Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles

Sara Moreira*

October, 2016

Abstract

This paper examines how the state of the economy when businesses begin operations affects their size and performance over the lifecycle. Using micro-level data that covers the entire universe of businesses operating in the U.S. since the late 1970s, I provide new evidence that businesses born in downturns start on a smaller scale and remain smaller over their entire lifecycle. In fact, I find no evidence that these differences attenuate even long after entry. Using new data on the productivity and composition of startup businesses, I show that this persistence is related to selection at entry and demand-side channels. In order to evaluate the relative importance of these two mechanisms, I build a model of firm dynamics that includes aggregate shocks, idiosyncratic productivity, and a demand accumulation process. When I mute the effects of selection mechanisms, I find that the average initial size differences are more procyclical, but they are less persistent over time. Finally, I use the model to quantitatively evaluate the role of the persistent effects of entry conditions in the propagation of the Great Recession. My current model simulations indicate that the impact of the crisis on the 2008–2009 cohorts reduces aggregate employment by at least one percentage point in the following ten years.

*University of Chicago, spmoreira@uchicago.edu. The research in this paper was conducted while I was Special Sworn Status researchers of the U.S. Census Bureau at the Chicago Census Research Data Center. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information has been disclosed. Support for this research at the Chicago RDC from NSF (ITR-0427889) is also gratefully acknowledged. All errors are mine. I thank the members of my dissertation committee for their support and guidance: Erik Hurst, Steven Davis, Ali Hortaçsu, and Chad Syverson. I also thank the participants in seminars at the University of Chicago for their comments, and Javier Miranda and Frank Limehouse for their help with the datasets. I thank the Kauffman Foundation, and the FCT for financial support.

Electronic copy available at: https://ssrn.com/abstract=3037178

1 Introduction

The 2007–2009 recession and anemic recovery have reinvigorated the study of persistence in business cycle fluctuations. This fundamental issue motivates an extensive theoretical literature, whose central goal is to understand propagation mechanisms. Recent empirical contributions by Fort, Haltiwanger, Jarmin, and Miranda (2013), and Decker, Haltiwanger, Jarmin, and Miranda (2014) suggest that the entry and exit dynamics of startup businesses is critically important for understanding business cycle fluctuations. Firm births account for a significant portion of total job creation and productivity growth in the U.S. economy. Despite the importance of entry and exit dynamics in business cycle fluctuations, little is known about the role of post-entry dynamics in the propagation of economic shocks over time.

In this paper, I provide new evidence that businesses born during recessionary periods start on a smaller scale and remain smaller over their entire lifecycle. In fact, I find no evidence that the average size of cohorts reverts toward the mean over time. This is a surprising result: in the absence of other frictions the wedge between the size of businesses born in good and bad times should gradually disappear as the initial shocks subside. The fact that it does not suggests that the growth dynamics of businesses over their lifecycle may play a complex and previously unexplored role in the persistence of economic fluctuations. To identify and understand the mechanisms that explain this empirical pattern, I take advantage of detailed micro-level data available from the Census Bureau's Longitudinal Business Database (LBD) on the location, industry, and organizational structure of every non-farm establishment operating in the United States between 1976 and 2011. Finally, I implement a theoretical framework to quantify the main economic mechanisms generating the empirical patterns found in the data.

I begin by assessing whether the average size of businesses that begin operations during expansions ("expansionary startups") differs from that of businesses that begin operations in recessions ("recessionary startups"). I estimate these effects using a model that decomposes longitudinal outcomes into age-period-cohort effects. This methodology is well suited to measure the effect of initial economic conditions faced by each generation of businesses while accounting for businesses's lifecycle patterns and the effect of current business cycle conditions on all businesses. The results are striking: a one percent deviation of the real gross domestic product (GDP) from its trend is associated with a one percent increase in the average size of the cohort born in the same year. The procyclicality of average cohort size is robust to using alternative measures of economic activity that capture fluctuations in aggregate uncertainty, labor market conditions, and access to credit markets. Overall, this evidence is consistent with the hypothesis that differences in initial economic conditions affect the initial investment decisions of startup businesses.

