting effects could be that wealth varies substantially across the would-be entrepreneurs. Using data on the capital structure of all firms in Sweden from 1998 to 2011 as well as on the wealth of individuals from 1999 to 2007, we can analyse the importance of these channels.

We can examine the capital structure of firms started by different entrepreneur types along with their wealth endowments, as well as the relationship between capital structure and firm performance. Furthermore our data allows us to investigate the cyclicality in the capital structure decisions of different entrepreneur types; for example whether firms started in downturns started with lower leverage or with lower total debt and equity. In conjunction with the relationship between capital structure decisions and firm growth allows us to empirically investigate whether cyclically varying financial constraints affect the composition of startups over the cycle.

2.8 Concluding remarks

In this paper we investigate the relationship between variations in entrepreneur characteristics and startup characteristics in the cross section and over time. Using Swedish linked firm employee microdata which allow us to identify owners of firms we show that entrepreneur characteristics, in particular their labor force status prior to starting a firm, substantially affect startup employment and survival. We divide the Swedish population into different segments according to their prior labor force status, subdividing previously employed into segments according to their previous earnings, the unemployed into segments according to their unemployment duration and those out of the labor force according to their education. We document substantial heterogeneity in terms of employment within the group of previously employed entrepreneurs, as entrepreneurs with higher previous income start significantly larger firms. The firms started by entrepreneurs with previously low income do not differ substantially from firms started by the previously unemployed, highlighting that a simple division of entrepreneurs into the previously employed and the previously unemployed is not very meaningful.

After establishing substantial differences between firms started by different entrepreneur segments in the cross section we investigate whether variation in the composition of entrepreneurs is an important driver of variation in employment created by startups. We document that 40% of the variation in employment created at birth is due to variation in the average size of startups. Decomposing variation in average firm size into variation within entrepreneur segments over time and into variation in the composition of entrepreneurs shows that between 13.5% and 23.5% of variation in startup size is due to variation in the composition of entrepreneurs. Controlling for additional characteristics like education or industry of the startup, as suggested in the empirical extensions, is likely to increase the contribution of compositional variation.

Employment creation by startups and its main margins, the number of firms and the average firm size exhibit a similar cyclicality in Sweden as in the US. Employment creation by startups falls in times of high unemployment both due to a significant decrease in the number of startups and their average size. Disaggregating these margins by entrepreneur segment, we document a significantly positive relation between the unemployment rate at birth and startup employment created out of unemployment, which is the result of a higher number of firms with slightly higher average size. On the other hand startup employment created by previously employed entrepreneurs significantly declines, particularly for entrepreneurs with high previous income. The number of firms started by these high income entrepreneurs does not change significantly but their average size decreases substantially. The significant decline in the average size of startups on aggregate is thus due to two effects, on the one hand a higher share of firms started by low quality entrepreneurs and on the other hand a decrease in the size of firms started by high quality entrepreneurs.

Further analysis, along the lines suggested in the last part of this paper is needed to shed lights on the mechanisms driving these results. A further disaggregation of entrepreneur segments for example by education or by industry would be beneficial to understand the contribution of ex-ante decisions and characteristics for firm size and growth. Analysis of demand shocks and financial constraints would highlight the main frictions firms and entrepreneurs of different types face. A particular emphasis in this analysis should be put on the question whether these frictions are mainly structural or cyclical and thus inform the question whether structural policies or cyclical policies are more adequate to foster entrepreneurship.

Chapter 3

Demand Learning with Nonperishable Products - Multiunit Firms and Inventory Replenishment

3.1 Introduction

When setting prices in order to maximize revenue, firms - while having a prior belief about the demand they face - are often uncertain about the actual demand for their product. Over time these firms can learn about the potential demand for their product by observing how well the product sells. If the product sells often the firm will revise its belief about the demand upwards, if sales are rare the belief will adjust downwards. Such a firm can control the speed of learning via the price it charges. Customers are heterogeneous in their valuation of the firms products and will buy the product only if the price is less than their valuation. Hence, charging a higher price reduces the number of customers that are willing to buy the product and thus also the speed of learning while a lower price increases the probability of a sale and the speed of learning. Thus learning introduces an additional tradeoff into the price setting problem of the firm: in addition to balancing the revenue from a sale and the probability of a sale, prices also affect the speed of learning. This problem has been considered for a firm trying to sell a single unit of its product while learning about demand by Mason and Välimäki (2011).

In this paper we first adapt the Mason and Välimäki (2011) model to a firm with multiple units in its inventory. In a second step we also consider such a multiunit firm that has the option to replenish its inventory in the case of a sale. Considering a multiunit firm introduces the possibility of upward adjustment of beliefs. Beliefs about demand being high adjust upwards only after a sale. As the problem stops immediately after a sale for a single unit firm, beliefs only adjust downwards in Mason and Välimäki (2011) which implies that the value of learning is unambiguously negative. They show that this implies that prices for a learning firm are higher than those of a non-learning firm as it tries to slow down learning. For a firm with multiple units in its inventory the possibility of upwards revisions in the belief imply that the value of learning could be positive. The goal of this paper is to qualitatively and quantitatively analyze the value of learning for multiunit firms in order to assess whether the positive effect of learning on prices is particular to the assumption of a single unit firm. We qualitatively characterize the optimal price set by the multiunit firm at different beliefs and different inventory levels. In line with intuition we find that the optimal price decreases as the inventory (=supply) increases; also the price increases as the belief about demand being high increases. We then proceed to a comparison between the price set by a firm that updates its beliefs about demand and a firm which keeps beliefs fixed. The countervailing effects on the value of learning complicate this analysis, however we can show that the value of learning is unambiguously negative and a learning firm slows down the sales process by increasing prices. The results from Mason and Välimäki (2011) thus extend to multiunit firms.

After the analysis for the simple multiunit firm without any possibility of production, we consider a firm that can replenish its inventory by one unit after every sale at some cost. This formulation of a production possibility allows us to use inductive arguments in order to qualitatively characterize optimal prices and replenishment probabilities for the firm. For a non-learning firm we find that prices and replenishment probabilities both increase in the belief that demand is high and decrease in the size of the inventory. Also the opportunity to replenish reduces the price charged by the firm which can replenish its inventory vs. a firm which can't. This result is due to convex replenishment costs which make replenishment a profitable opportunity at the optimal replenishment rate and thus increase the benefit of selling. Introducing replenishment also affects the learning opportunities for the firm. Most obvious is the fact that even a firm with a single unit inventory can now experience positive a learning effect when it sells the unit and restocks its inventory. Thus even for a single unit firm the comparison between a learning and a non-learning firm becomes ambiguous as the effect of learning could be positive or negative. We analytically show that for very high or very low demand beliefs the learning effect for such a firm is unambiguously negative. For multiunit firms and for intermediate beliefs we will have to resort to numerical analysis.

In order to analyze the features of the model which we could not characterize analytically in the general case, we present a numerical example. We first look at the value functions of learning and non-learning firms. We find that the value function of a learning firm is convex, consistent with our analytical findings. The non-learning firm has a concave value function which implies that the non-learning value function always lies above the learning firm's value function. This has direct consequences on the relative prices set by the learning and the non-learning firm: identical to the single unit setup of Mason and Välimäki (2011), the learning firm always sets a higher price than the non-learning firm. Furthermore the learning monopolist also chooses a higher replenishment rate than the non-learning monopolist. In order to understand these findings, we have to examine the value of information and the marginal effects of the price and the replenishment rate on the value of information. The downwards drift of beliefs if no sale occurs leads to a positive effect of a price increase on the value of information. Intuitively, an upwards adjustment of the price reduces the informativeness of the event "no sale" and thus slows down the arrival of negative information. On the other hand the possibility of an upwards adjustment of beliefs after a sale leads to a positive effect of a price decrease on the value of information: a lower price increases the probability of a sale and thus the arrival of positive information. While in Mason and Välimäki (2011) only the first effect is present, with multiunit firms and the possibility of upwards belief adjustment there is ambiguity about which effect dominates. We find that in our example the first effect dominates which leads to the higher price charged by the learning monopolist vs. the non-learning monopolist. Finally we run a simulation to compare the learning and non-learning optimal pricing strategies. We find that the pricing strategy of a learning monopolist leads to higher average profits and a lower standard deviation of profits. Thus while the "subjective" value function of the non-learning monopolist is higher, learning improves the realized profits.

This paper is mainly related to the literature on revenue management under uncertain demand. A survey of some of the literature can be found in Araman and Caldentey (2011) who mainly concentrate on contributions from the operations research literature. Overall the contributions of this literature have been focussed on the numerical comparison of approximate heuristic pricing rules vs. optimal pricing and not on the effects of learning on prices. Araman and Caldentey (2009) analyze optimal pricing of a multiunit firm selling nonperishable products that learns about the demand intensity it faces. In principle the formulation of their model is quite similar to ours with the notable difference that they assume a final reward once all units have been sold. This final reward implies that prices increase with inventory size for low beliefs about demand and decrease for high beliefs while in our model prices always decrease. They do not analyze the effect of learning on prices but numerically investigate the performance of different approximate pricing rules vs. the true optimal pricing rule. Aviv and Pazgal (2005) also analyze demand learning by a multiunit firm. They also analyze the performance of different pricing heuristics in a setting with a specific valuation distribution and a prior about the arrival rate of customers that follows a gamma distribution. While they quantify the value of learning i.e. compare the expected revenue of a non-learning firm with that of a learning firm they do not look at how learning affects prices. Farias and Roy (2010) extend the model of Aviv and Pazgal (2005) by considering a prior that follows a finite mixture of gamma distributions. They analyze the performance of different pricing heuristics but do not analyze the effect of learning on prices.

Our paper is organized as follows: In section 3.2 we analyze the pricing problem of a multiunit firm. First we consider the problem without learning (3.2.1) then with learning (3.2.2) in order to compare the optimal pricing rules (3.2.3). In section 3.3 we analyze the problem of a multiunit firm that can replenish its inventory after a sale. First we look at the problem without learning (3.3.1) and then at the problem with learning (3.3.2). In section 3.4 we present a numerical example in order to analyze features of the models which we could not characterize analytically in the general setting. Section 3.5 concludes with an outlook on remaining questions.

3.2 Demand learning without production

We adapt the model from Mason and Välimäki (2011) of a monopolist selling a single good to a market with unknown demand by allowing for multiple units to be sold by the monopolist. Time is denoted by $t = 0, 1, ..., \infty$ and periods are of arbitrarily small length dt. The monopolist faces a stream of short lived consumers, each period a consumer arrives with probability λdt where λ is unknown to the monopolist. We assume that $\lambda \in \{\lambda_L, \lambda_H\}$ with $\lambda_H > \lambda_L$. As we will take the continuous time limit by letting $dt \to 0$, we neglect the possibility of more than one consumer arriving in a period as the probability of this event is of order dt^2 i.e. the probability of more than one consumer arriving becomes arbitrarily small compared to the probability of a single or no consumer arriving. A consumer who arrives in the market in period t randomly draws his valuation v of the good from a commonly known distribution F with density f. We assume that F has an increasing hazard rate i.e. $\frac{f(v)}{1-F(v)}$ is increasing. The consumer exits the market at the end of the period, i.e. the purchasing decision cannot be delayed. Thus the consumer buys the good if his valuation exceeds the current price p_t i.e. if $v \leq p_t$.

In the baseline model we consider an infinitely lived monopolist who has a stock of N identical goods for which he has a valuation of 0. The monopolist does not know the arrival rate λ of customers with certainty but has a belief π_t of the probability with which the arrival rate is λ_H. His only decision is which price p_t to post in the current period.

3.2.1 Pricing without learning about the arrival rate

We first consider the problem of a monopolist who has a constant belief $\pi_t = \pi \forall t$ in order to analyze the effect of learning on optimal price setting by comparing the learning and no learning problems. We denote the expected arrival rate of a monopolist with belief π as $\lambda(\pi) = \pi\lambda_H + (1 - \pi)\lambda_L$. For a particular price p the probability that a buyer with a valuation higher than p arriving in a period of length dt is $\lambda(\pi)(1 - F(p))dt$. Assuming that the monopolist discounts profits at the rate r, the value function of a monopolist with current stock N and belief π can be written as:

$$V(N,\pi)=\max_p\{\lambda(\pi)(1-F(p))dt(p+(1-rdt)V(N-1,\pi))+(1-rdt)(1-\lambda(\pi)(1-F(p))dt)V(N,\pi)\}$$

Neglecting all terms of order dt^2 or higher, the value function can be rewritten in flow value form as:

$$rV(N,\pi)=\max_p\{\lambda(\pi)(1-F(p))[p+V(N-1,\pi)-V(N,\pi)]\} \tag{3.1}$$

Let us first establish two basic intuitive properties of the value function:

Proposition 1 (Properties of the non-learning value function)**.** *The value function of the problem of a monopolist does not update his belief about the arrival rate is*

1. *increasing in the size of the inventory N*
2. *increasing in the belief π*

Using these basic properties we can characterize the optimal price set by the monopolist.

Proposition 2 (Properties of the non-learning optimal price)**.** *The optimal price $p^{NL}(N,\pi)$ of a monopolist who does not update his belief about the arrival rate is*

1. *decreasing in the size of the inventory N*
2. *increasing in the belief π*

Using the results about optimal pricing we can further characterize the value function as being concave. This characterization will become important in order to compare the optimal price set by the ignorant monopolist with the optimal price set by the monopolist who learns about the arrival rate.

Proposition 3 (Concavity of the non-learning value function)**.** *The value function $V^{NL}(N,\pi)$ is concave in π for all N.*

3.2.2 Pricing with learning about the arrival rate

Let us now consider the problem when the monopolist updates his belief about the arrival rate. The monopolist only observes whether a unit has been purchased in the current period or not i.e. consumers who enter the market but decide to not purchase the good are unobserved by the monopolist. Thus beliefs are updated depending on whether a sale has happened in the period or not. Denoting $\Delta\lambda = \lambda_H - \lambda_L$, by Bayes' rule, the posterior belief if a unit has been sold and the initial belief was π_t is given by:

$$\pi_{t+1} = P(\lambda = \lambda_H | S, \pi_t, p_t) = \frac{P(S|\lambda_H, p_t)\pi_t}{P(S|\lambda_H, p_t)\pi_t + P(S|\lambda_L, p_t)(1-\pi_t)} = \frac{\lambda_H \pi_t}{\lambda_L + \Delta\lambda\pi_t}$$

Thus the belief discretely adjusts upwards as $\frac{\lambda_H}{\lambda_L + \Delta\lambda\pi} \geq 1 \forall \pi \in [0,1]$. This adjustment is independent of the price charged in the current period. This is due to the fact that the informativeness of the event "sale" about the event "customer present" is independent of the price. Analogously we can calculate the posterior belief if no sale occurred in the period:

$$\pi_{t+1} = \frac{P(NS|\lambda_H, p_t)\pi_t}{P(NS|\lambda_H, p_t)\pi_t + P(NS|\lambda_L, p_t)(1-\pi_t)} = \frac{[1 - \lambda_H(1 - F(p_t))dt]\pi_t}{1 - \lambda_H(1-F(p_t))dt\pi_t - \lambda_L(1-F(p_t))dt[1-\pi_t]}$$

Thus the belief change can be written as:

$$d\pi_t = \pi_{t+1} - \pi_t = \frac{-\Delta\lambda(1-F(p_t))\pi_t(1-\pi_t)dt}{1-(1-F(p_t))[\lambda_H\pi_t + \lambda_L(1-\pi_t)]dt}$$

Then

$$\lim_{dt \to 0} d\pi_t/dt = -\Delta\lambda(1-F(p_t))\pi_t(1-\pi_t) \leq 0$$

identical to equation (1) in Mason and Välimäki (2011). Thus the belief continuously adjusts downwards as long as no sale occurs and jumps upwards once a sale occurred. This revision is stronger for a lower price as the informativeness of the event "no sale" for the event "no customer present" is higher, the lower the price. Denoting $\lambda(\pi) = \pi\lambda_H + (1-\pi)\lambda_L$ and assuming that the seller discounts future payoffs at rate r we can write the value function of a seller who has a stock of N units and belief π as:

$$\begin{aligned} V(N,\pi) &= \max_p \{\lambda(\pi)(1-F(p))dt[p + (1-rdt)V(N-1,\pi_S)] \\ &\quad + (1-rdt)(1-\lambda(\pi)(1-F(p))dt)V(N, \pi + d\pi_{NS}(p))\} \end{aligned}$$

We can establish the following properties of the value function:

Proposition 4 (Properties of $V(N,\pi)$). *The value of a monopolist with inventory N and belief π is*

1. *nondecreasing in the size of the inventory N*

2. *nondecreasing and convex in the belief π*

Using the fact that $V(N,\pi)$ is convex in π implies that $V(N,\pi)$ is differentiable almost everywhere with respect to π, we can take a Taylor expansion to write $V(N, \pi + d\pi_{NS}) \approx$

$V(N,\pi) + d\pi_{NS} V_\pi(N,\pi)$. Then, omitting all terms of order dt^2 and higher, we can rewrite the value function as:

$$V(N,\pi) = \max_p \left\{ \frac{\lambda(\pi)(1-F(p))[p + V(N-1,\pi_S(\pi))] - \Delta\lambda\pi(1-\pi)(1-F(p))V_\pi(N,\pi)}{r + \lambda(\pi)(1-F(p))} \right\} \tag{3.2}$$

Taking the first order condition with respect to p yields:

$$\lambda(\pi)(1-F(p)) = rf(p)\frac{\lambda(\pi)[p + V(N-1,\pi_S(\pi))] - \Delta\lambda\pi(1-\pi)V_\pi(N,\pi)}{r + \lambda(\pi)(1-F(p))}$$

At the optimal price $p^*(N,\pi)$, the first order condition can be rewritten to:

$$\frac{[1-F(p^*(N,\pi))]^2}{f(p^*(N,\pi))} = \frac{rV(N,\pi)}{\lambda(\pi)} \tag{3.3}$$

This form of the first order condition allows us to establish the following result regarding the optimal price:

Proposition 5 (Properties of $p^*(N,\pi)$)**.** *The optimal price $p^*(N,\pi)$ set by the monopolist with belief π and inventory size N is*

1. *nonincreasing in the number of units in the inventory N*
2. *nondecreasing in the belief π of the probability with which the arrival rate is high*

These properties are quite intuitive, as the price is nonincreasing in supply which here is measured by the inventory of the firm and nondecreasing in the belief about demand being high. This result also establishes that after a sale occurs, the optimal price increases due to an increase in the belief that the arrival rate is high and due to a decrease of the inventory.

3.2.3 Optimal pricing with learning vs. no learning

Following Mason and Välimäki (2011) we can compare the optimal prices of a monopolist who updates his beliefs with the optimal prices of a monopolist who does not update his beliefs. It is useful to rewrite the value functions in a comparable way. Let us first define a common component:

$$V_R(p,\pi) = \frac{\lambda(\pi)(1-F(p))p}{r + \lambda(\pi)(1-F(p))}$$

We can then write the value function without learning as:

$$V^{NL}(N,\pi) = \max_p \left\{ V_R(p,\pi) + \frac{\lambda(\pi)(1-F(p))V^{NL}(N-1,\pi)}{r + \lambda(\pi)(1-F(p))} \right\}$$

Similarly we can write the value function with learning as:

$$V^L(N,\pi) = \max_p \Bigg\{ V_R(p,\pi) + \frac{\lambda(\pi)(1-F(p))V^L(N-1,\pi)}{r+\lambda(\pi)(1-F(p))} + \frac{\overbrace{\lambda(\pi)(1-F(p))(V^L(N-1,\pi_S)-V^L(N-1,\pi))}^{\text{(1) Optimism after sale}} - \overbrace{\Delta\lambda\pi(1-\pi)(1-F(p))V^L_\pi(N,\pi)}^{\text{(2) Pessimism after no sale}}}{r+\lambda(\pi)(1-F(p))} \Bigg\}$$

Comparing the value functions of both problems, we see that learning affects the value function in two ways: (1) by increasing the value through an upward shift in beliefs after a sale and (2) by decreasing the value through a downward shift in beliefs if no sale occurs. While (1) is only relevant for a monopolist with inventory $N \geq 2$, effect (2) affects the value function for all inventory levels.

A monopolist with inventory $N = 1$ only experiences the negative "increasing pessimism" effect (2). This directly implies that for $N = 1$ the value of the objective function of the learning problem is smaller than the value of the objective function in the non-learning problem at all prices p. Thus the maximum of the learning problem must be smaller i.e. $V^L(1,\pi) \leq V^{NL}(1,\pi)\forall\pi$. Using the first order conditions of the problems (equation (C.1) for the non-learning problem and equation (3.3) for the learning problem) this observation delivers Theorem 1 from Mason and Välimäki (2011) that the learning price is greater than the non-learning price for the $N = 1$ monopolist.

A similar analysis is more complicated for monopolists with inventory $N \geq 2$ as effects (1) and (2) act in opposite directions and it is thus not obvious whether $V^L(N,\pi)$ is larger or smaller than $V^{NL}(N,\pi)$. However we can use the properties of both value functions to show that $V^{NL}(N,\pi) > V^L(N,\pi)$. This result then has direct implications for the prices set by the non-learning vs. the learning monopolist.

Proposition 6 (Price with learning vs. no learning)**.** *For the problems of the non-learning vs. the learning monopolist the following holds for all N:*

1. $V^{NL}(N,\pi) \geq V^L(N,\pi)$
2. $p^{NL}(N,\pi) \leq p^L(N,\pi)$

Both inequalities hold strictly for all $\pi \in (0,1)$.

3.3 Demand learning with replacement

In this section we extend the model by allowing the monopolist to invest into replacement i.e. the monopolist can choose a probability with which the inventory does not decrease in the case of a sale. This is a way of modelling reproduction of products that facilitates the analysis of the pricing and production choice problem as we can rely on inductive arguments for characterizing the value function and optimal prices. The general problem of a monopolist who chooses a production rate and thus can increase his inventory can no longer be characterized as easily. The general problem is considered in the appendix.

3.3.1 Replacement without learning

We first consider the problem without demand learning. In order to better compare the results with those of the learning case, we still consider a monopolist with belief π and current inventory N. The belief π translates into a customer arrival rate of $\lambda(\pi)$. Such a monopolist can choose the price he charges p and the replacement probability γ. In the case of a sale the monopolist who chose γ has to pay the convex "replacement cost" $c(\gamma)$ with $c(0) = 0$; the inventory decreases by 1 unit with probability $1 - \gamma$ and remains constant at N units with probability γ. This setup translates into the following value function:

$$rV^{RN}(N,\pi) = \max_{p,\gamma}\{\lambda(\pi)(1-F(p))[p+(1-\gamma)V^{RN}(N-1,\pi)+\gamma V^{RN}(N,\pi)-V^{RN}(N,\pi)-c(\gamma)]\}$$

Arguments similar to the proof of Proposition 1 which characterizes the value function of the non-learning problem without replacement can be used to show that $V^{RN}(N,\pi)$ is increasing in N and π.

The first order conditions of the problem are:

$$\begin{aligned}
\frac{\partial}{\partial\gamma} : 0 &= V^{RN}(N,\pi) - V^{RN}(N-1,\pi) - c'(\gamma) \\
\frac{\partial}{\partial p} : 0 &= 1 - F(p) - f(p)[p + (1-\gamma)(V^{RN}(N-1,\pi) - V^{RN}(N,\pi)) - c(\gamma)]
\end{aligned}$$

Using the first order condition we can show the following properties of the optimal choices:

Proposition 7 (Price and production without learning)**.** *The optimal price $p^{RN}(N,\pi)$ and the optimal rate of production $\gamma^{RN}(N,\pi)$ move in the same direction. They are decreasing in the current inventory and increasing in the belief about the arrival rate of customers.*

We can also compare the price set by a monopolist who has the ability to replace a unit he just sold with the optimal price of a monopolist who does not possess this ability. The opportunity cost of selling a unit for the monopolist who cannot reproduce is given by $V^{NL}(N,\pi) - V^{NL}(N-1,\pi)$ while for the monopolist with reproduction the opportunity cost is:

$$V^{RN}(N,\pi) - V^{RN}(N-1,\pi) - \gamma(V^{RN}(N,\pi) - V^{RN}(N-1,\pi)) + c(\gamma)$$

At the optimal γ^*, the opportunity cost will be lower for the seller with reproduction capacities, thus he should charge a lower price.

Proposition 8 (Optimal price with replacement vs. no replacement)**.** *The monopolist with reproduction possibility charges a lower price than the monopolist without reproduction possibility.*

3.3.2 Replacement with learning

Let us now consider the problem of a monopolist with inventory N who updates his belief π conditional on observing a sale or no sale in a period and at the same time can choose to replace a sold unit with some probability γ by incurring a replacement cost $c(\gamma)$. This problem basically combines the problems of section 3.2.2 and 3.3.1. The flow value of a monopolist with current belief π and inventory N is derived similar to the previous problems and is given by:

$$\begin{aligned} rV^{RL}(N,\pi) =& \max_{p,\gamma}\{\lambda(\pi)(1-F(p))[p+(1-\gamma)V^{RL}(N-1,\pi_S)+\gamma V^{RL}(N,\pi_S)-V^{RL}(N,\pi)-c(\gamma)] \\ & -\Delta\lambda(1-F(p))(1-\pi)\pi V^{RL}_{\pi}(N,\pi)\} \end{aligned}$$

The first order conditions of this problem are:

$$\begin{aligned} \frac{\partial}{\partial\gamma}: 0 =& c'(\gamma) - [V^{RL}(N,\pi_S) - V^{RL}(N-1,\pi_S)] \\ \frac{\partial}{\partial p}: 0 =& \lambda(\pi)[1-F(p)-f(p)[p+V^{RL}(N-1,\pi_S)-V^{RL}(N,\pi) \\ & +\gamma(V^{RL}(N,\pi_S)-V^{RL}(N-1,\pi_S))-c(\gamma)]-\Delta\lambda f(p)\pi(1-\pi)V^{RL}_{\pi}(N,\pi) \end{aligned}$$

First of all note that by the same arguments as in the proof of Proposition 7 we can show that the $p^{RL}(N,\pi) < p^{L}(N,\pi)$. We can combine both first order conditions which yields:

$$\begin{aligned} p - \frac{[1-F(p)]}{f(p)} &= V^{RL}(N,\pi) - V^{RL}(N-1,\pi_S) + [c(\gamma)-\gamma c'(\gamma)] - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^{RL}_{\pi}(N,\pi) \\ &= [c(\gamma)+(1-\gamma)c'(\gamma)] + \underbrace{V^{RL}(N,\pi) - V^{RL}(N,\pi_S) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^{RL}_{\pi}(N,\pi)}_{<0} \end{aligned}$$

Comparing this combined first order condition with (C.5), we see that the right hand side of this FOC is smaller than that of (C.5). Furthermore note that the left hand side is increasing in p and the right hand side is increasing in γ. This observation yields the following proposition:

Proposition 9. *The following implications regarding the price and production rate set by the learning vs. the non-learning monopolist hold:*

1. *If* $p^{RL}(N,\pi) > p^{RN}(N,\pi)$, *then* $\gamma^{RL}(N,\pi) > \gamma^{RN}(N,\pi)$
2. *If* $\gamma^{RL}(N,\pi) < \gamma^{RN}(N,\pi)$ *then* $p^{RL}(N,\pi) < p^{RN}(N,\pi)$

Note that at $\pi \in \{0,1\}$ the learning and non-learning value functions are identical which implies that the production rate and price set at these beliefs are identical. We can use comparative statics to analyze how the price and production rate change close to these points. A full analysis as in the setup without replacement is difficult already for the $N = 1$ case as with replacement positive updating can occur for all N and not only for $N > 1$. Thus the learning effect is not unambiguously negative for $N = 1$ as in the no replacement case. Overall the possibility of replacement offers more opportunities for positive updating than the no replacement case. Furthermore note that learning should be pushing the replacement rate upwards vs. the non-learning case as replacement only occurs conditional on a sale i.e. conditional on good news which makes the remaining units more valuable. This is exactly what we find in the neighborhood of $\pi = 0$ and $\pi = 1$ for $N = 1$.

Proposition 10 (Comparative statics for $N = 1$ in the vicinity of $\pi \in \{0,1\}$)**.** *Consider a monopolist with* $N = 1$*. For* π *close to* 0 *or close to* 1 *the replacement probability and the optimal price of a non-learning monopolist is smaller than the replacement rate and the optimal price of a learning monopolist, i.e. for* ε *small:*

1. $\gamma^{RL}(1,\varepsilon) > \gamma^{RN}(1,\varepsilon)$ *and* $p^{RL}(1,\varepsilon) > p^{RN}(1,\varepsilon)$
2. $\gamma^{RL}(1,1-\varepsilon) > \gamma^{RN}(1,1-\varepsilon)$ *and* $p^{RL}(1,1-\varepsilon) > p^{RN}(1,1-\varepsilon)$

3.4 Numerical analysis

In the previous sections we were able to qualitatively characterize the price set by the monopolist with and without learning as well as with and without reproduction possibilities. However except the monopolist without reproduction possibilities we were not able to make general statements about the relative price set by learning vs. non-learning monopolists. In this section we transform the monopolists problem into a ordinary differential equation of the value function in the belief π. We eliminate the optimal price such that the differential equation only depends on $V(N,\pi), V_\pi(N,\pi)$ and $V(N-1,\pi)$. Then we solve for $V(N,\pi)$ recursively. Using the value function we can recover the optimal pricing and production policies of the monopolists. Additionally we are able to calculate the value of information.

3.4.1 The numerical approach

Solving the non-learning problem is straightforward as conditional on $V(N-1,\pi)$ the problem is a simple maximization problem with 1 variable in the baseline case and 2 variables in the case with replenishment. The learning problem is more difficult to solve as the maximization program which defines the value function depends on the derivative of the value function itself (see eq.(3.2)). At the optimal price $p(N,\pi)$ this functional equation becomes an ordinary differential equation. However to solve for the optimal price we need to know the value function. Fortunately the maximization program is a Legendre-Fenchel-Transform which we can invert, which allows us to transform the functional equation into an ordinary differential equation without having to solve for the optimal price explicitly. For a function $f(x)$ the Legendre-Fenchel transform is defined by

$$f^*(k) = \max_{x\in\mathbb{R}}\{kx - f(x)\}$$

Now consider the maximization program for the learning firm without replacement in flow value form:

$$\frac{rV^L(N,\pi)}{\lambda(\pi)} = \max_p\{(1-F(p)[p + V^L(N-1,\pi_S) - V^L(N,\pi) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^L_\pi(N,\pi)]\}$$

Defining $z := 1 - F(p)$, noting that $p = F^{-1}(1-z)$[1] and furthermore defining:

$$k(N,\pi) = V^L(N-1,\pi_S) - V^L(N,\pi) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^L_\pi(N,\pi)$$

we see that the value function is actually the Legendre-Fenchel transform of

$$G(z) = -zF^{-1}(1-z) = -(1-F(p))p$$

i.e.:

$$\frac{rV^L(N,\pi)}{\lambda(\pi)} = G^*(k(N,\pi)) = \max_z\{k(N,\pi)z - G(z)\}$$

[1]Obviously the optimal p corresponding to z is given by $\sup\{p : 1 - F(p) = z\}$. For simplicity we assume for the rest of the paper that $F^{-1}(1-z)$ is single valued.

If $G^*(k)$ is invertible, reorganizing the above equation yields the nonlinear ODE for $V^L(N,\pi)$ given $V^L(N-1,\pi)$:

$$\frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V_\pi(N,\pi) = V^L(N-1,\pi_S) - V^L(N,\pi) - (G^*)^{-1}\left(\frac{rV^L(N,\pi)}{\lambda(\pi)}\right)$$

subject to the boundary conditions $V^L(N,0) = V^{NL}(N,0)$ and $V^L(N,1) = V^{NL}(N,1)$. Having solved for $V^{NL}(N,\pi)$ in the first step, the natural strategy is to solve for $V^L(N,\pi)$ recursively beginning with $V^L(0,\pi) = 0$. One problem that arises is that the boundary points of this ODE are singular points as at both points $\pi(1-\pi) = 0$, which premultiplies V^L_π. We can overcome this problem by using the fact that V^L is convex and find an upper and a lower bound for the value function by solving the ODE on the interval $[\varepsilon, 1-\varepsilon]$ for ε small, using as initial condition $V^L(N,\varepsilon) = V^L(N,0)$ for the lower bound and $V^L(N,\varepsilon) = V^L(N,0) + \varepsilon(V^L(N,1) - V^L(N,0))$ for the upper bound. In our example we assume that the valuations of consumers are uniformly distributed on $[0,1]$, which implies that $F(p) = p$ and $f(p) = 1$ on $[0,1]$. Using the specific distribution we can easily calculate $(G^*)^{-1}(x) = 2\sqrt{x} - 1$ which yields the ODE

$$V^L_\pi(N,\pi)\frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi) = V^L(N-1,\pi_S) - V^L(N,\pi) - 2\sqrt{\frac{rV^L(N,\pi)}{\lambda(\pi)}} + 1$$

which we solve using MATLAB's solver for stiff ODEs ("ode15s").

For the setup with inventory replenishment we have to slightly adjust our numerical strategy. In order to make use of the same LF transformation as above, we first have to eliminate the replenishment rate from the flow value function equation. The first order condition for γ^{RL} was given by:

$$c'(\gamma) = V^{RL}(N,\pi_S) - V^{RL}(N-1,\pi_S)$$

As we proceed recursively as in the no-replenishment setup, we already know $V^{RL}(N-1,\pi_S)$ at this step. We do not know $V^{RL}(N,\pi)$ yet, as this is the function we ultimately want to recover. In order to recover the value function, we use iteration starting from an initial linear guess of the value function: given the current guess $V^{RL}_n(N,\pi)$ of the value function, we recover the optimal replenishment probability $\gamma^{RL}_n(N,\pi)$ for this guess. We can then use the same LF transformation as above to get the ODE:

$$V^{RL}_\pi(N,\pi)\frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi) = V^{RL}(N-1,\pi_S) - V^{RL}(N,\pi) + \gamma c'(\gamma) - c(\gamma) - 2\sqrt{\frac{rV^{RL}(N,\pi)}{\lambda(\pi)}} + 1$$

We solve this ODE to get the guess $V^{RL}_{n+1}(N,\pi)$ at the next iteration step, from which we calculate the new implied optimal replenishment probability. We stop when $\sup_{\pi\in[0,1]} |\gamma^{RL}_n(N,\pi) -$

$\gamma^{RL}_{n+1}(N,\pi)| < 10^-5$. In our numerical example we make use of the cost function $c(\gamma) = \gamma^2$.

3.4.2 The numerical example

In order to numerically calculate the value function, we have to fix several objects:

1. The consumer valuations are distributed uniformly across $[0, 1]$
2. The arrival rates are $\{\lambda_L, \lambda_H\} = \{3, 6\}$
3. The monopolist's discount rate is $r = 0.2$
4. The cost function for replenishment is $c(\gamma) = \gamma^2$

Given this parametrization we can calculate the value function. An example of the value functions for the setup without replenishment are shown in figure 1. With replenishment the value functions take on higher values but are of similar shape. As we showed

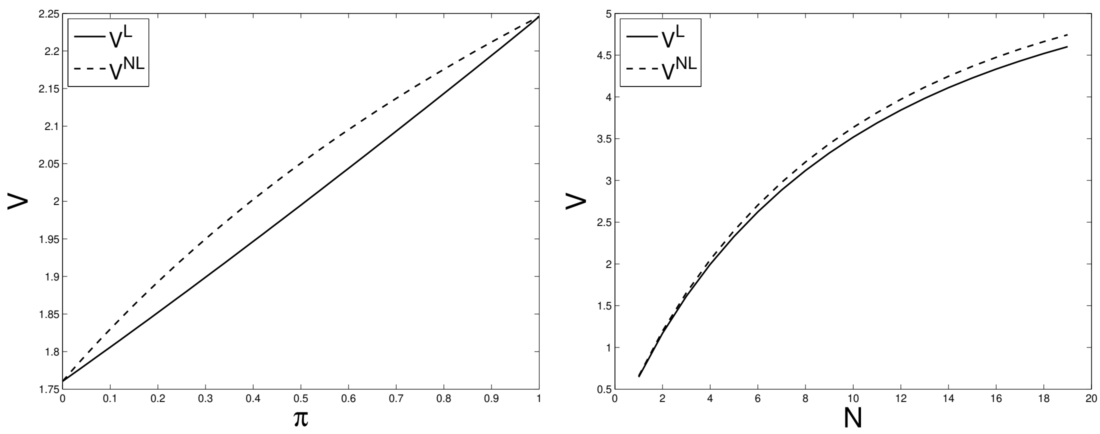


Figure 3.1: Value function for the monopolist without replenishment possibilities. The left panel shows the value functions for $N = 4$, letting π vary. The right panel shows the value functions for $\pi = 0.5$, letting N vary.

earlier, the value function with learning is convex in π. The value function of the non-learning monopolist however is concave. Together with equality of the value functions at the boundary points $\pi = 0$ and $\pi = 1$ this implies that the non-learning monopolist's value function is above that of the learning monopolist. Note however that this does not mean that the realized profit of the non-learning monopolist is higher. In fact, the expected profit of a monopolist who adjusts his belief conditional on observing sales is higher. While the learning monopolist takes into account that his belief is not equal to the true arrival rate, the non-learning monopolist does not. Acknowledging that any intermediate belief is suboptimal leads to the lower values attached to those beliefs. As we have shown earlier, the fact that the value function of the non-learning monopolist takes on higher values than that of the learning monopolist has direct effects on the prices charged by the different sellers.