After documenting that the average business size across cohorts is significantly affected by aggregate economic conditions at inception, I evaluate the evolution of this effect over the businesses' lifecycle. I find that the size gap is highly persistent as cohorts age, and does not attenuate over time. A one

Electronic copy available at: https://ssrn.com/abstract=3037178

percent increase in the initial aggregate real GDP is associated with a 1.0 percent increase in average size in the short run and a 1.2 percent increase in the long run. These results are not consistent with the notion that the effects of initial economic conditions on firm dynamics are short lived. Rather, they suggest that the average business size is strongly influenced by temporary economic shocks at the time of inception and, as such, the cohort lifecycle dynamics are path dependent and exhibit hysteresis.[1]

There are at least two broad economic forces that can explain the observed differences in firm dynamics across cohorts. First, systematic differences in the quality of businesses entering during economic booms and recessions could create differences in initial investment and growth patterns. The expected value of a new business and the outside options of entrepreneurs depend heavily on current economic conditions. Therefore, the quality of new entrepreneurs may vary systematically with the business cycle. To empirically examine this mechanism, I develop new data that tracks annual revenue at the firm level and build measures of revenue per employee in the LBD. I document that recessionary startups are, on average, more productive than expansionary startups. In addition, I find that the composition of businesses born during economic downturns is tilted toward sectors that require a greater amount of technical skill and entrepreneurial quality. Overall, these results suggest that the average quality of new entrepreneurs is countercyclical, which means that other economic forces must be responsible for the observed differences in initial investment and growth over time. The differences in average entrepreneurial quality should, if anything, mitigate the initial size discrepancy between expansionary and recessionary cohorts.

Second, the observed differences in firm dynamics across cohorts could also be caused by economic constraints on the ability of businesses to adjust their size following an initial investment. These mechanisms could take the form of capital adjustments costs or, alternatively, could be related to the inability of businesses to overcome demand-side constraints. I exploit differences in physical capital intensity across sectors (e.g. Hall, 2004) to formally examine if capital adjustment costs affect the observed persistence in size differences across cohorts. The findings suggest that differences in physical capital intensity are not significantly associated with differences in size persistence. I then use sectoral differences in the importance of product reputation and customer relations to gauge whether demand-side channels might explain the size persistence observed in the data. I employ two proxies for the sectoral importance of these demand-side channels: the share of total inputs spent on advertising and the product differentiation index developed by Gollop and Monahan (1991). I find that these proxies are significantly related to the persistence of size differences across cohorts. My empirical findings suggest that the process of demand accumulation through consumer reputation and brand awareness affects the ability of recessionary startups to catch up to their expansionary counterparts.

[1] This result is somewhat related to the extant literature documenting that wages of individuals graduating in recessions are negatively affected by the initial conditions even after these conditions have subsided (Kahn, 2010; Oreopoulos et al., 2012).

In the second part of the paper, I propose an analytically tractable framework that formally illustrates and quantifies the role of selection at entry and demand-side channels in explaining the entry and post-entry dynamics of cohorts that start in different stages of the business cycle. I develop a model of firm dynamics with endogenous entry and exit in the spirit of Hopenhayn (1992), in an environment where exogenous aggregate shocks affect the profitability of businesses. The model features an endogenous productivity distribution of entrants: potential entrants receive a signal about their idiosyncratic productivity and decide whether to enter the market based on the current state of aggregate demand. Incumbent firms are endowed with heterogeneous productivities that evolve over time, and they must decide how much to produce and whether to exit. In line with the observed empirical patterns, I introduce size persistence over the lifecycle by allowing the demand for each firm's differentiated product to be dynamic, such that if the business sells more today, it accumulates reputational capital and expands its future demand.