3.4.3 Comparing learning vs. non-learning policies

After having calculated the value functions, we can recover the optimal pricing and replenishment policies. The structure of the problem implies for all setups the same relation between the optimal price and the value function:

$$\frac{(1 - F(p^*))^2}{f(p^*)} = \frac{rV}{\lambda(\pi)}$$

As we found before that both for the setup with and without replenishment the learning value function was less than the non-learning value function, the above condition implies that the price set by the monopolist who updates his beliefs is higher than the price set by the non-learning monopolist. This is shown for the case with replenishment in figure 2:

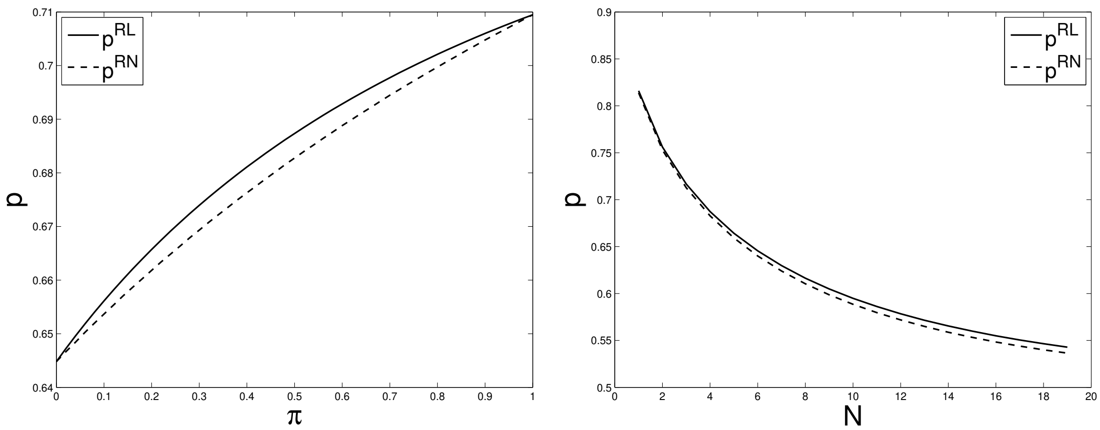


Figure 3.2: Optimal pricing policy for the monopolist with replenishment possibilities. The left panel shows the pricing policy for $N = 4$, letting π vary. The right panel shows the pricing policy for $\pi = 0.5$, letting N vary.

In the setup with replenishment of the inventory we can also look at the implications of learning for the replenishment rate. In our numerical example, the finding from proposition 9 extends to larger N and to the entire range $\pi \in [0, 1]$: The replenishment rate of the learning monopolist is higher than the replenishment rate of the non-learning monopolist.

Inspecting the left panels of figures 2 and 3 we see that the impact of learning on the optimal policies varies with the belief π. Learning affects the policies most in the intermediate range of beliefs where the uncertainty faced by the learning monopolist about the true customer arrival rate is highest.

3.4.4 The value of information

Several contributions on learning through experimentation by a monopolist have established relations between the value of improved information and the direction of exper-

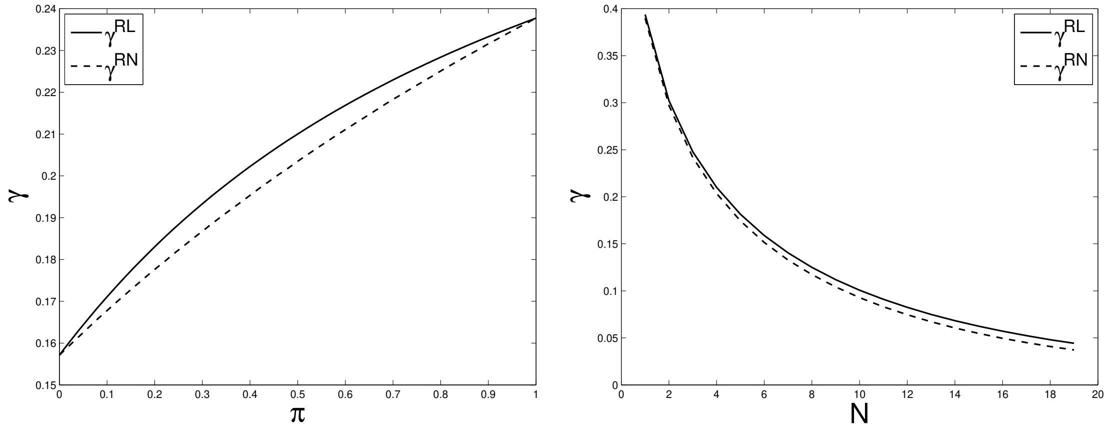


Figure 3.3: Optimal replenishment policy for the monopolist with replenishment possibilities. The left panel shows the replenishment probability for $N = 4$, letting π vary. The right panel shows the replenishment probabilities for $\pi = 0.5$, letting N vary.

imentation. In the models of Grossman et al. (1977) and Trefler (1993) find that an experimenting monopolist moves his price relative to a non-experimenting monopolist into the direction that increases the value of information. If the value of information is negative, the price moves into the direction that reduces the flow of information while when the value of information is positive, the price moves into the direction that increases the flow of information.

The expected value of information is defined as:

$$I(p, \pi) := \mathbb{E}[V(\pi + d\pi)|\pi] - V(\pi)$$

Neglecting terms in dt^2, we can write this expression for the setup without replacement as:

$$I(p, N, \pi) = \lambda(\pi)(1 - F(p))[V^L(N-1, \pi_S) - V^L(N, \pi) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^L_\pi(N, \pi)]dt \quad (3.4)$$

Notice that the value function in flow value form can be rewritten as:

$$rV^L(N, \pi) = \max_p\{\lambda(\pi)(1 - F(p))p + I(p, N, \pi)\}$$

For our numerical example with $F(p) = p$ we can thus write:

$$p^L(N, \pi) = \frac{1 + I_p(p, N, \pi)/\lambda(\pi)}{2}$$

Thus if the value of information is an increasing function of the price, the learning price is higher than the static[2] price ($p = 1/2$), while if it is a decreasing function of the price,

[2]The static price can be seen as the price set by a monopolist with infinite inventory.

the experimenting monopolist will set a lower price than the static price. This result is in line with the results from Trefler (1993). Differentiation of equation (3.4) and noting that $f(p) = 1$ yields:

$$I_p(p, N, \pi) = -\lambda(\pi)[V^L(N-1, \pi_S) - V^L(N, \pi) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^L_\pi(N, \pi)]$$

Exemplary plots of this function are shown in figure 3.4. In our numerical example this derivative is always positive. It is increasing in π and decreasing in N, in line with the findings that price increases in π and decreases in N.

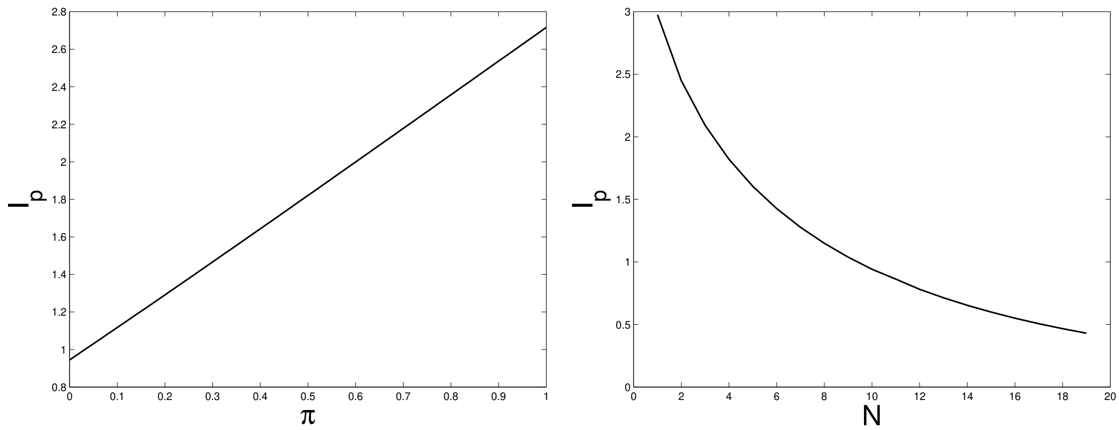


Figure 3.4: Effect of price on the value of information for the monopolist without replenishment possibilities. The left panel shows the effect for $N = 4$, letting π vary. The right panel shows the effect for $\pi = 0.5$, letting N vary.

In addition to the effect of the value of information on the learning price vs. the static price, we can take a look on how increased optimism after a sale and increased pessimism after no sale affect the relation between the learning price and the dynamic non-learning price. Consider the formulation of the value function of a learning monopolist at the beginning of section 3.2.3. In that equation we defined the increased optimism after a sale as:

$$\lambda(\pi)(1 - F(p))(V^L(N-1, \pi_S) - V^L(N-1, \pi))$$

and the increased pessimism after no sale as

$$-\Delta\lambda\pi(1-\pi)(1-F(p))V^L_\pi(N, \pi)$$

The sum of both effects can also be interpreted as the expected value of information, acknowledging that even without learning the value function changes when a unit is sold. Note that the only difference to the original definition (3.4) is that $V^L(N, \pi)$ is replaced by $V^L(N-1, \pi)$. The sum is:

$$\tilde{I}(p, N, \pi) = \mathbb{E}[V^L(N + dN, \pi + d\pi)] - \mathbb{E}[V^L(N + dN, \pi)]$$

$$= \lambda(\pi)(1 - F(p))[V^L(N-1, \pi_S) - V^L(N-1, \pi) - \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V^L_\pi(N, \pi)]$$

In the figures 3.5 and 3.6 below we plot the derivative of this value of information measure together with the difference between the learning and the non-learning price. Both functions match quite well, the difference coming from the fact that learning also affects the optimal price through its effect on $V^L(N-1, \pi)$ as can be seen from inspecting the two value functions at the beginning of section 2.3. As $V^L(N-1, \pi) < V^{NL}(N-1, \pi)$, this effect further increases the difference between p^L and p^{NL}.

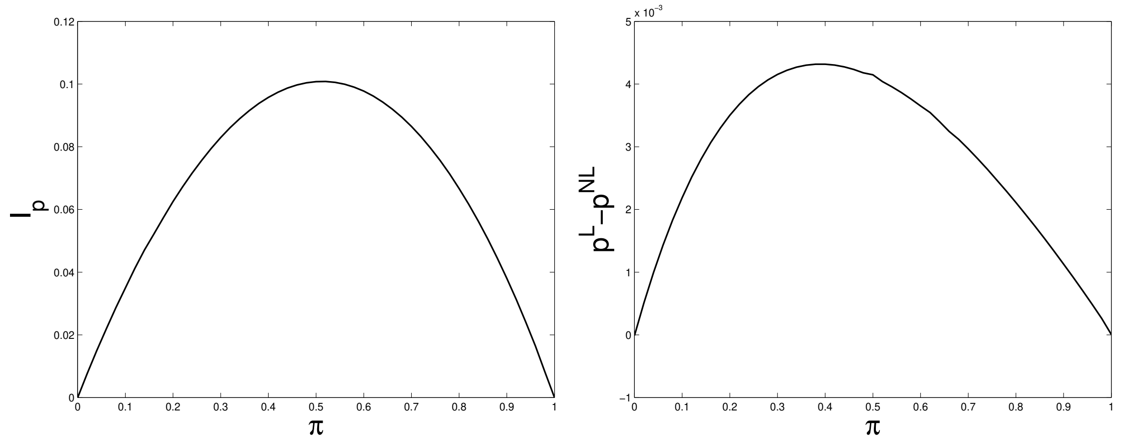


Figure 3.5: Value of information and price difference between the learning and non-learning monopolist without replenishment. The left panel shows the marginal effect of a price change on the value of information. The right panel shows the difference between the price set by the learning and the non-learning monopolist. Both graphs are for $N = 4$, letting π vary.

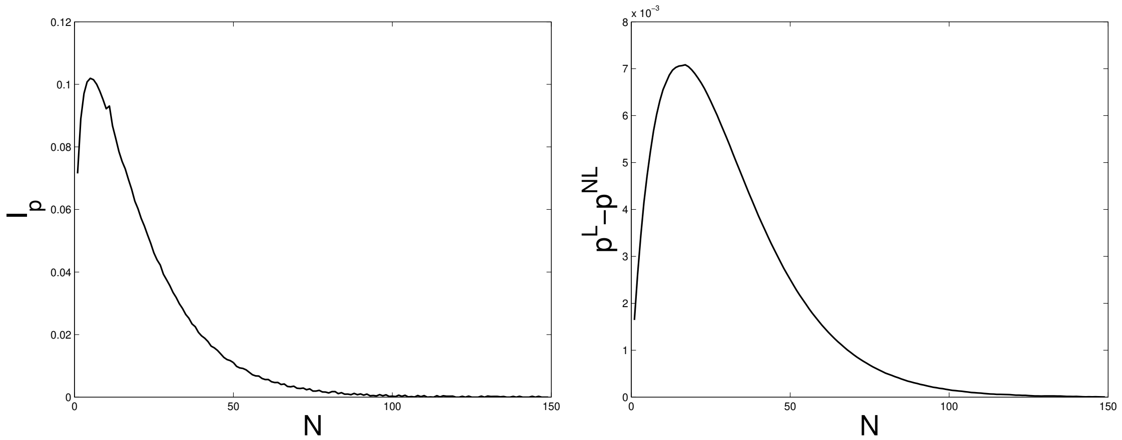


Figure 3.6: Value of information and price difference between the learning and non-learning monopolist without replenishment. The left panel shows the marginal effect of a price change on the value of information. The right panel shows the difference between the price set by the learning and the non-learning monopolist. Both graphs are for $\pi = 0.5$, letting N vary.

Let us finally consider the value of information for the setup with replenishment. As a static version of the model with replenishment is not interesting as there will be no positive probability of replenishment, we only consider the comparison between the learning and the non-learning dynamic models. Thus we resort to the value of information definition from the last paragraphs:

$$\begin{aligned}
I^R(p,\gamma,N,\pi) =&\mathbb{E}[V^{RL}(N+dN,\pi+d\pi)] - \mathbb{E}[V^{RL}(N+dN,\pi)] \\
=&\lambda(\pi)(1-F(p))[(1-\gamma)[V^{RL}(N-1,\pi_S) - V^{RL}(N-1,\pi)] \\
&+ \gamma[V^{RL}(N,\pi_S) - V^{RL}(N,\pi)] \\
&- \frac{\Delta\lambda}{\lambda(\pi)}\pi(1-\pi)V_\pi^{RL}(N,\pi)]
\end{aligned}$$

In Figures 3.7 and 3.8 we plot the derivative of this measure of the value of information with respect to p and γ together with the difference of the optimal policies of a learning monopolist and those of a non-learning monopolist. As in the previous paragraphs, the derivative of the value of information measure and the difference in policies match quite well. As the value of information increases in both the price and the replenishment probability, the learning monopolist sets a higher price than the non-learning monopolist and also chooses a higher replenishment probability than the non-learning monopolist.

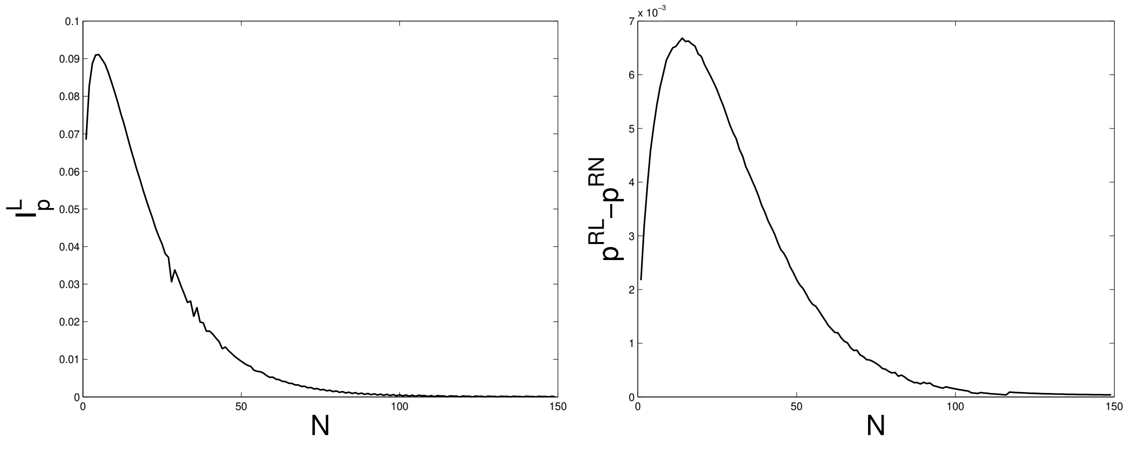


Figure 3.7: Value of information and price difference between the learning and non-learning monopolist without replenishment. The left panel shows the marginal effect of a price change on the value of information. The right panel shows the difference between the price set by the learning and the non-learning monopolist. Both graphs are for $N = 4$, letting π vary.

3.4.5 Simulation

In this section we look at the implications of our model for belief dynamics and prices set by the learning monopolist without replenishment. We use the same parametrization of the model as in the previous sections. Starting with an initial belief of $\pi = 0.5$ we

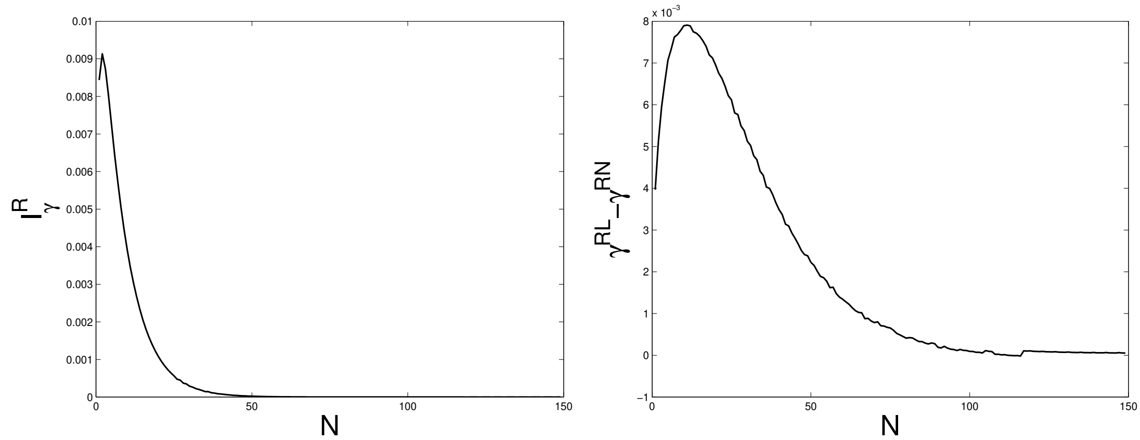


Figure 3.8: Value of information and replenishment probability difference between the learning and non-learning monopolist without replenishment. The left panel shows the marginal effect of a replenishment probability change on the value of information. The right panel shows the difference between the replenishment probability chosen by the learning and the non-learning monopolist. Both graphs are for $\pi = 0.5$, letting N vary.

simulate the model for 5000 periods, each of length $dt = 0.001$. This implies that in the high demand $\lambda_H = 6$ state the expected number of customers arriving over the simulation horizon is 30 while in the low state $\lambda_L = 3$ the expected number is 15. In figure 3.9, the left panel shows the belief dynamics, the middle panel shows the remaining stock and the right panel shows the current price set by the monopolist.

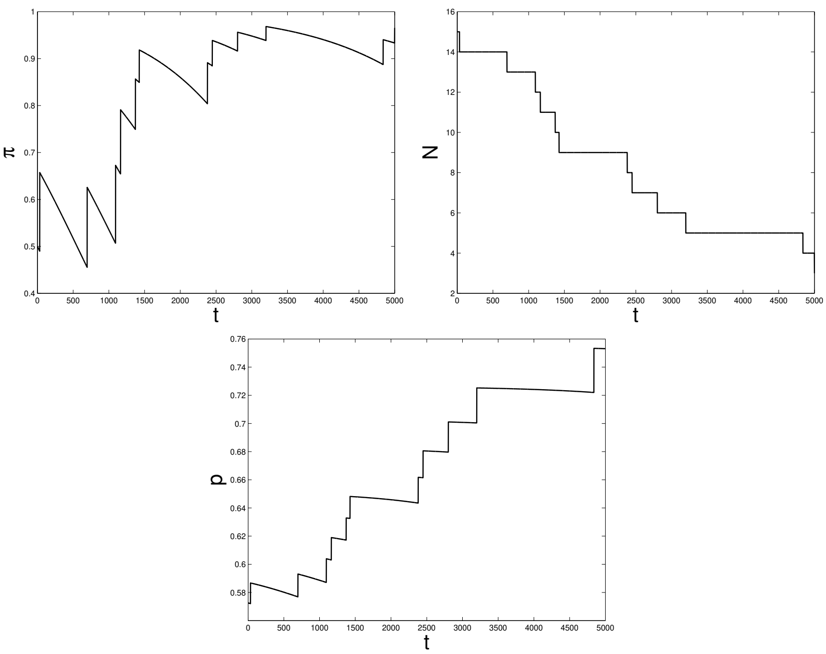


Figure 3.9: Simulation of a high demand market for a learning monopolist without replenishment possibility with initial belief π and initial stock 15. The left panel shows the belief dynamics, the middle panel the remaining stock and the right panel the current price for each period. The period length is $dt = 0.001$ and a total of 5000 periods are simulated.

The simulation in Figure 3.9 was run for a setup with high demand, i.e. the true λ was λ_H. We see that the monopolist's belief approaches 1 over time. Thus there is steady convergence towards the true state. The price posted by the monopolist also increases over time as (1) the belief that demand is high increases and (2) supply/inventory decreases. Finally we can look at how learning affects realized profits of a monopolist. To do so we simulate 500 markets, i.e. customer arrival times and customer valuations. These markets are identical for the learning and the non-learning monopolist. We draw the true customer arrival rate randomly for each market. The probability of $\lambda = \lambda_H$ is 0.5 consistent with an initial belief of $\pi = 0.5$. We then simulate customer arrival for each market. The learning monopolist adjusts his beliefs over time and posts the price $p^L(N_t, \pi_t)$ whereas the non-learning monopolist posts $p^{NL}(N_t, 0.5)$. Comparing realized profits we see that the learning monopolist has (1) higher average realized profits and (2) a lower standard deviation of realized profits.

	Learning	Nonlearning
$mean(\Pi)$	4.0486	4.0383
$\sigma(\Pi)$	1.1949	1.1997

Table 3.1: Mean and standard deviation of realized profits for the learning and non-learning monopolists. Values are calculated from a sample of realized profits of 500 simulated markets.

While these differences are small for the current parametrization, a bigger difference between λ_L and λ_H increases the realized profit advantage of a learning monopolist. While ignorance is bliss with regard to the value function -V^{NL} is larger than V^L - learning pays off when looking at realized profits.

3.5 Conclusion

This paper analyzed demand learning of multiunit monopolists with and without the opportunity of restocking the inventory subsequent to a sale. We first focussed on an analytical characterization of optimal pricing and restocking policies. Subsequently we analyzed the effect of learning on the optimal policies by contrasting the behavior of learning and non-learning firms. The possibility of upward and downward belief shifts result in an ambiguity about the effect of learning on prices. In the single unit setup of Mason and Välimäki (2011) beliefs could only drift downwards and a price increase unequivocally increased the value of learning, which is no longer true in the multiunit setup. Thus our setup introduces potential ambiguity about the direction of experimentation. However, we show that the results from Mason and Välimäki (2011) carry over to the multiunit case in general and that the learning monopolist optimally slows down the speed of learning relative to the non-learning monopolist.

In order to further shed light on the effect of learning on prices we examined a numerical example, relying on methods similar to Araman and Caldentey (2009). Even though the possibility of upwards belief shifts in principle introduces an ambiguity about the effect of learning on prices, we found that the learning monopolist always charges a higher price than the non-learning monopolist in our example. Thus the effect of learning on prices in the multiunit setup is identical to the single unit setup of Mason and Välimäki (2011). Furthermore in the setup with replenishment, the replenishment rate of the learning monopolist is also higher than that of the non-learning monopolist. These differences in the optimal policies are caused by the fact that the value of information is increasing n the price the monopolist charges as well as in the replenishment rate. In line with the results from Grossman et al. (1977) and Trefler (1993) the "experimenting" monopolist experiments towards the direction of an increasing value of information.

Appendix A

Appendix to Chapter 1

A.1 Numerical equilibrium search strategy

The key endogenous variables are the wage, relative prices of the two segments and the masses of entry attempts for each firm type. The solution algorithm iterates over the wage until the labor market clearing wage is found. Within each iteration, I solve for the relative prices p_P/p_S which given the current wage satisfy the correct relative supply of goods, such that $p_P y_S^P = \frac{\beta}{1-\beta} p_S y_S^S$. Then once I have found the correct wage w_i, and the correct relative price, which satisfies the correct relative supply of goods, by Walras Law, not only the labor market but also the goods market clears.

1. Start with a wage guess w_i
2. Guess a price $p_{P,0}$ which implies a price $p_{S,0}$ through (1.4)
3. Solve for optimal firm decisions
4. Given the firm values solve for the number of startup attempts using the free entry condition
5. Given the number of startup attempts calculate the firm type distribution and per capita profits
6. Calculate the relative supply of goods, if $y_S^P > \frac{\beta}{1-\beta} y_S^S$, adjust p_P downwards, otherwise upwards.
7. Iterate until $p_P(w_i)$ is found.
8. Calculate total labor demand for w_i and $p_P(w_i)$
9. If $l^D(w_i) > 1$, adjust w_i upwards, if $l^D(w_i) < 1$ adjust w downwards.
10. Stop once $\frac{w_{i+1}-w_i}{w_i} < 10e^{-5}$

A.2 Calculation of model statistics for calibration

This section describes how to calculate the model statistics which are used for calibrating the model.

Productivity distributions for the firms The distributions are truncated pareto distributions. The non-entrepreneurial firms' distribution starts at 1 and is truncated above by z_S. We thus get for the non-entrepreneurial firms:

$$g_M(z) = g_S(z) = \frac{b_S z^{-b_S-1}}{1 - z_S^{-b_S}} \text{ for } z \in [1, z_S]$$

The distribution for the entrepreneurial firms is truncated below by z_E and we thus get:

$$g_E(z) = \frac{b_E z_E^{b_E}}{z^{b_E+1}} \text{ for } z \geq z_E$$

Total mass of firms To calculate the total mass of firms we calculate the mass of firms of a given type and then add the firm masses of the three firm types. Consider type $i \in \{E, M, S\}$. We first calculate the number of startup attempts a_i. We then get the mass of entrepreneurs successfully finding an idea as $m_E^i = a_j^\gamma \psi_i^{1-\gamma}$. We then calculate the cutoff z_j^* and get the stationary firm mass as:

$$\nu(z, j) = \frac{1 + g_p}{g_p + \delta_j} m_E^j g_j(z) \mathbb{1}(z \geq z_j^*)$$

The per capita mass of firms of type j is then:

$$m_j = \int_{z_j^*}^{\infty} \frac{1 + g_p}{g_p + \delta_j} m_E^j g_j(z)$$

We get then for entrepreneurial firms:

$$m_E = \frac{1 + g_p}{g_p + \delta_E} m_E^E z_E^{b_E} (z_E^*)^{-b_E}$$

And for non-entrepreneurial firms:

$$m_j = \frac{1 + g_p}{g_p + \delta_j} m_E^j \frac{(z_j^*)^{-b_S} - z_S^{-b_S}}{1 - z_S^{-b_S}}$$

In general we can calculate the mass of firms between two productivity levels as follows. For entrepreneurial firms

$$m_E(z_l, z_h) = \frac{1+g_p}{g_p+\delta_E} m_E^E z_E^{b_E} (z_l^{-b_E} - z_h^{-b_E})$$

For non-entrepreneurial firms:

$$m_j(z_l, z_h) = \frac{1+g_p}{g_p+\delta_j} m_E^j \frac{(z_l^*)^{-b_S} - z_h^{-b_S}}{1 - z_S^{-b_S}}$$

These results can be used to calculate the conditional firm shares for firms with more or less than x employees, by using the type specific productivity level z_x^j for which firms of type j have exactly x employees as:

$$FS(e \leq x) = \frac{m_E(z_E^*, z_x^E) + m_M(z_M^*, z_x^M) + m_S(z_S^*, z_x^S)}{m_E + m_M + m_S}$$

and

$$FS(e \geq x) = \frac{m_E(z_x^E, \infty) + m_M(z_x^M, z_S) + m_S(z_x^S, z_S)}{m_E + m_M + m_S}$$

Employment at firms between two cutoffs To calculate the average firm size between two cutoffs z_l and z_h, we calculate the total per capita labor demand of type $j \in \{E, M, S\}$ firms between these cutoffs for each type. Then we add up the calculated labor demands and divide by the mass of firms.

For non-entrepreneurial firms, labor demand between two cutoffs z_l and z_h can be calculated as:

$$\begin{aligned}
l^j(z_l, z_h) &= \int_{z_l}^{z_h} \left(c_e^j + l(z, j)\right) d\nu(z, j) \\
&= \int_{z_l}^{z_h} \left(c_e^j + \left[\frac{z\alpha_j p_j}{w}\right]^{\frac{1}{1-\alpha_j}}\right) d\nu(z, j) \\
&= \left[c_e^j + \left(\frac{\alpha_j p_j}{w}\right)^{\frac{1}{1-\alpha_j}} \frac{b_S}{\frac{1}{1-\alpha_j} - b_S} \frac{(z_h)^{\frac{1}{1-\alpha_j} - b_S} - (z_l)^{\frac{1}{1-\alpha_j} - b_S}}{((z_l)^{-b_S} - (z_h)^{-b_S})}\right] m_j(z_l, z_h)
\end{aligned}$$

Entrepreneurial firms with productivity lower than z_c are constrained by the requirement that $e \geq 1$ and have an establishment size given by

$$z^{\frac{1}{1-\alpha_E\eta_E}} \left(\frac{\alpha_E \eta_E p_P}{w}\right)^{\frac{1}{1-\alpha_E\eta_E}} + c_f^E + c_e^E$$

For these constrained firms, we can calculate total labor demand similarly to single establishment firms as:

$$l_c^E(z_l, z_c) = \left[c_f^E + c_e^E + \left(\frac{\alpha_E \eta_E p_E}{w}\right)^{\frac{1}{1-\alpha_E\eta_E}} \frac{b_E}{\frac{1}{1-\alpha_E\eta_E} - b_E} \frac{(z_c)^{\frac{1}{1-\alpha_E\eta_E}-b_E} - (z_l)^{\frac{1}{1-\alpha_E\eta_E}-b_E}}{((z_l)^{-b_E} - (z_c)^{-b_E})}\right] m_E(z_l, z_c)$$

Unconstrained entrepreneurial firms with productivity between z_c and z_h operate multiple establishments of identical size

$$\frac{c_e^E \alpha_E}{1-\alpha_E}$$

Their labor demand is given by:

$$\begin{aligned}
l_u^E(z_c, z_h) &= \int_{z_c}^{z_h} c_f^E + \left(\frac{c_e^E \alpha_E}{1-\alpha_E} + c_e^E\right) e(z, E) d\nu(z, E) \\
&= \int_{z_c}^{z_h} c_f^E + \left(\frac{c_e^E \alpha_E}{1-\alpha_E} + c_e^E\right) z^{\frac{1}{1-\eta_E}} \left(\frac{p_P}{w}\right)^{\frac{1}{1-\eta_E}} \left(\frac{\eta_E(1-\alpha_E)}{c_e^E}\right)^{\frac{1-\alpha_E\eta_E}{1-\eta_E}} (\alpha_E\eta_E)^{\frac{\alpha_E\eta_E}{1-\eta_E}} d\nu(z, E) \\
&= \left(c_f^E + \frac{b_E}{\frac{1}{1-\eta_E} - b_E}\left[\frac{p_P}{w}\eta_E \alpha_E^{\alpha_E\eta_E}\left(\frac{c_e^E}{1-\alpha_E}\right)^{\eta_E(\alpha_E-1)}\right]^{\frac{1}{1-\eta_E}} \frac{z_h^{\frac{1}{1-\eta}-b_E} - z_c^{\frac{1}{1-\eta_E}-b_E}}{(z_c)^{-b_E} - z_h^{-b_E}}\right) m_E(z_c, z_h)
\end{aligned}$$

Similar to the calculations for the share of firms with more or less than x employees we can also calculate the share of employment at firms with more or less than x employees, by replacing the functions $m_j(z_l, z_h)$ for the mass of firms with the functions $l_j(z_l, z_h)$ for labor demand.

Average size of firms with more or less than x employees The results from above can be used to calculate the average firm size of firms between two cutoffs z_l and z_h as:

$$s(z_l, z_h) = \frac{l_c^E(z_l, \min\{z_c, z_h\}) + l_u^E(\min\{z_c, z_h\}, z_h) + l^M(z_l, z_h) + l^S(z_l, z_h)}{l^E(z_l, z_h)) + l^M(z_l, z_h) + l^S(z_l, z_h)}$$

To calculate the average size of firms with less than x employees we use z_x^j the type specific productivity level for which firms of type j have x employees and then calculate:

$$AS(e \leq x) = \frac{l_c^E(z_E^*, \min\{z_c, z_x^E\}) + l_u^E(\min\{z_c, z_x^E\}, z_x^E) + l^M(z_M^*, z_x^M) + l^S(z_S^*, z_x^S)}{l^E(z_E^*, z_x^E)) + l^M(z_M^*, z_x^M) + l^S(z_S^*, z_x^S)}$$

Similarly to calculate the average firm size of firms with more than x employees we have to use the following relationship:

$$AS(e \geq x) = \frac{l_c^E(z_x^E, \max\{z_c, z_x^E\}) + l_u^E(\max\{z_c, z_x^E\}, \infty) + l^M(z_x^M, z_S) + l^S(z_x^S, z_S)}{l^E(z_x^E, \infty)) + l^M(z_x^M, z_S) + l^S(z_x^S, z_S)}$$

Entry Rate We calculate the entry rate as the mass of succesful entrants relative to the total mass of firms. The mass of successful entrants is given for entrepreneurial firms as:

$$m_E^{SE} = m_E^E z_E^{b_E} (z_E^*)^{-b_E}$$

and for non-entrepreneurial firms as:

$$m_j^{SE} = m_E^j \frac{(z_j^*)^{-b_S} - z_S^{-b_S}}{1 - z_S^{-b_S}}$$

The mass of total firms has been calculated above. The entry rate is then:

$$\frac{m_E^{SE} + m_M^{SE} + m_S^{SE}}{m_E + m_M + m_S}$$

Share of firms below some age We first calculate the mass of firms for each firm type below a certain age. To do this we first calculate the mass of succesful entrants and can then calculate the mass of firms below a certain age as:

$$m_j(a^*) = \sum_{a=0}^{a^*} \left(\frac{1 - \delta_j}{1 + g_p} \right)^a m_j^{SE} = m_j^{SE} \frac{(1 + g_p) \left[1 - \left(\frac{1-\delta_j}{1+g_p} \right)^{a^*+1} \right]}{g_p + \delta_j}$$

The share of firms below a certain age is then given by:

$$FS(a \leq a^*) = \frac{m_E(a^*) + m_M(a^*) + m_S(a^*)}{m_E + m_M + m_S}$$

Exit Rate for firms in particular age and size class In order to calculate the exit rates for firms in a particular age and size class we first calculate the per-capita mass of firms in a particular age and size class. Let z_x^j the productivity level for which firms of type j have exactly x employees. The per-capita mass of entrepreneurial entrants with less than x employees is given by:

$$m_E^{SE}(e \leq x) = m_E^E z_E^{b_E} ((z_E^*)^{-b_E} - (z_x^E)^{-b_E})$$

the per-capita mass of non-entrepreneurial entrants with less than x employees is given by:

$$m_j^{SE}(e \leq x) = m_E^j \frac{(z_j^*)^{-b_S} - (z_x^E)^{-b_S}}{1 - z_S^{-b_S}}$$

The per-capita mass of firms of type j which are of age a and below size x is then given by:

$$m_j^a(e \leq x) = m_j^{SE}(e \leq x) \left(\frac{1 - \delta_j}{1 + g_p} \right)^a$$

The exit rate of firms between ages a_l and a_h with size less than x is then given by the weighted sum of the exit rates:

$$ER(a_l \leq a \leq a_h, e < x) = \frac{\sum_{a=a_l}^{a_h} \sum_{j \in \{E,M,S\}} m_j^a(e \leq x)\delta_j}{\sum_{a=a_l}^{a_h} \sum_{j \in \{E,M,S\}} m_j^a(e \leq x)}$$

Appendix B

Appendix to Chapter 2

B.1 Robustness of Swedish evidence

B.1.1 Time series break in 2004

Definitions of entrepreneurs, workers and firms in the LISA database changed in 2004. This leads to a visible time series break in various characteristics of startup cohorts. While we control for common breaks in the time series across counties using year fixed effects in the main text, in this section we additionally demean the dependent variable at the county level separately for the period before 2004 and after 2004. Qualitatively the main results of our preferred specifications (6) that include year fixed effects remain unchanged, the magnitudes however are smaller. The elasticity of total startup employment with respect to the unemployment rate is -1.87, compared to -2.86 in the main specification.

Dep. var: $\log(Emp)_{t,0,m}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.9405***		1.1957***	0.1973		0.0536
	(0.3419)		(0.3397)	(0.3874)		(0.3915)
County unemp. rate		0.9171***	1.0994***		-1.8967**	-1.8749**
		(0.2755)	(0.2753)		(0.8868)	(0.9024)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.02	0.04	0.07	0.00	0.02	0.02
No. observations	300	300	300	300	300	300

Table B.1: The relationship between total cohort employment at birth and economic conditions at birth

Similarly the elasticity of the number of firms with respect to the unemployment rate is negative at -0.85 which compares to 1.37 in the main specification.