I calibrate the model to the U.S. economy and use the simulated data to estimate the age-period-cohort models used in the empirical section. When I parameterize the model to match a set of observed moments concerning the entry, survival, and size of businesses in the economy, the framework successfully replicates the growth and exit patterns by age. The simulated data generates estimates of the effect of economic conditions on the average size that are similar to those obtained from the actual data. In particular, the initial cohort size is also procyclical and the differences in the size of cohorts do not dissipate over time. I use the model to construct a counterfactual scenario in which I isolate the effect of reputational capital on firm dynamics by holding constant the composition of entrants across cohorts, effectively muting the impact of selection at entry. The wedge between the initial investment made by expansionary and recessionary startups is more than twice as large in this counterfactual scenario as when selection at entry effects are not muted. This result is due to the fact that favorable aggregate demand conditions reduce the average cohort size by incentivizing smaller and less productive businesses to enter the market.

I also employ my framework to analyze the persistence of size differences between cohorts over their lifecycle. In the counterfactual scenario, the difference in average business size between recessionary and expansionary cohorts decreases by approximately three percent per year. By contrast, the simulated average size difference in the baseline scenario is initially smaller but does not decrease over time. This result shows that the demand accumulation process induces a slow convergence of the initial size differences created by temporary aggregate demand shocks. Moreover, the endogenous entry and exit of businesses create a compositional effect that conceals this underlying convergence. The intuition for this result is subtle: a larger share of less productive businesses, which would not be started during bad periods, enter during economic booms. These businesses, which are also smaller and less resilient are more likely to exit as cohorts age, thereby further slowing the size convergence across cohorts. Furthermore, the idiosyncratic productivity process is mean-reverting, which also contributes to the dissipation of initial differences in the distribution of idiosyncratic productivities.

Finally, I use the model to quantitatively evaluate the role of the persistent effects of entry conditions

in the propagation of the Great Recession. I produce an exogenous shock in the model's aggregate demand to mimic the decline in the entrance of new firms that was observed during the Great Recession. My simulations show that the impact of the crisis on businesses born in 2008 and 2009 implied a reduction in aggregate employment by at least 1 percentage point in the following ten years. This accounts for a permanent loss of 1.2 million jobs in the economy.

This paper relates to a number of existing literatures. First, I contribute to the empirical literature on aggregate employment dynamics and its links to business microstructure. Previous literature recognizes that there is substantial heterogeneity in the composition of businesses in an economy and that understanding such variation is critical to study how economic shocks affect employment. Moscarini and Postel-Vinay (2012) documents that the negative correlation between net job creation and the unemployment rate is more pronounced for large firms. By contrast, Fort et al. (2013) shows evidence that younger and smaller businesses are more sensitive to business cycle shocks because the difference between the net job creation rate of young/small and large/mature businesses declines during recessions. More recently, Haltiwanger et al., 2013 provide evidence that age is more important than size in explaining employment creation by firms. Ouimet and Zarutskie (2014) and Davis and Haltiwanger (2014) suggest that the secular decline in the share of young businesses disproportionately affects the younger and less-educated individuals who are more likely to be hired by these firms. I contribute to this work by arguing that the cohort composition of businesses in an economy has important implications for its aggregate gross and net job creation. Moreover, my results suggest that the lower business dynamism of recessionary cohorts could have negative implications for the labor market outcomes of individuals that enter the labor market during downturns.

The facts presented in the paper also speak to an important literature on the propagation of aggregate economic shocks (Cogley and Nason, 1995; King and Rebelo, 1999). The notion that the economic conditions faced by a startup business at inception can have long-lasting consequences for their performance has received little attention by the literature. The exception is Sedlácek and Sterk (2014).[2] This paper studies cohort-level employment in the U.S., and documents that total employment variations across cohorts are found to be persistent, and largely driven by differences in average firm size. Unlike this study, I use micro-level data to empirically estimate how much aggregate economic conditions affect the average size and growth trajectory of businesses. Moreover, I use the granularity of the data to examine the mechanisms that govern this relationship and, guided by these empirical findings, I build a model that allows me to study shock propagation and quantitatively evaluate the role of persistent effects of entry conditions in the amplification and persistence of aggregate shocks.