Dep. var: $\log(Firms_{t,0,m})$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.1869		0.5421**	-0.0395		-0.1043
	(0.2582)		(0.2395)	(0.2177)		(0.2206)
County unemp. rate		1.4479***	1.5306***		-0.8028	-0.8453*
		(0.1920)	(0.1941)		(0.4998)	(0.5085)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.00	0.16	0.17	0.00	0.01	0.01
No. observations	300	300	300	300	300	300

Table B.2: The relationship between the number of firms at birth and economic conditions at birth

Dep. var: $\log(\frac{Emp_{t,0,m}}{Firms_{t,0,m}})$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.7563***		0.6606***	0.2358		0.1579
	(0.2069)		(0.2089)	(0.3086)		(0.3130)
County unemp. rate		-0.5132***	-0.4125**		-1.0813	-1.0170
		(0.1688)	(0.1693)		(0.7093)	(0.7215)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.04	0.03	0.06	0.00	0.01	0.01
No. observations	300	300	300	300	300	300

Table B.3: Relationship between firm size at birth and economic conditions at birth

Dep. var: $\log(Emp)_{t+4,4,m}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.9461		2.1206***	-0.0297		-0.2222
	(0.7332)		(0.7395)	(0.9363)		(0.9462)
County unemp. rate		2.2014***	2.7121***		-3.1079	-3.1955
		(0.5298)	(0.5515)		(2.3874)	(2.4211)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.07	0.10	0.00	0.01	0.01
No. observations	240	240	240	240	240	240

Table B.4: Relationship between cohort employment after four years and economic conditions at birth

Dep. var: $\log(\frac{Emp_{t+4,4,m}}{Firms_{t+4,4,m}})$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.7360		1.0343	0.2287		0.1437
	(0.6185)		(0.6521)	(0.9104)		(0.9226)
County unemp. rate		0.4397	0.6888		-1.4671	-1.4104
		(0.4617)	(0.4863)		(2.3279)	(2.3609)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	No	Yes	Yes	Yes
R-squared	0.01	0.00	0.01	0.00	0.00	0.00
No. observations	240	240	240	240	240	240

Table B.5: Relationship between average size after five years and economic conditions at birth

Dep. var: $\frac{Firms_{t+4,4,m}}{Firms_{t,0,m}}$	(1)	(2)	(3)	(4)	(5)	(6)
County empl. growth	0.2074**		0.2236**	-0.0593		-0.0634
	(0.0984)		(0.1041)	(0.1069)		(0.1084)
County unemp. rate		-0.0164	0.0375		-0.0430	-0.0680
		(0.0740)	(0.0776)		(0.2737)	(0.2774)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.02	0.00	0.02	0.00	0.00	0.00
No. observations	240	240	240	240	240	240

Table B.6: Relationship between five year firm survival rates and economic conditions at birth

B.1.2 Finer disaggregation of entrepreneur types

In the main text we performed the analysis for income quintiles of previously employed, two groups of unemployed and one group of persons previously out of the labor force. Here we do the same analysis with a finer disaggregation at the income decile level, four groups of unemployed and two groups of persons previously out of the labor force.

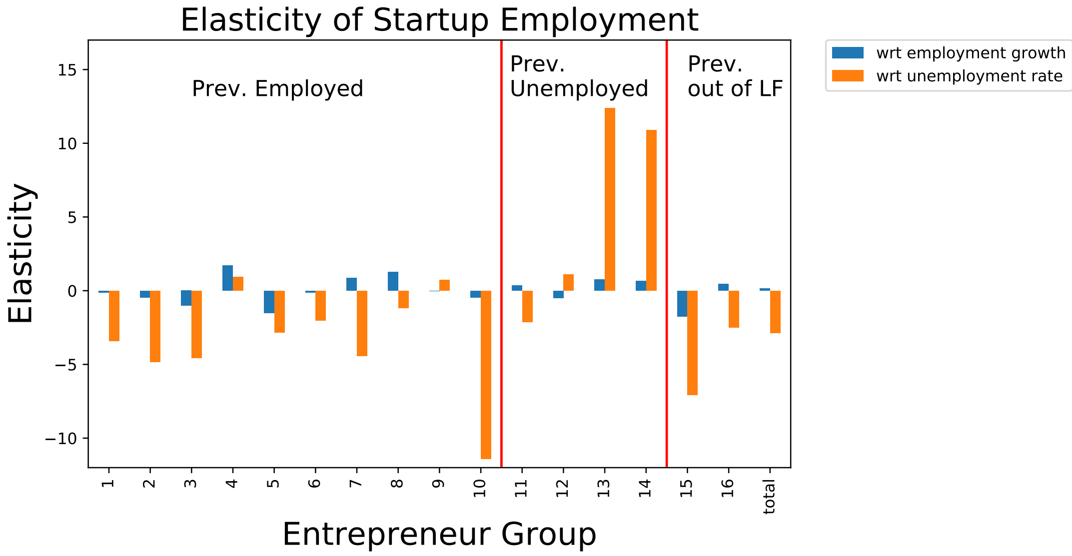


Figure B.1: Partial elasticity of total cohort employment at birth with respect to employment growth and the unemployment rate.

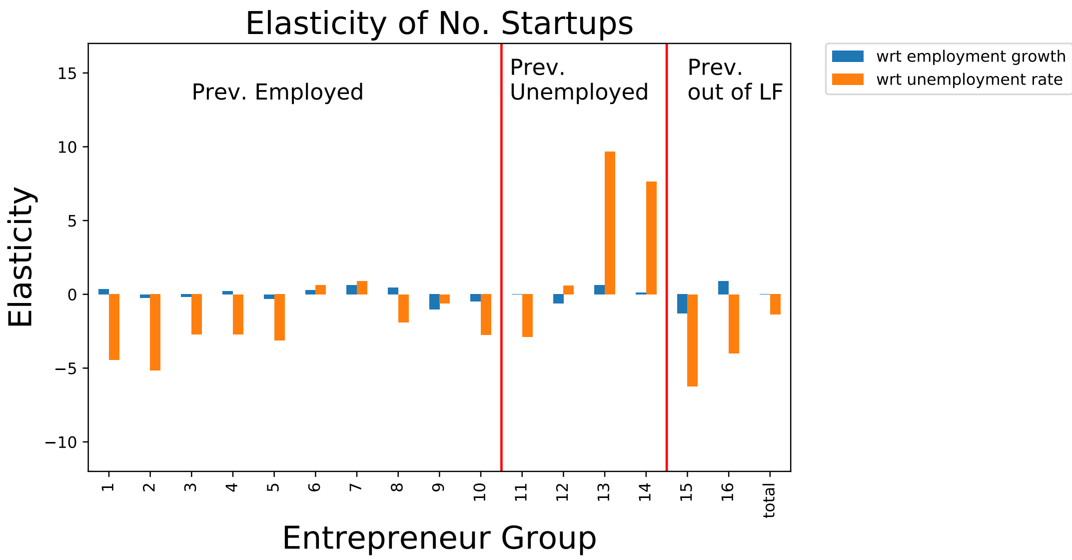


Figure B.2: Partial elasticity of the number of firms at birth with respect to employment growth and the unemployment rate.

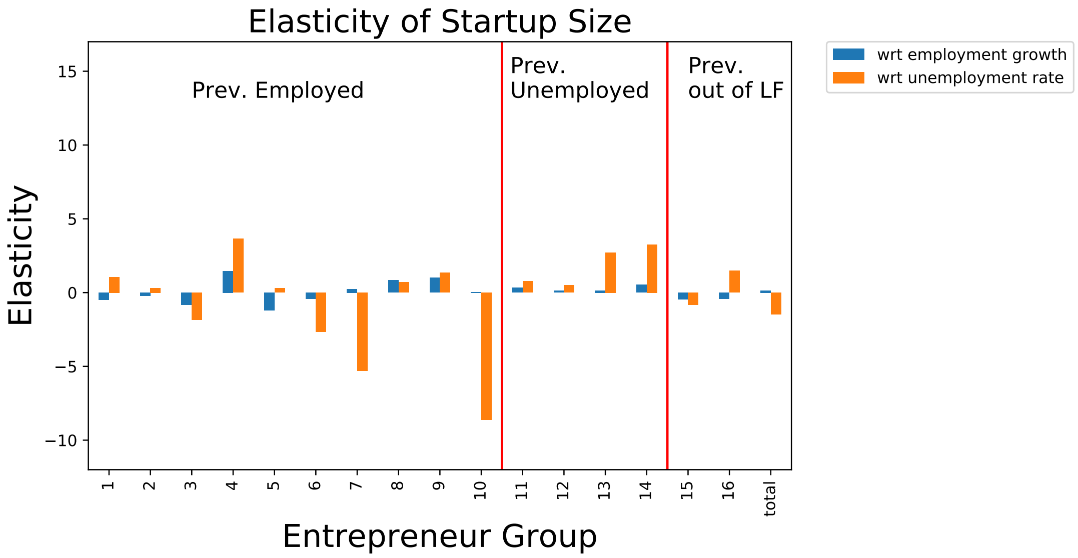


Figure B.3: Partial elasticity of the average firm size at birth with respect to employment growth and the unemployment rate.

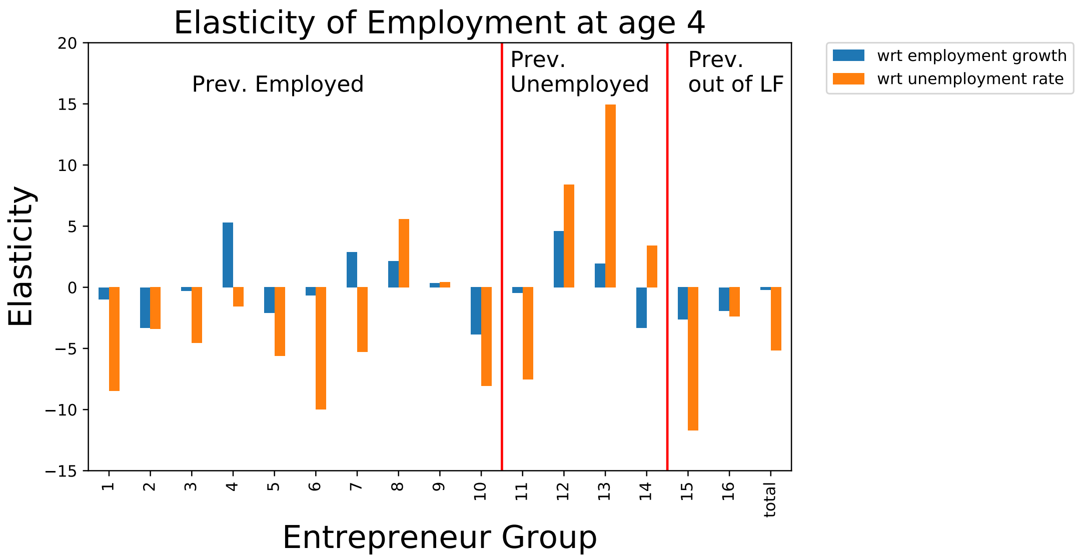


Figure B.4: Partial elasticity of cohort employment at age 4 with respect to employment growth and the unemployment rate.

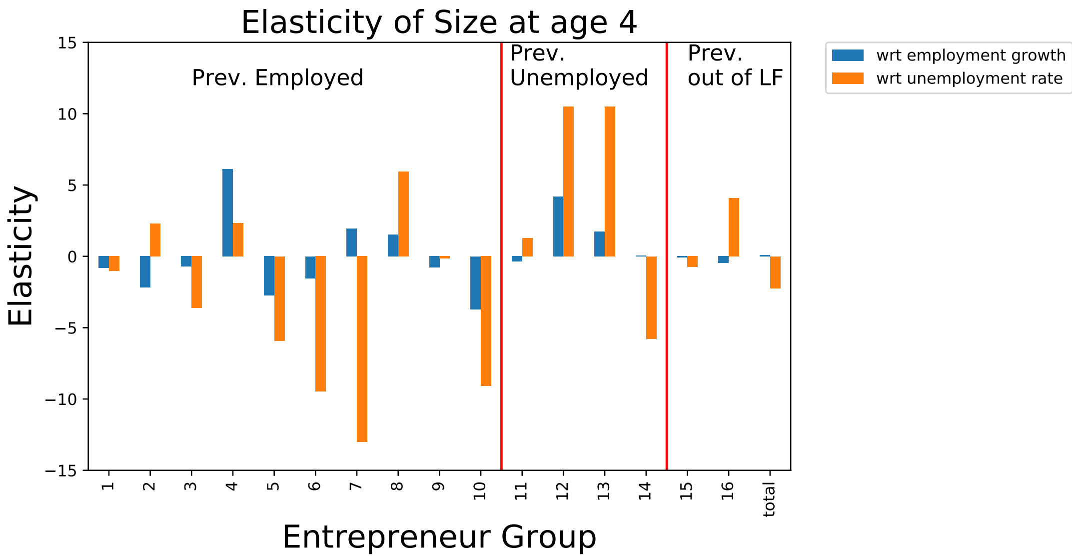


Figure B.5: Partial elasticity of average firm size at age 4 with respect to employment growth and the unemployment rate.

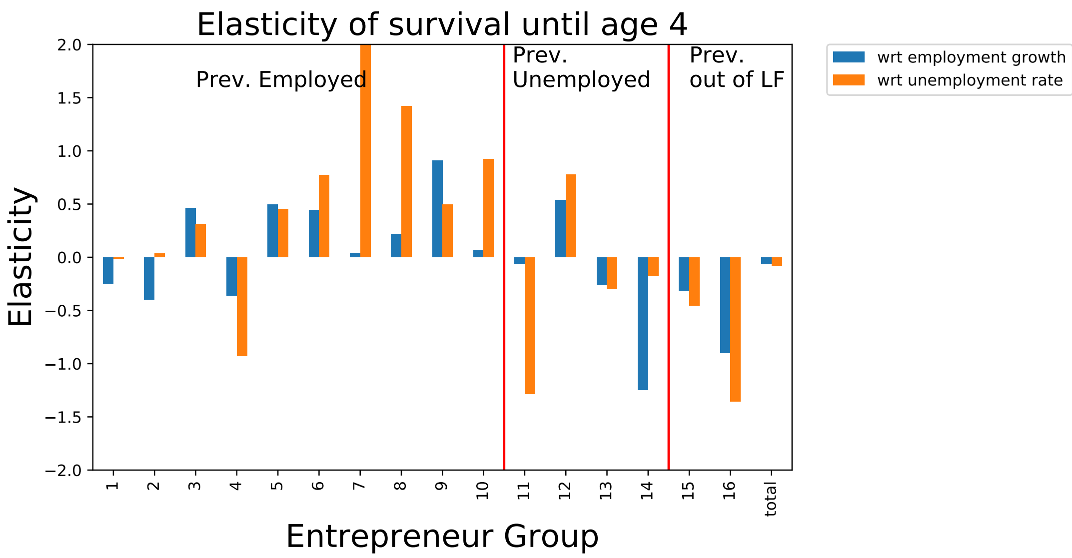


Figure B.6: Partial elasticity of the survival rate until age 4 with respect to employment growth and the unemployment rate.

B.1.3 Controlling for initial conditions in exitrate analysis

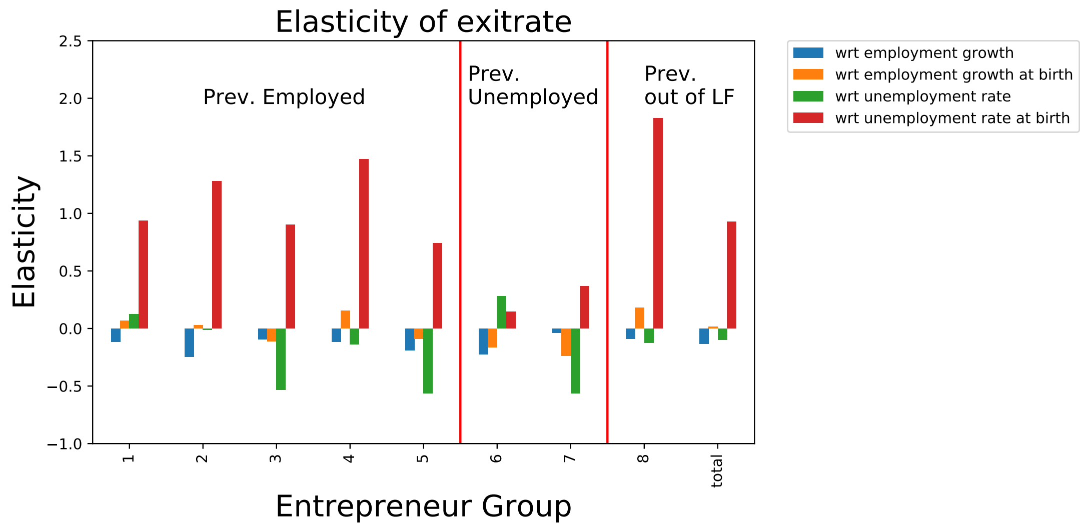


Figure B.7: Elasticity of the exit rate with respect to employment growth and the unemployment rate in the current period and at birth.

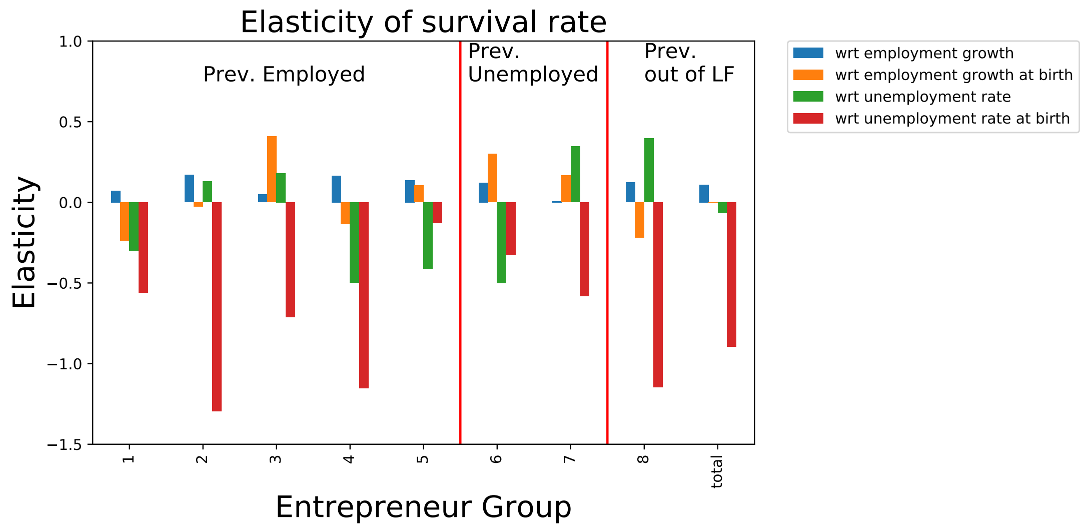


Figure B.8: Elasticity of the survival rate since birth with respect to employment growth and the unemployment rate in the current period and at birth.

B.2 Robustness of US evidence

B.2.1 MSA level detrending

Some of the MSAs in the US data exhibit long run growth, particularly in the number of new firms and employment at new firms. In this section we cleanse out these long run trends by HP-filtering both the dependent variables and the independent variables using a filtering parameter of 6.5. Then we regress deviations from trend of the dependent variables on deviations from trend of the independent variables.

Dep. var: $\log(Emp)_{t,0,m}$	(1)	(2)	(3)
MSA empl. growth		0.0140***	0.0138***
		(0.0005)	(0.0006)
MSA unemp. rate	-0.0102***		-0.0067***
	(0.0016)		(0.0015)
No. observations	7992	7992	7992
R-squared	0.01	0.08	0.08

Table B.7: US: Total cohort employment of startups and economic conditions at birth using detrended variables

Dep. var: $\log(Firms_{t,0,m})$	(1)	(2)	(3)
MSA empl. growth		0.0085***	0.0080***
		(0.0002)	(0.0002)
MSA unemp. rate	-0.0193***		-0.0173***
	(0.0007)		(0.0006)
No. observations	7992	7992	7992
R-squared	0.09	0.14	0.22

Table B.8: US: Number of startups and economic conditions at birth using detrended variables

Dep. var: $\log(\frac{Emp_{t,0,m}}{Firms_{t,0,m}})$	(1)	(2)	(3)
MSA empl. growth		0.0055***	0.0059***
		(0.0005)	(0.0005)
MSA unemp. rate	0.0092***		0.0106***
	(0.0014)		(0.0014)
No. observations	7992	7992	7992
R-squared	0.01	0.01	0.02

Table B.9: US: Startup size at birth and economic conditions at birth using detrended variables

Dep. var: $\log(Emp)_{t+5,5,m}$	(1) I	(2)	(3)
MSA empl. growth		0.0031***	0.0032***
		(0.0009)	(0.0009)
MSA unemp. rate	0.0005		0.0025
	(0.0028)		(0.0028)
No. observations	6323	6323	6323
R-squared	0.00	0.00	0.00

Table B.10: US: Cohort employment after five years and economic conditions at birth using detrended variables

Dep. var: $\log(\frac{Emp_{t+5,5,m}}{Firms_{t+5,5,m}})$	(1)	(2)	(3)
MSA empl. growth		-0.0000	0.0002
		(0.0008)	(0.0008)
MSA unemp. rate	0.0034		0.0035
	(0.0026)		(0.0027)
No. observations	6323	6323	6323
R-squared	0.00	0.00	0.00

Table B.11: US: Average size after five years and economic conditions at birth using detrended variables

Dep. var: $\frac{Firms_{t+5,5,m}}{Firms_{t,0,m}}$	(1)	(2)	(3)
MSA empl. growth		-0.0297***	-0.0518***
		(0.0102)	(0.0103)
MSA unemp. rate	-0.3309***		-0.3622***
	(0.0319)		(0.0324)
No. observations	6323	6323	6323
R-squared	0.02	0.00	0.02

Table B.12: US: Five year firm survival rates and economic conditions at birth using detrended variables

B.3 Variance decomposition of cohort employment for US metros and US states

In this section we decompose the variance of the employment impact of cohorts at birth into the variation in firm size at birth and variation in the number of firms at birth according to equation 2.4. We do so both for US metros and US states and consider two measures. One is the variance of log deviations from the region specific mean, and the second measure is the variance of log deviations from the region specific HP trend. As the number of new firms moves closely with the population size of a region and the different MSAs and states exhibit long run trends in population, detrending substantially reduces the variation in the number of firms. Firm size at birth on the other hand does not exhibit a similar long run comovement with population size. So detrending substantially reduces the share of variance that is attributed to variation in the number of firms both at the MSA and the state level.

	Deviations from MSA mean	Deviations from MSA trend
No. of firms	39.7%	15.0%
Firm size	60.2%	85.0%

Table B.13: Variance decomposition of employment at birth at the MSA level

	Deviations from state mean	Deviations from state trend
No. of firms	53.6%	38.5%
Firm size	46.4%	61.5%

Table B.14: Variance decomposition of employment at birth at the state level

B.4 Derivation of variance decomposition

Let e_t the average firm size at time t, $e_{g,t}$ the average firm size of firms by entrepreneurs from group g and $s_{g,t}$ the share of firms by entrepreneurs from group g. Denote the means of these variables over time by bars i.e. $\bar{e}_t$ the mean average firm size over time. Then the following holds:

$$\begin{aligned} e_t - \bar{e} &= \sum_g e_{g,t}s_{g,t} - \bar{e} \\ &= \sum_g \left[(e_{g,t} - \bar{e}_g)s_{g,t} + (\bar{e}_g - \bar{e})s_{g,t}\right] \end{aligned}$$

$$= \sum_g \left[(e_{g,t} - \bar{e}_g)\bar{s}_g + (\bar{e}_g - \bar{e})s_{g,t} + (e_{g,t} - \bar{e}_g)(s_{g,t} - \bar{s}_g)\right]$$

$$= \sum_g \left[(e_{g,t} - \bar{e}_g)\bar{s}_g + (\bar{e}_g - \bar{e})(s_{g,t} - \bar{s}_g) + (e_{g,t} - \bar{e}_g)(s_{g,t} - \bar{s}_g) + \underbrace{(\bar{e}_g - \bar{e})\bar{s}_g}_{=0}\right]$$

The variance of deviations of the average firm size from its county specific mean is given by:

$$\begin{aligned} Var(e_{t,r} - \bar{e}_r) &= \frac{\sum_r \sum_t \left(e_{t,r} - \bar{e}_r - \frac{\sum_r \sum_t e_{t,r} - \bar{e}_r}{T*R}\right)^2}{T*R} \\ &= \frac{\sum_r \sum_t (e_{t,r} - \bar{e}_r)^2}{T*R} \end{aligned}$$

The second equality is due to the fact that for all regions

$$\sum_c e_{c,r} - \bar{e}_r = 0$$

Now we can use the decomposition from above to rewrite the variance as:

$$\begin{aligned} Var(e_{t,r} - \bar{e}_r) = &\frac{\sum_r \sum_t (e_{t,r} - \bar{e}_r) \sum_g \left[(e_{g,t,r} - \bar{e}_{g,r})\bar{s}_{g,r} + (\bar{e}_{g,r} - \bar{e}_r)(s_{g,t,r} - \bar{s}_{g,r}) + (e_{g,t,r} - \bar{e}_{g,r})(s_{g,t,r} - \bar{s}_{g,r})\right]}{T*R} \\ = &\frac{\sum_g \sum_r \sum_t (e_{t,r} - \bar{e}_r) \left[(e_{g,t,r} - \bar{e}_{g,r})\bar{s}_{g,r} + (\bar{e}_{g,r} - \bar{e}_r)(s_{g,t,r} - \bar{s}_{g,r}) + (e_{g,t,r} - \bar{e}_{g,r})(s_{g,t,r} - \bar{s}_{g,r})\right]}{T*R} \end{aligned}$$

B.5 Business cycle indicators used in other studies

In many studies, following the convention from the 1980s and 1990s real business cycle literature, business cycles are defined as deviations from detrended GDP. Usually national quarterly GDP numbers are detrended using a Hodrick-Prescott filter. However the HP filter has been shown to be unsuitable for short time series and be especially problematic at the limits of the time series. Furthermore studies at the regional level usually don't detrend using the HP filter but use demeaned changes in GDP instead.

Alternative cyclical indicators based on labor market conditions or income have been used in the literature. We use the demeaned level of the unemployment rate and the demeaned change of the unemployment rate as labor market based cyclical indicators. These measures have been used by Moreira (2016) and Fairlie (2013) in recent studies on the cyclicality of entrepreneurship. In addition we also use demeaned changes in log local employment, similar to Sedláček and Sterk (2017).

Group	Measure	Reference
GDP / income	$\Delta_t \log(GDP)$	Lee and Mukoyama (2015), Moreira (2016), Sedláček and Sterk (2017)
	$\Delta_t \log(PInc)$	Moreira (2016)
Labor Market	$\Delta_t u$	Lee and Mukoyama (2015), Moreira (2016)
	$\Delta_t \log(u)$	Moreira (2016)
	u	Fairlie (2013)
	$\Delta_t \log(E)$	Moreira (2016), Sedláček and Sterk (2017)

Table B.15: Cyclical indicators

Appendix C

Appendix to Chapter 3

C.1 Proofs

C.1.1 Proof of Proposition 1

Proof. 1. We can rewrite the value function as

$$V(N,\pi) = \max_p \left\{ \frac{\lambda(\pi)(1-F(p))[p+V(N-1,\pi)]}{r+\lambda(\pi)(1-F(p))} \right\}$$

we will prove the claim by induction. Denote by $p(N,\pi)$ the solution to the monopolists problem. Note that $V(1,\pi) > 0 = V(0,\pi)$. Assume that the claim holds for $N-1$, then:

$$\begin{aligned} V(N,\pi) &\geq \frac{\lambda(\pi)(1-F(p(N-1,\pi)))[p(N-1,\pi)+V(N-1,\pi)]}{r+\lambda(\pi)(1-F(p(N-1,\pi)))} \\ &> \frac{\lambda(\pi)(1-F(p(N-1,\pi)))[p(N-1,\pi)+V(N-2,\pi)]}{r+\lambda(\pi)(1-F(p(N-1,\pi)))} \\ &= V(N-1,\pi) \end{aligned}$$

Where the first inequality holds by suboptimality of $p(N-1,\pi)$ and the second inequality holds due to the induction assumption. Thus by induction $V(N,\pi) > V(N-1,\pi)\forall N$.

2. Now consider the flow value function (3.1) for $N=1$. Let $\pi' < \pi$ and assume that $V(1,\pi) \leq V(1,\pi')$. Note that $\lambda(\pi) > \lambda(\pi')$. Then:

$$\begin{aligned} rV(1,\pi) &\geq \lambda(\pi)[1-F(p(1,\pi'))][p(1,\pi')-V(1,\pi)] \\ &\geq \lambda(\pi)[1-F(p(1,\pi'))][p(1,\pi')-V(1,\pi')] \\ &> \lambda(\pi')[1-F(p(1,\pi'))][p(1,\pi')-V(1,\pi')] = rV(N,\pi') \end{aligned}$$

where the first inequality is due to suboptimality of $p(1, \pi')$ the second inequality is due to assumption and the last inequality is due to $\lambda(\pi) > \lambda(\pi')$. The result is a contradiction to the assumption $V(1, \pi) \leq V(1, \pi')$ and thus it must be that $V(1, \pi) > V(1, \pi')$. Inspecting the flow value function (3.1) we see quite clearly that if $V(N-1, \pi) > V(N-1, \pi')$ the same argument as for $N = 1$ delivers that $V(N, \pi) > V(N, \pi')$. We can thus conclude by induction that for all N, $\pi > \pi'$ implies that $V(N, \pi) > V(N, \pi')$, as claimed.

□

C.1.2 Proof of Proposition 2

Proof. The first order condition of the monopolist's problem is given by:

$$\frac{\partial}{\partial p} : 0 = 1 - F(p) - f(p)[p + V(N-1, \pi) - V(N, \pi)] =: G(p, N, \pi)$$

1. At the optimum we have that:

$$rV(N, \pi) = \lambda(\pi)(1 - F(p^{NL}(N, \pi)))[p^{NL}(N, \pi) + V(N-1, \pi) - V(N, \pi)]$$

   Thus we can rewrite the first order condition as:

$$\frac{[1 - F(p^{NL}(N, \pi))]^2}{f(p^{NL}(N, \pi))} = \frac{rV(N, \pi)}{\lambda(\pi)} \tag{C.1}$$

   Note that by the assumption of an increasing hazard rate, the left hand side of this equation is decreasing in p. The right hand side of this equation is increasing in N as $V(N, \pi)$ is increasing in N. Thus as the first order condition must hold for all N and π it cannot be that $p^{NL}(N, \pi)$ is increasing in N as this would yield a contradiction. Thus $p^{NL}(N, \pi)$ is decreasing in N.

2. In order to prove the second part of the proposition we make use of the implicit function theorem. By the implicit function theorem:

$$\frac{\partial p^{NL}(N, \pi)}{\partial \pi} = - \left. \frac{\partial G / \partial \pi}{\partial G / \partial p} \right|_{p^{NL}(N,\pi), N, \pi}$$

   As p^{NL} is a maximizer, we get by the second order condition:

$$\left. \frac{\partial G(p, N, \pi)}{\partial p} \right|_{p^{NL}(N,\pi), N, \pi} < 0$$

   Thus it remains to show that $\partial G / \partial \pi > 0$. Taking the partial derivative of $G(p, N, \pi)$

with respect to π yields:

$$\frac{\partial G}{\partial \pi} = f(p)[V_\pi(N,\pi) - V_\pi(N-1,\pi)]$$

As $f(p^{NL}(N,\pi)) > 0$ we only have to show that $V_\pi(N,\pi) - V_\pi(N-1,\pi) > 0$. We will show this by induction: for $N = 1$ we have that $V_\pi(1,\pi) > 0$ as $V(1,\pi)$ is increasing in π, furthermore as $V(0,\pi) = 0$, we have that $V_\pi(1,\pi) - V_\pi(0,\pi) > 0$. Using the envelope condition yields the following expression for $rV_\pi(N,\pi)$:

$$\begin{aligned} rV_\pi(N,\pi) =& \Delta\lambda(1 - F(p^{NL}(N,\pi)))[p^{NL}(N,\pi) + V(N-1,\pi) - V(N,\pi)] \\ &+ \lambda(\pi)(1 - F(p^{NL}(N,\pi)))[V_\pi(N-1,\pi) - V_\pi(N,\pi)] \\ =& \Delta\lambda\frac{[1 - F(p^{NL}(N,\pi))]^2}{f(p^{NL}(N,\pi))} \\ &+ \lambda(\pi)(1 - F(p^{NL}(N,\pi)))[V_\pi(N-1,\pi) - V_\pi(N,\pi)] \end{aligned} \tag{C.2}$$

Taking the difference between the above expressions for N and $N-1$ yields:

$$\begin{gathered} [r + \lambda(\pi)(1 - F(p^{NL}(N,\pi)))][V_\pi(N,\pi) - V_\pi(N-1,\pi)] \\ = \\ \Delta\lambda\left[\tfrac{[1-F(p^{NL}(N,\pi))]^2}{f(p^{NL}(N,\pi))} - \tfrac{[1-F(p^{NL}(N-1,\pi))]^2}{f(p^{NL}(N-1,\pi))}\right] \\ +\lambda(\pi)(1 - F(p^{NL}(N-1,\pi)))[V_\pi(N-1,\pi) - V_\pi(N-2,\pi)] \end{gathered}$$

By the fact that $p^{NL}(N,\pi)$ is decreasing in N and the hazard rate is increasing in p, the first term on the right hand side of the equation is positive. The second term is positive by the induction assumption. Thus we can conclude by induction that $V_\pi(N,\pi) - V_\pi(N-1,\pi) > 0$ for all N and π. This implies that $\frac{\partial G}{\partial \pi} > 0$ and thus $\frac{dp^{NL}(N,\pi)}{d\pi} > 0$ which proves the claim.

□

C.1.3 Proof of Proposition 3

Proof. We already know that $V^{NL}(N,\pi)$ is increasing in π. Thus it remains to show that the derivative is decreasing in π i.e. that $V_{\pi\pi}(N,\pi) < 0$. Consider equation C.2. Taking the derivative with respect to π yields:

$$\begin{gathered} (r + \lambda(\pi)(1 - F(p^{NL}(N,\pi))V^{NL}_{\pi\pi}(N,\pi) \\ +(\Delta\lambda(1 - F(p^{NL}(N,\pi))) - \lambda(\pi)f(p^{NL}(N,\pi))\tfrac{\partial p^{NL}(N,\pi)}{\pi})V^{NL}_\pi(N,\pi) \\ = \end{gathered}$$

$$\frac{(1-F(p^{NL}(N,\pi))^2/f(p^{NL}(N,\pi))}{\partial p}\frac{\partial p^{NL}(N,\pi)}{\partial \pi}+\lambda(\pi)(1-F(p^{NL}(N,\pi)V_{\pi\pi}^{NL}(N-1,\pi)$$
$$+V_{\pi}^{NL}(N-1,\pi)(\Delta\lambda(1-F(p^{NL}(N,\pi)))-\lambda(\pi)f(p^{NL}(N,\pi))\frac{\partial p^{NL}(N,\pi)}{\pi})$$

Letting $\psi=(r+\lambda(\pi)(1-F(p^{NL}(N,\pi))))$ which is larger than 0 and rearranging yields:

$$\begin{aligned}V_{\pi\pi}^{NL}(N,\pi)=&\psi^{-1}\Bigg[\frac{(1-F(p^{NL}(N,\pi))^2/f(p^{NL}(N,\pi))}{\partial p}\frac{\partial p^{NL}(N,\pi)}{\partial \pi}\\&+\lambda(\pi)(1-F(p^{NL}(N,\pi)V_{\pi\pi}^{NL}(N-1,\pi)\\&+(V_{\pi}^{NL}(N-1,\pi)-V_{\pi}^{NL}(N,\pi))\\&\times(\Delta\lambda(1-F(p^{NL}(N,\pi)))-\lambda(\pi)f(p^{NL}(N,\pi))\frac{\partial p^{NL}(N,\pi)}{\pi})\Bigg]\end{aligned}\tag{C.3}$$

Now let us consider the terms of the above equation. We will show that $V_{\pi\pi}^{NL}(N-1,\pi)\leq 0$ implies that $V_{\pi\pi}^{NL}(N,\pi)<0$. As $V_{\pi\pi}^{NL}(0,\pi)=0$ the claim then follows by induction. By the fact that the hazard rate of F is increasing and the optimal price is increasing in π, the first term is negative. By the induction assumption, the second term is less or equal to 0. For the third term, note that we have shown in the proof of part b) of proposition 2, that $V_{\pi}^{NL}(N-1,\pi)-V_{\pi}^{NL}(N,\pi)<0$ and thus it only remains to show that

$$\Delta\lambda(1-F(p^{NL}(N,\pi)))-\lambda(\pi)f(p^{NL}(N,\pi))\frac{\partial p^{NL}(N,\pi)}{\pi}>0$$

. Note that this is the derivative of the subjective probability of a sale with respect to the belief π. Using equation C.1 we can write the subjective probability of a sale as:

$$\lambda(\pi)(1-F(p^{NL}(N,\pi)))=rV^{NL}(N,\pi)\frac{f(p^{NL}(N,\pi))}{1-F(p^{NL}(N,\pi))}$$

Taking the derivative of the right hand side with respect to π yields:

$$r\underbrace{V_{\pi}^{NL}(N,\pi)}_{>0}\frac{f(p^{NL}(N,\pi))}{1-F(p^{NL}(N,\pi))}+rV^{NL}(N,\pi)\underbrace{\frac{\partial\frac{f(p^{NL}(N,\pi))}{1-F(p^{NL}(N,\pi))}}{\partial p}}_{>0}\underbrace{\frac{\partial p^{NL}(N,\pi)}{\partial\pi}}_{>0}$$

as the value function is increasing in π by proposition 1, the price is increasing in π by proposition 2 and the hazard rate of F is increasing by assumption. This implies that the last term in equation C.3 is also less than 0 which concludes the proof. □

C.1.4 Proof of Proposition 4

Proof. Let $p_N:\mathbb{N}\times\mathbb{N}^N\to\mathbb{R}_+,(t,t_1,...,t_N)\mapsto p$ a pricing function, giving the price a monopolist with an initial stock of N units charges in period t if the ith unit has been sold in period $t_i < t$, where $t_i = 0$ if the unit has not been sold yet. Let $W(N, \pi, p_N)$ be the expected profit of a monopolist with initial stock of N units and initial belief π who follows the price policy p_N.