This paper is also related to a literature that studies the relationship between entrepreneurial activity and business cycles (e.g. Caballero and Hammour (1994), Bernanke and Gertler (1989) Rampini

[2]There are also papers inspired by the organizational ecology literature, which suggest that founding conditions can have long-lasting effects because of the importance of initial strategic choices (e.g. Boeker, 1989) or initial stocks of financial and human capital (Cooper, Gimeno-Gascon, and Woo, 1994).

(2004)). Caballero and Hammour (1994) uses a "vintage model of creative destruction" to study the entry and exit of firms in response to fluctuations in aggregate demand and finds that startup rates in the economy depend heavily on the structure of startup costs. The literature also suggests that during economic booms, agency frictions between financial intermediaries and entrepreneurs are less pronounced, which allows entrepreneurs to borrow and invest more (e.g. Bernanke and Gertler (1989)). In this setting, there is a lagged procyclical mechanism running from output to entrepreneurship. Rampini (2004) proposes a model in which positive shocks to an economy increase the productivity of business activities, thereby making agents more willing to tolerate the risk of starting a firm. I add to this literature by examining not only how startup rates vary over the business cycle but also how the quality and evolution of these startups depend on initial business cycle conditions.

The paper also contributes to the literature that examines the role of cohort effects in the labor market. This literature suggests that the economic conditions faced by workers when they enter the labor market significantly affect their future earnings. Baker et al. (1994) documents that the initial wage of workers within a firm is affected by the business cycle and that workers who start with higher wages maintain this advantage over time. More recently, Kahn (2010) and Oreopoulos, von Wachter, and Heisz (2012) offer causal evidence that initial labor market conditions significantly affect the lifecycle wage profile of graduating students. Other work suggests that these effects are also found in the market for CEOs. Schoar and Zuo (2013) finds that the economic conditions when individuals become CEOs have a lasting effect on their careers and managerial styles. My paper expands this line of analysis to firm dynamics and shows that firms' post-entry dynamics are also significantly influenced by economic conditions at their time of entry.

Finally, this paper is also related to previous theoretical work on firm dynamics. Lee and Mukoyama (2008) and Clementi and Palazzo (2013) also use Hopenhayn's framework to study entry and exit decisions over the business cycle. More specifically, Clementi and Palazzo (2013) uses a model that includes capital adjustments and decreasing returns to scale, finding that entry and exit enhance the effects of aggregate shocks. In line with the empirical evidence documented in Foster et al. (2008) and Foster et al. (2015), firm growth dynamics are primarily driven by idiosyncratic demand shocks in their model. My empirical findings suggest that differences in input adjustment cost are not significantly associated with differences in size persistence. The model developed in this paper studies firm dynamics in a monopolistic competitive environment where demand plays a role in hindering the ability of businesses to adjust their size.

I begin the remainder of the paper by documenting the main empirical regularities on which the paper is based. In sections 3 and 4, I investigate the economic mechanisms that could explain the empirical patterns observed in the data. In section 5, I develop a model framework to assess the role of these economic mechanisms in explaining the documented empirical facts. Section 6 concludes.

2 The Persistent Effect of Entry Conditions

In this section I show that the average size of businesses varies with business cycle conditions at inception. I document that, at any stage of their lifecycle, businesses born in downturns are smaller than comparable businesses born in expansionary periods. To evaluate the magnitude of this correlation, I estimate cohort effects by implementing age-period-cohort models.

2.1 Empirical Methodology

Age-period-cohort models are used to isolate the effect of belonging to a certain cohort from the effects of the aging process and current economic conditions. This methodology is commonly used in the literature on individuals' lifecycle consumption and income dynamics. I extend it to the study of firm dynamics. In this setting, the cohort effects capture differences in average size across generations, irrespective of the particular period or age at which a business is observed. The age fixed-effects capture the impact of lifecycle patterns on the size of businesses, and the period fixed-effects capture aggregate shocks that affect the size of all businesses during a particular period, irrespective of their cohort and age.