1. Consider a seller with N goods with belief π. We want to show that $V(N, \pi) \geq V(N-1, \pi)$. Let p^*_{N-1} the optimal pricing function of the monopolist with $N-1$ units and initial belief π. Assume that the monopolist with N units follows the pricing function

$$\tilde{p}(t, t_1, ..., t_N) = \begin{cases} p^*_{N-1}(t, t_1, ..., t_{N-1}) & \text{if } t_{N-1} = 0 \\ 0 & \text{if } t_{N-1} > 0 \end{cases}$$

Then $W(N, \pi, \tilde{p}) = V(N-1, \pi)$. Note that $\tilde{p}$ is a feasible pricing policy for the monopolist with stock N and belief π and thus the monopolist will do at least as well with the optimal pricing policy. Thus we can conclude that

$$V(N, \pi) \geq W(N, \pi, \tilde{p}) = V(N-1, \pi)$$

2. Consider a seller with N units. Let π, π' two beliefs, wlog $\pi > \pi'$ and let $\alpha \in (0, 1)$. Let $\pi^\alpha = \alpha\pi + (1-\alpha)\pi'$. Denote by $W^H(p_N, N)$ the expected profit of a seller with N units who follows pricing rule p_N if the true demand state is $\lambda = \lambda^H$. Analogously denote by $W^L(p_N, N)$ the expected profit if the true arrival rate is low. Obviously $W^H(p_N, N) \geq W^L(p_N, N) \forall p_N$ as consumers arrive earlier with a higher probability if the arrival rate is high. Let $p_N(\pi)$ the optimal pricing rule of a seller with an initial stock of N units and initial belief π. Then by optimality of $p_N(\pi)$:

$$V(N, \pi) = \pi W^H(p_N(\pi), N) + (1-\pi)W^L(p_N(\pi), N) \geq \pi W^H(p_N(\pi'), N) + (1-\pi)W^L(p_N(\pi'), N)$$

By $W^H(p_N, N) \geq W^L(p_N, N)$ and $\pi > \pi'$:

$$\pi W^H(p_N(\pi'), N) + (1-\pi)W^L(p_N(\pi'), N) > \pi' W^H(p_N(\pi'), N) + (1-\pi')W^L(p_N(\pi'), N) = V(N, \pi')$$

and thus

$$V(N, \pi) > V(N, \pi')$$

In order to show convexity we can use the optimality of $p_N(\pi)$ and $p_N(\pi')$ which yields:

$$\begin{aligned} V(N, \pi) &\geq \pi W^H(p_N(\pi^\alpha), N) + (1-\pi)W^L(p_N(\pi^\alpha), N) \\ V(N, \pi') &\geq \pi' W^H(p_N(\pi^\alpha), N) + (1-\pi')W^L(p_N(\pi^\alpha), N) \end{aligned}$$

Thus:

$$\begin{aligned}\alpha V(N,\pi)+(1-\alpha)V(N,\pi') &\geq (\alpha\pi+(1-\alpha)\pi')W^H(p_N(\pi^\alpha),N)\\ &\quad +(\alpha(1-\pi)+(1-\alpha)(1-\pi'))W^L(p_N(\pi^\alpha),N)\\ &\geq \pi^\alpha W^H(p_N(\pi^\alpha),N)+(1-\pi^\alpha)W^L(p_N(\pi^\alpha),N)\\ &\geq V(N,\pi^\alpha)\end{aligned}$$

which implies that $V(N,\pi)$ is convex in π.

□

C.1.5 Proof of Proposition 5

Proof. In order to prove this proposition we make use of the properties of $V(N,\pi)$ and the assumption of an increasing hazard rate.

1. Assume to the contrary that $p^*(N,\pi)$ was increasing in N. Note that as $V(N,\pi)$ is increasing in N, the right hand side of (3.3) is increasing in N. Due to the increasing hazard rate, $\frac{1-F(p)}{f(p)}$ is decreasing in p, as is $1-F(p)$. Thus the left hand side of (3.3) is decreasing in p. As we assumed that $p^*(N,\pi)$ is increasing in N, the left hand side would be decreasing in N. This contradicts (3.3) holding for all (N,π). Thus $p^*(N,\pi)$ must be nonincreasing in N.

2. Proof for p^* increasing: We can rewrite the value function in flow value form as:

$$rV(N,\pi)=\max_p\{\lambda(\pi)(1-F(p))[p+V(N-1,\pi_S)-V(N,\pi)]-\Delta\lambda\pi(1-\pi)(1-F(p))V_\pi(N,\pi) \tag{C.4}$$

The derivative of objective function with respect to p is:

$$G(p,N,\pi)=\lambda(\pi)(1-F(p)-f(p)[p+V(N-1,\pi_S)-V(N,\pi)])+\Delta\lambda f(p)\pi(1-\pi)V_\pi(N,\pi)$$

At the optimal price the usual first order condition must hold:

$$G(p^*(N,\pi),N,\pi)=0$$

By the implicit function Theorem we can calculate the derivative of $p^*(N,\pi)$ with respect to π as:

$$\frac{dp^*(N,\pi)}{d\pi}=-\left.\frac{\partial G/\partial\pi}{\partial G/\partial p}\right|_{p^*(N,\pi),N,\pi}$$

As $p^*(N,\pi)$ is the maximum, we know that $\partial G/\partial p<0$. Thus it remains to establish that $\partial G/\partial\pi > 0$. We have:

$$\begin{aligned}\frac{\partial G}{\partial \pi} &= \Delta\lambda(1-F(p)-f(p)[p+V(N-1,\pi_S)-V(N,\pi)]) \\ &\quad -\lambda(\pi)f(p)[V_\pi(N-1,\pi_S)\frac{d\pi_S}{d\pi}-V_\pi(N,\pi) \\ &\quad +\Delta\lambda f(p)\frac{d}{d\pi}\pi(1-\pi)V_\pi(N,\pi)\end{aligned}$$

At $\pi=1$, $p=p^*(N,1)=: p_H$ the above becomes:

$$\begin{aligned}\partial G/\partial\pi &= \Delta\lambda\overbrace{(1-F(p_H)-f(p_H)[p_H+V(N-1,1)-V(N,1)])}^{=0} \\ &\quad -\lambda_H f(p_H)V_\pi(N-1,1)\frac{\lambda_L}{\lambda_H}+\lambda_H f(p_H)V_\pi(N,1)-\Delta\lambda f(p_H)V_\pi(N,1) \\ &= \lambda_L f(p_H)(V_\pi(N,1)-V_\pi(N-1,1))\end{aligned}$$

Next we have to show that $V_\pi(N,1)-V_\pi(N-1,1)>0\forall N$. We will prove this claim by induction starting from $N=1$. Note that $V(0,\pi)=0$, thus $V_\pi(0,1)=0$. As $V_\pi(1,1)>0$ due to $V(1,\pi)$ being increasing in π the claim holds true for $N=1$. For $N>1$ we first have to find a tractable expression for $V_\pi(N,1)$. Using the envelope condition on (C.4), we have:

$$\begin{aligned}rV_\pi(N,1) &= \Delta\lambda(1-F(p_H))\overbrace{[p_H+V(N-1,1)-V(N,1)]}^{=(1-F(p_H))/f(p_H)} \\ &\quad -\lambda_H(1-F(p_H))[V_\pi(N-1,1)\frac{\lambda_L}{\lambda_H}-V_\pi(N,1)] \\ &\quad +\Delta\lambda(1-F(p_H))V_\pi(N,1) \\ &= \frac{\Delta\lambda[1-F(p_H)]^2}{f(p_H)}+\lambda_L(1-F(p_H))[V_\pi(N-1,1)-V_\pi(N,1)]\end{aligned}$$

Thus we get:

$$\begin{aligned}&[r+\lambda_L(1-F(p_H(N))](V_\pi(N,\pi)-V_\pi(N-1,\pi)) = \\ &\qquad \Delta\lambda\left[\frac{[1-F(p_H(N))]^2}{f(p_H(N))}-\frac{[1-F(p_H(N-1))]^2}{f(p_H(N-1))}\right] \\ &\qquad +\lambda_L(1-F(p_H(N-1)))[V_\pi(N-1,1)-V_\pi(N-2,1)]\end{aligned}$$

Due to $p_H(N) < p_H(N-1)$ and $(1-F(p))^2/f(p)$ being decreasing in p, the first term on the right hand side is positive. The second term is also positive due to the induction assumption. Thus $V_\pi(N,\pi)-V_\pi(N-1,\pi)>0$, which implies that for all N:

$$\left.\frac{\partial G}{\partial\pi}\right|_{p^*(N,1),N,1} > 0$$

Using the expression for $V_\pi(N, \pi)$ we can rewrite

$$\left.\frac{\partial G}{\partial \pi}\right|_{p^*(N,\pi),N,\pi} = \Delta\lambda(1 - F(p^*(N,\pi))) - \frac{f(p^*(N,\pi))}{1 - F(p^*(N,\pi))} r V_\pi(N,\pi)$$

We know already that $\partial p^*/\partial \pi$ is increasing at $\pi = 1$. Thus for small ϵ, $p^*(N, 1-\epsilon) < p^*(N, 1)$ but this implies that $(1 - F(p^*(N, 1))) < (1 - F(p^*(N, 1-\epsilon))$. Furthermore by $f(p)/(1 - F(p))$ being an increasing function and V_π being increasing in π due to convexity of V in π, we have that

$$\frac{f(p^*(N,1))}{1 - F(p^*(N,1))} r V_\pi(N,1) > \frac{f(p^*(N,1-\epsilon))}{1 - F(p^*(N,1-\epsilon))} r V_\pi(N,1-\epsilon)$$

Thus we have that:

$$\left.\frac{\partial G}{\partial \pi}\right|_{p^*(N,1-\epsilon),N,1-\epsilon} > \left.\frac{\partial G}{\partial \pi}\right|_{p^*(N,1),N,1} > 0$$

and thus

$$\left.\frac{dp^*(N,\pi)}{d\pi}\right|_{\pi=1-\epsilon} > 0$$

This argument can be extended to the whole interval $[0, 1]$, which establishes the claim.

□

C.1.6 Proof of Proposition 6

Proof. First of all note that due to Bayes' rule, the belief of a learning monopolist will never change if $\pi = 1$ or $\pi = 0$, i.e. if the monopolist is certain that the arrival rate is high or low he will never update his belief no matter which sequence of arrivals he observes. Thus there is no difference between the problems of a learning and an ignorant monopolist which directly implies that $V^{NL}(N, 0) = V^L(N, 0)$ and $V^{NL}(N, 1) = V^L(N, 1)$. Now we can use the fact that $V^{NL}(N, \pi)$ is concave and $V^L(N, \pi)$ is convex in π to conclude that

$$V^{NL}(N,\pi) > \pi V^{NL}(N,1) + (1-\pi)V^{NL}(N,0) = \pi V^L(N,1) + (1-\pi)V^L(N,0) > V^L(N,\pi)$$

From the structure of the optimality conditions C.1 and 3.3 in conjunction with the assumption of a decreasing hazard rate we can directly conclude that

$$p^L(N,\pi) > p^{NL}(N,\pi)$$

□

C.1.7 Proof of Proposition 7

Proof. We first show the properties of p^{RN} and then show that γ^{RN} increases iff p^{RN} increases. Similar to the setting without replacement we can rewrite the first order condition for p as:

$$\frac{[1 - F(p^{RN})]^2}{f(p^{RN})} = \frac{rV(N, \pi)}{\lambda(\pi)}$$

As $V^{RN}(N, \pi)$ is increasing in N, and the left hand side is decreasing in p, by the same arguments as in the proof of proposition 2 a), it must be that p^{RN} is decreasing in N. Also similarly to the setting without replacement we can use the implicit function theorem to show that p^{RN} is increasing in in π if and only if the partial derivative of the first order condition with respect to π is positive. Thus we have to show that

$$f(p)(1 - \gamma)[V_\pi^{RN}(N, \pi) - V_\pi^{RN}(N - 1, \pi)] > 0$$

Thus we have to show that $V_\pi^{RN}(N, \pi) - V_\pi^{RN}(N - 1, \pi) > 0$ which we will show by induction. At $N = 1$ by the same arguments as in the proof of proposition 2 we have that $V_\pi^{RN}(1, \pi) - V_\pi^{RN}(0, \pi) = V_\pi^{RN}(1, \pi) > 0$. Using the envelope condition on the value function we can write:

$$rV_\pi^{RN}(N, \pi) = \frac{\Delta\lambda}{\lambda(\pi)} V^{RN}(N, \pi) + \lambda(\pi)(1 - F(p^{RN}(N, \pi)))(1 - \gamma^{RN}(N, \pi))(V_\pi^{RN}(N - 1, \pi) - V_\pi^{RN}(N, \pi))$$

Subtracting $rV_\pi^{RN}(N - 1, \pi)$ from $rV_\pi^{RN}(N, \pi)$ implies:

$$\begin{gathered}[r + \lambda(\pi)(1 - F(p^{RN}(N, \pi)))(1 - \gamma^{RN}(N, \pi))](V_\pi^{RN}(N, \pi) - V_\pi^{RN}(N - 1, \pi)) \\ = \tfrac{\Delta\lambda}{\lambda(\pi)}[V^{RN}(N, \pi) - V^{RN}(N - 1, \pi)] \\ +\lambda(\pi)(1 - F(p^{RN}(N - 1, \pi))(1 - \gamma^{RN}(N - 1, \pi))[V_\pi^{RN}(N - 1, \pi) - V_\pi^{RN}(N - 2, \pi)]\end{gathered}$$

The first term on the right hand side is positive as $V^{RN}(N, \pi)$ is increasing in N and the second term is positive by the induction assumption. Thus we can conclude by induction that $V_\pi^{RN}(N, \pi) - V_\pi^{RN}(N - 1, \pi) > 0$ for all N, π which establishes that p^{RN} is increasing in π.

Plugging in the first order condition for γ into the FOC for p, we can rewrite the first order conditions as:

$$\frac{1 - F(p^{RN})}{f(p^{RN})} - p^{RN} = (1 - \gamma^{RN})(-c'(\gamma^{RN})) - c(\gamma^{RN}) \tag{C.5}$$

We already know that the left hand side of this equation is decreasing in p. The derivative of the right hand side with respect to γ is:

$$(\gamma - 1)c''(\gamma)$$

As $c''(\gamma) > 0$ by convexity and $\gamma^{RN} < 1$, the right hand side is decreasing in γ. Using the implicit function theorem, this establishes that γ^{RN} increases iff p^{RN} increases. □

C.1.8 Proof of Proposition 8

Proof. Note that the value of a monopolist with reproduction possibility is always at least as high as the value of a monopolist without reproduction. This is due to the fact that the the value of a monopolist with reproduction possibilities who chooses $\gamma = 0$ and prices optimally is the same as the value of a monopolist without reproduction possibilities. Thus at the optimal γ the monopolist with reproduction possibilities must achieve weakly better value than at $\gamma = 0$.

Then we can use the fact that the structure of the first order conditions for both problems is identical i.e.:

$$\frac{[1 - F(p^{NL}(N, \pi))]^2}{f(p^{NL}(N, \pi))} = \frac{rV^{NL}(N, \pi)}{\lambda(\pi)}$$

and

$$\frac{[1 - F(p^{RN}(N, \pi))]^2}{f(p^{RN}(N, \pi))} = \frac{rV^{RN}(N, \pi)}{\lambda(\pi)}$$

We already know that the left hand side is decreasing in p and thus as $V^{RN}(N, \pi) \geq V^{NL}(N, \pi)$ the proposition holds. □

C.1.9 Proof of Proposition 10

Proof. First note that $\gamma^{RL}(1, 0) = \gamma^{RN}(1, 0)$. We thus have to show that $\left.\frac{\partial \gamma^{RL}}{\partial \pi}\right|_{(1,0)} > \left.\frac{\partial \gamma^{RN}}{\partial \pi}\right|_{(1,0)}$. Note that by the implicit function theorem and using $\frac{\partial \pi_S}{\partial \pi} = \frac{\lambda_H \lambda_L}{\lambda(\pi)}$, we have:

$$\left.\frac{\partial \gamma^{RL}}{\partial \pi}\right|_{(1,0)} = \frac{V_\pi^{RL}(1, 0)\frac{\lambda_H}{\lambda_L}}{c''(\gamma^{RL}(1, 0))} \text{ and } \left.\frac{\partial \gamma^{RL}}{\partial \pi}\right|_{(1,0)} = \frac{V_\pi^{RN}(1, 0)}{c''(\gamma^{RN}(1, 0))}$$

Thus we have to show that

$$V_\pi^{RL}(1, 0)\frac{\lambda_H}{\lambda_L} > V_\pi^{RN}(1, 0) \tag{C.6}$$

Using the envelope condition and the first order condition for p we can derive a simple expression for $V_\pi^{RN}(1,\pi)$:

$$\begin{aligned}
rV_\pi^{RN}(1,\pi) &= \frac{\Delta\lambda[1-F(p^{RN}(1,\pi))]^2}{f(p^{RN}(1,\pi))} + \lambda(\pi)[1-F(p^{RN}(1,\pi))][\gamma^{RN}(1,\pi)-1]V_\pi^{RN}(1,\pi) \\
\Leftrightarrow V_\pi^{RN}(1,\pi) &= \frac{1}{r+\lambda(\pi)(1-F(p^{RN}(1,\pi)))(1-\gamma^{RN}(1,\pi))}\frac{\Delta\lambda[1-F(p^{RN}(1,\pi))]^2}{f(p^{RN}(1,\pi))}
\end{aligned}$$

The general structure of $V_\pi^{RL}(1,\pi)$ is more complicated. However at $\pi \in \{0,1\}$ we can exploit amongst other things the fact that $\pi_S = \pi$ which simplifies the resulting expression. At $\pi = 0$ we have:

$$\begin{aligned}
rV_\pi^{RL}(1,0) =& \frac{\Delta\lambda[1-F(p^{RL}(1,0))]^2}{f(p^{RL}(1,0))} + \lambda_L[1-F(p^{RL}(1,0))]\left[\gamma^{RL}(1,0)\frac{\lambda_H}{\lambda_L}-1\right]V_\pi^{RL}(1,0) \\
&- \Delta\lambda(1-F(p^{RL}(1,0)))V_\pi^{RL}(1,0) \\
\Leftrightarrow V_\pi^{RL}(1,0) =& \frac{1}{r+\lambda_H(1-F(p^{RL}(1,0)))(1-\gamma^{RL}(1,0))}\frac{\Delta\lambda[1-F(p^{RL}(1,0))]^2}{f(p^{RL}(1,0))}
\end{aligned}$$

Noting that $p^{RL}(1,0) = p^{RN}(1,0)$ we thus get that:

$$\frac{V_\pi^{RN}(1,0)}{V_\pi^{RL}(1,0)} = \frac{r+\lambda_H(1-F(p^{RL}(1,0)))(1-\gamma^{RL}(1,0))}{r+\lambda_L(1-F(p^{RN}(1,0)))(1-\gamma^{RN}(1,0))} > 1 \tag{C.7}$$

Then equation (C.6) holds if

$$\begin{aligned}
&\frac{\lambda_H}{\lambda_L} > \frac{r+\lambda_H(1-F(p^{RL}(1,0)))(1-\gamma^{RL}(1,0))}{r+\lambda_L(1-F(p^{RN}(1,0)))(1-\gamma^{RN}(1,0))} \\
\Leftrightarrow &\frac{\lambda_H}{\lambda_L}r > r
\end{aligned}$$

This condition holds by assumption as $\lambda_H > \lambda_L$. Thus we can conclude that for π small $\gamma^{RL}(1,\pi) > \gamma^{RN}(1,\pi)$ as claimed.

Furthermore equation (C.7) implies that $V^{RL}(1,\pi) < V^{RN}(1,\pi)$ for π small. Noting that the first order condition for p can be rewritten as in the nonreplacement problems:

$$\frac{[1-F(p^{RL}(1,\pi))]^2}{f(p^{RL}(1,\pi))} = \frac{rV^{RL}(1,\pi)}{\lambda(\pi)}$$

and

$$\frac{[1-F(p^{RN}(1,\pi))]^2}{f(p^{RN}(1,\pi))} = \frac{rV^{RN}(1,\pi)}{\lambda(\pi)}$$

The observation that the left hand side is decreasing in p then yields that $p^{RL}(1,\pi) >$

$p^{RN}(1, \pi)$ for π small.

Let us now consider the case for π close to 1. As $V^{RL}(1, 1) = V^{RN}(1, 1)$, we have to show that

$$V_{\pi}^{RL}(1, 1)\frac{\lambda_L}{\lambda_H} < V_{\pi}^{RN}(1, 1) \tag{C.8}$$

We already derived the expression for $V_{\pi}^{RN}(1, 1)$ above as:

$$V_{\pi}^{RN}(1, 1) = \frac{1}{r + \lambda_H(1 - F(p^{RN}(1, 1)))(1 - \gamma^{RN}(1, 1))} \frac{\Delta\lambda[1 - F(p^{RN}(1, 1))]^2}{f(p^{RN}(1, 1))}$$

. Furthermore we have:

$$\begin{aligned} rV_{\pi}^{RL}(1, 1) =& \frac{\Delta\lambda[1 - F(p^{RL}(1, 1))]^2}{f(p^{RL}(1, 1))} + \lambda_H(1 - F(p^{RL}(1, 1)))(\gamma\frac{\lambda_L}{\lambda_H} - 1)V_{\pi}^{RL}(1, 1) \\ &+ \Delta\lambda(1 - F(p^{RL}(1, 1)))V_{\pi}^{RL}(1, 1) \\ \Leftrightarrow V_{\pi}^{RL}(1, 1) =& \frac{1}{r + \lambda_L(1 - F(p^{RL}(1, 1)))(1 - \gamma^{RL}(1, 1))} \frac{\Delta\lambda[1 - F(p^{RL}(1, 1))]^2}{f(p^{RL}(1, 1))} \end{aligned}$$

Thus the condition from equation (C.8) becomes:

$$\begin{aligned} \frac{1}{r + \lambda_L(1 - F(p^{RL}(1, 1)))(1 - \gamma^{RL}(1, 1))} \frac{\lambda_L}{\lambda_H} &< \frac{1}{r + \lambda_H(1 - F(p^{RN}(1, 1)))(1 - \gamma^{RN}(1, 1))} \\ \Leftrightarrow r\frac{\lambda_L}{\lambda_H} &< r \end{aligned}$$

This condition holds by assumption as $\lambda_L < \lambda_H$. This concludes that for ε small, $\gamma^{RL}(1, 1 - \varepsilon) > \gamma^{RN}(1, 1 - \varepsilon)$. Furthermore we have that $V_{\pi}^{RL}(1, 1) > V_{\pi}^{RN}(1, 1)$ which implies that $V^{RL}(1, 1 - \varepsilon) < V^{RN}(1, 1 - \varepsilon)$ and thus by the same arguments as above $p^{RL}(1, 1 - \varepsilon) > p^{RN}(1, 1 - \varepsilon)$. $\square$

Bibliography

Adelino, Manuel, Song Ma, and David Robinson, "Firm age, investment opportunities, and job creation," *The Journal of Finance*, 2017, *72* (3), 999–1038.

Aghion, Philippe, Antonin Bergeaud, PeterJ TimoBoppart, and HuiyuLi Klenow, "A Theory of Falling Growth and Rising Rents," Working Paper 2019.

Albert, Christoph and Andrea Caggese, "Cyclical Fluctuations, Financial Shocks, and the Entry of Fast Growing Entrepreneurial Startups," Working Paper 2019.

Andersson, Fredrik W. and Jan Andersson, "Företagsledarna i Sverige – En algoritm för att peka ut företagens operativa ledare i näringslivet," Technical Report, Statistics Sweden 2012.

Araman, Victor F. and René Caldentey, "Dynamic Pricing for Nonperishable Products with Demand Learning," *Operations Research*, 2009, *57* (5), 1169–1188.

_ and _ , "Revenue management with incomplete demand information," *Wiley Encyclopedia of Operations Research and Management Science*, 2011.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," Working Paper, National Bureau of Economic Research 2017.

Aviv, Yossi and Amit Pazgal, "Dynamic pricing of short life-cycle products through active learning," *Olin School Business, Washington Univ., St. Louis, MO*, 2005.

Bagwell, Kyle, Garey Ramey, and Daniel F Spulber, "Dynamic retail price and investment competition," *The RAND Journal of Economics*, 1997, pp. 207–227.

Baker, George P and Thomas N Hubbard, "Contractibility and asset ownership: On-board computers and governance in US trucking," *The Quarterly Journal of Economics*, 2004, *119* (4), 1443–1479.

Barlevy, Gadi, "The sullying effect of recessions," *The Review of Economic Studies*, 2002, *69* (1), 65–96.

Bartik, Timothy J, *Who benefits from state and local economic development policies?*, WE Upjohn Institute for Employment Research, 1991.

Behrens, Kristian and Vera Sharunova, "Inter-and intra-firm linkages: Evidence from microgeographic location patterns," Working Paper 2015.

Bento, Pedro, "Niche firms, mass markets, and income across countries: Accounting for the impact of entry costs," *Journal of Economic Dynamics and Control*, 2014, *48*, 147–158.

Berger, Allen N, "The economic effects of technological progress: Evidence from the banking industry," *Journal of Money, credit and Banking*, 2003, pp. 141–176.

Bernanke, Ben and Mark Gertler, "Agency costs, net worth, and business fluctuations," *American Economic Review*, 1989, *79* (1), 14–31.

Bernstein, Shai, Emanuele Colonnelli, Davide Malacrino, and Timothy McQuade, "Who Creates New Firms When Local Opportunities Arise?," Technical Report, National Bureau of Economic Research 2018.

Blanchflower, David G and Andrew J Oswald, "What makes an entrepreneur?," *Journal of labor Economics*, 1998, *16* (1), 26–60.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are ideas getting harder to find?," Working Paper, National Bureau of Economic Research 2017.

Brynjolfsson, Erik and Lorin M Hitt, "Beyond computation: Information technology, organizational transformation and business performance," *Journal of Economic Perspectives*, 2000, *14* (4), 23–48.

_ , Thomas W Malone, Vijay Gurbaxani, and Ajit Kambil, "Does information technology lead to smaller firms?," *Management science*, 1994, *40* (12), 1628–1644.

Caballero, Ricardo and Mohamad L Hammour, "The Cleansing Effect of Recessions," *American Economic Review*, 1994, *84* (5), 1350–68.

Carlstrom, Charles T and Timothy S Fuerst, "Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis," *The American Economic Review*, 1997, pp. 893–910.

Catalini, Christian, Jorge Guzman, and Scott Stern, "Passive Versus Active Growth: Evidence from Founder Choices and Venture Capital Investment," Working Paper 26073, National Bureau of Economic Research July 2019.

Chatterjee, Satyajit and Burcu Eyigungor, "Declining Interest Rates and Firm Dynamics: Falling Startups and Rising Concentration," Working Paper 2018.

Choi, Joonkyu, "Entrepreneurial Risk-Taking, Young Firm Dynamics, and Aggregate Implications," *Working Paper*, 2017.

Clementi, Gian Luca and Berardino Palazzo, "Entry, exit, firm dynamics, and aggregate fluctuations," *American Economic Journal: Macroeconomics*, 2016, *8* (3), 1–41.

Congregado, Emilio, Antonio A Golpe, and Simon C Parker, "The dynamics of entrepreneurship: hysteresis, business cycles and government policy," *Empirical Economics*, 2012, *43* (3), 1239–1261.

Council of Economic Advisors, "Benefits of competition and indicators of market power," CEA Issue Brief 2016.

Davis, Steven J and John Haltiwanger, "Labor market fluidity and economic performance," Working Paper, National Bureau of Economic Research 2014.

_ , _ , Ron Jarmin, and Javier Miranda, "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," *NBER macroeconomics annual*, 2006, *21*, 107–179.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "Where has all the skewness gone? The decline in high-growth (young) firms in the US," *European Economic Review*, 2016, *86*, 4–23.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, "The role of entrepreneurship in US job creation and economic dynamism," *The Journal of Economic Perspectives*, 2014, *28* (3), 3–24.

Dillon, Eleanor W and Christopher T Stanton, "Self-employment dynamics and the returns to entrepreneurship," Technical Report, National Bureau of Economic Research 2017.

Engbom, Niklas, "Firm and Worker Dynamics in an Aging Labor Market," Working Paper 2017.

Fairlie, Robert W, "Entrepreneurship, economic conditions, and the great recession," *Journal of Economics & Management Strategy*, 2013, *22* (2), 207–231.

_ and Frank M Fossen, "Defining Opportunity versus Necessity Entrepreneurship: Two Components of Business Creation," Working Paper 26377, National Bureau of Economic Research October 2019.

Farias, Vivek F. and Benjamin Van Roy, "Dynamic Pricing with a Prior on Market Response," *Operations Research*, 2010, *58* (1), 16–29.

Fossen, Frank M, "Entrepreneurship over the Business Cycle in the United States: A Decomposition," Working Paper 2019.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *The American Economic Review*, 2008, *98* (1), 394–425.

_ , _ , and Cornell J Krizan, "Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s," *The Review of Economics and Statistics*, 2006, *88* (4), 748–758.

Friedrich, Benjamin, "Trade shocks, firm hierarchies and wage inequality," Working Paper, Yale University 2015.

Garcia, Sebastian Camarero and Martin Murmann, "Unemployment Benefit Duration and Startup Success," Working Paper 2019.

Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of political economy*, 2000, *108* (5), 874–904.

_ and Esteban Rossi-Hansberg, "Organization and inequality in a knowledge economy," *The Quarterly Journal of Economics*, 2006, *121* (4), 1383–1435.

Giroud, Xavier, "Proximity and investment: Evidence from plant-level data," *The Quarterly Journal of Economics*, 2013, *128* (2), 861–915.

Goldschlag, Nathan and Alex Tabarrok, "Is regulation to blame for the decline in American entrepreneurship?," *Economic Policy*, 2018, *33* (93), 5–44.

Grossman, Sanford J, Richard E Kihlstrom, and Leonard J Mirman, "A Bayesian approach to the production of information and learning by doing," *The Review of Economic Studies*, 1977, pp. 533–547.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, "Are US industries becoming more concentrated?," Working Paper 2017.

Gutiérrez, Germán and Thomas Philippon, "Declining Competition and Investment in the US," Working Paper, National Bureau of Economic Research 2017.

Guzman, Jorge and Scott Stern, *Nowcasting and Placecasting Entrepreneurial Quality and Performance*, University of Chicago Press, August

_ and _ , "The State of American Entrepreneurship: New Estimates of the Quality and Quantity of Entrepreneurship for 32 US States, 1988-2014," Working Paper 22095, National Bureau of Economic Research March 2016.

Haltiwanger, John, Ron S Jarmin, Robert B Kulick, and Javier Miranda, "High growth young firms: Contribution to job, output and productivity growth," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-49*, 2016.

Harrison, J. D., "The decline of American entrepreneurship — in five charts," *The Washington Post*, 2015.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, "The elements of statistical learning: prediction, inference and data mining," *Springer-Verlag, New York*, 2009.

Hathaway, Ian and Robert E Litan, "Declining business dynamism in the United States: A look at states and metros," *Brookings Institution*, 2014.

_ and _ , "The other aging of America: the increasing dominance of older firms," *Brookings Institution*, 2014.

_ and _ , "What's Driving the Decline in the Firm Formation Rate? A Partial Explanation," *The Brookings Institution*, 2014.

_ , Mark E Schweitzer, and Scott Shane, "The Shifting Source of New Business Establishments and New Jobs," *Economic Commentary*, 2014.

Holmes, Thomas J, "Bar codes lead to frequent deliveries and superstores," *RAND Journal of Economics*, 2001, pp. 708–725.

_ and John J Stevens, "An alternative theory of the plant size distribution, with geography and intra-and international trade," *Journal of Political Economy*, 2014, *122* (2), 369–421.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, "Can unemployment insurance spur entrepreneurial activity?," Technical Report, National Bureau of Economic Research 2014.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg, "The Industrial Revolution in Services," Working Paper 25968, National Bureau of Economic Research June 2019.

_ and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

_ and _ , "The life cycle of plants in India and Mexico," *The Quarterly Journal of Economics*, 2014, *129* (3), 1035–1084.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

Hurst, Erik and Benjamin Wild Pugsley, "What Do Small Businesses Do?," *Brookings Papers on Economic Activity*, 2011, *2.*

Hurst, Erik G and Benjamin W Pugsley, "Wealth, Tastes, and Entrepreneurial Choice," in "Measuring Entrepreneurial Businesses: Current Knowledge and Challenges," University of Chicago Press, 2016, pp. 111–151.

Hyatt, Henry R and James R Spletzer, "The recent decline in employment dynamics," *IZA Journal of Labor Economics*, 2013, *2* (1), 5.

Jiang, Helu and Faisal Sohail, "Skill Biased Entrepreneurial Decline," Working Paper 2017.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin, "Demographic Origins of the Startup Deficit," Working Paper 2018.

Kehrig, Matthias, "The cyclical nature of the productivity distribution," *Earlier version: US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15*, 2015.

Kopecky, Joseph, "An aging dynamo: demographic change and the decline of entrepreneurial activity in the United States," Working Paper 2017.

Kozeniauskas, Nicholas, "What's Driving the Decline in Entrepreneurship," Working Paper 2017.

Lee, Yoonsoo and Toshihiko Mukoyama, "Entry and exit of manufacturing plants over the business cycle," *European Economic Review*, 2015, *77*, 20–27.

Levine, Ross and Yona Rubinstein, "Smart and illicit: who becomes an entrepreneur and do they earn more?," *The Quarterly Journal of Economics*, 2017, *132* (2), 963–1018.

Loecker, Jan De and Jan Eeckhout, "The rise of market power and the macroeconomic implications," Working Paper, National Bureau of Economic Research 2017.

Mason, Robin and Juuso Välimäki, "Learning about the Arrival of Sales," *Journal of Economic Theory*, 2011, *146* (4), 1699–1711.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Molloy, Raven, Riccardo Trezzi, Christopher L Smith, and Abigail Wozniak, "Understanding declining fluidity in the US labor market," *Brookings Papers on Economic Activity*, 2016, *2016* (1), 183–259.

Moreira, Sara, "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles," *Working Paper*, 2016.

Moscarini, Giuseppe and Fabien Postel-Vinay, "The contribution of large and small employers to job creation in times of high and low unemployment," *The American Economic Review*, 2012, *102* (6), 2509–2539.

National Academies of Sciences, Engineering and Medicine, *Information Technology and the US Workforce: Where Are We and Where Do We Go from Here?*, National Academies Press, 2017.

Neira, Julian and Rish Singhania, "The Role of Corporate Taxes in the Decline of the Startup Rate," Working Paper 2017.

Ouyang, Min, "The scarring effect of recessions," *Journal of Monetary Economics*, 2009, *56* (2), 184–199.

Poschke, Markus, "Who becomes an entrepreneur? Labor market prospects and occupational choice," *Journal of Economic Dynamics and Control*, 2013, *37* (3), 693–710.

Pugsley, Benjamin W, Petr Sedlacek, and Vincent Sterk, "The nature of firm growth," Discussion Paper 12670, CEPR 2018.

Rampini, Adriano A, "Entrepreneurial activity, risk, and the business cycle," *Journal of Monetary Economics*, 2004, *51* (3), 555–573.

Reedy, EJ and Robert J Strom, "Starting smaller; staying smaller: America's slow leak in job creation," in "Small Businesses in the Aftermath of the Crisis," Springer, 2012, pp. 71–85.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration," Working Paper, National Bureau of Economic Research 2018.

Saint-Paul, Gilles, "Employment protection, international specialization, and innovation," *European Economic Review*, 2002, *46* (2), 375–395.

Salgado, Sergio, "Technical Change and Entrepreneurship," Working Paper 2018.

Saunders, Adam, "Has Information Technology Leveled the Competitive Playing Field?," *Essays on Information Technology and Intangible Capital*, 2011, p. 59.

Schoar, Antoinette, "The divide between subsistence and transformational entrepreneurship," *Innovation policy and the economy*, 2010, *10* (1), 57–81.

Schumpeter, Joseph Alois, *Business cycles*, McGraw-Hill New York, 1939.

Sedláček, Petr and Vincent Sterk, "The growth potential of startups over the business cycle," *American Economic Review*, 2017, *107* (10), 3182–3210.

Smith, Noah, "Why America Stopped Being a Startup Nation," *Bloomberg View*, 2016.

Svaleryd, Helena, "Self-employment and the local business cycle," *Small Business Economics*, 2015, *44* (1), 55–70.

Syverson, Chad, "What determines productivity?," *Journal of Economic literature*, 2011, *49* (2), 326–65.

Tåg, Joacim, Thomas Åstebro, and Peter Thompson, "Hierarchies and entrepreneurship," *European Economic Review*, 2016, *89*, 129–147.

Tambe, Prasanna and Lorin M Hitt, "The productivity of information technology investments: New evidence from IT labor data," *Information Systems Research*, 2012, *23* (3-part-1), 599–617.

The Economist, "American dynamism dimmed," *The Economist*, 2014.

Traina, James, "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," Working Paper 2018.

Trefler, Daniel, "The ignorant monopolist: Optimal learning with endogenous information," *International Economic Review*, 1993, pp. 565–581.

Vardishvili, Ia, "Entry Decision, Option Value of Delay and Business Cycles," *Working Paper*, 2018.

Vereshchagina, Galina and Hugo A Hopenhayn, "Risk taking by entrepreneurs," *American Economic Review*, 2009, *99* (5), 1808–30.

Williamson, Oliver E, "Hierarchical control and optimum firm size," *Journal of political economy*, 1967, *75* (2), 123–138.

Yellen, Janet, "Perspectives on Inequality and Opportunity from the Survey of Consumer Finances," Speech 821, Board of Governors of the Federal Reserve System 2014.

Curriculum Vitae

since 09/2013	Ph.D. Studies in Economics Center for Doctoral Studies in Economics (CDSE) Graduate School of Economics and Social Sciences (GESS) University of Mannheim *Thesis: Essays on Entrepreneurship and Pricing*
09/2014 - 05/2015	Visiting PhD Student, Yale University, New Haven.
08/2012 - 07/2013	M.Sc. in Economics, Barcelona Graduate School of Economics, Barcelona.
04/2009 - 03/2012	B.Sc. in Economics, Ludwig-Maximilians-University, Munich.
09/2010 - 04/2011	Visiting Student, University of British Columbia, Vancouver.
10/2008 - 03/2009	Studies in Mathematics, Technical University Munich, Munich.
11/2007 - 06/2008	Civilian Service, Horst-Schmidt-Kliniken, Wiesbaden.
10/2007 - 11/2007	Military Service, 5./ArtBtl 295, Sigmaringen.
08/2004 - 06/2007	Abitur, Internatsschule Schloss Hansenberg, Geisenheim.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Außerdem erkläre ich mich damit einverstanden, dass die Universität meine Dissertation zum Zwecke des Plagiatsabgleichs in elektronischer Form speichert, an Dritte versendet, und Dritte die Dissertation zu diesem Zwecke verarbeiten.

Mannheim, den 6.12.2019

Niklas Karl Garnadt