I formally present the age-period-cohort models by expressing the outcome of interest (Y) of business j of type x as a linear summation of age (a), period (t) and cohort (c) functions, a constant, and an error term ($u^x_{j,ct}$):

$$Y^x_{j,ct} = \alpha^x + \Gamma^x_a + \Lambda^x_t + B^x_c + u^x_{j,ct}$$

in which Γ^x_a, λ^x_t, B^x_c represent a set of functions capturing the age, time, and cohort-effects, and x represents a homogeneous group of businesses that share the same industry, location, legal status, and multi-unit status.

Because there is no obvious pattern for the functions Γ^x_a, λ^x_t, B^x_c, I would ideally implement fixed-effects for age, time, and cohort and allow the data to determine the functional form. However, there is an exact linear relationship between the three effects. As a result the identification of the level effects for these three factors requires additional normalization or exclusion assumptions. The extant literature has dealt with this problem using multiple approaches. One possible approach is to treat age, cohort, and period effects as proxy variables for underlying variables which are not themselves linearly dependent (Heckman and Robb, 1985). In this paper, I use this approach by treating indicators of business cycle conditions at the time of inception as proxies for cohort fixed-effects. An alternative approach is to make case-specific normalizations. Deaton (1997) suggests a normalization that makes the year effects average zero over the sample period and be orthogonal to a time trend, such that all growth is attributed to age and cohort effects. The underlying assumption is that the current cyclical fluctuations are zero on average over the long run. I employ a similar approach in the paper as a robustness check.

The universe of businesses encompasses substantial heterogeneity that can affect the age, cohort, and time profiles. Using granular information from the LBD on the sector, location, and organizational form of each business, I account for some of that heterogeneity to better isolate the cohort effects.

In the baseline specification, I implement the age-period-cohort model presented above by assuming that cohort effects are a proxy for underlying economic conditions, which are measured using business cycle indicators at the time of inception. I formally evaluate the correlation between indicators of business conditions in the period when businesses begin operating and the size of these businesses throughout their lifecycle by estimating:[3]

$$\ln Y_{j,ct} = \alpha + \gamma_a + \lambda_t + \beta \ln Z_c + \text{controls} + u_{j,ct} \tag{1}$$

in which the functions γ_a and λ_t are age and period fixed-effects, respectively. Under the standard exogeneity restrictions, the main coefficient of interest, β, measures the average percent change in size resulting from a one percent cohort-specific variation in the business cycle indicator Z_c. This coefficient captures the elasticity of average size with respect to the economic conditions at entry after controlling for the economic conditions that each business faces over its entire lifecycle and for the effects of the aging process of the firm.

The model specified above does not allow the effect of cohort on size to vary over the lifecycle of the firm. However, it is plausible that the effects of economic conditions at inception gradually subside as a business becomes older. To investigate whether cohort effects disappear over the lifecycle of the business, I implement the following specification in which $\ln Z_c$ interacts with age:

$$\ln Y_{j,ct} = \alpha + \gamma_a + \lambda_t + b \ln Z_c + \kappa_a \ln Z_c \times Age + \text{controls} + u_{j,ct} \tag{2}$$

in which the parameter b is the elasticity of size with respect to fluctuations in Z_c in the first year of activity, and the parameter κ captures how the average elasticity evolves with each extra year of activity. I also consider a non-parametric specification in which I allow for cohort effects to vary non-monotonically with age:

$$\ln Y_{j,ct} = \alpha + \gamma_a + \lambda_t + \sum_{a=1}^{A} \kappa_a D_a \ln Z_c + \text{controls} + u_{j,ct} \tag{3}$$

in which D_a are indicator variables that take the value of one if the business establishment is a years of age. The series κ_a measures the effect of economic conditions at inception when the business is a years of age. This series allows me to examine the economic and statistical significance of the average elasticity of size with respect to business cycle conditions at inception at any point of the lifecycle of the business.

[3] I use the logarithm specification because the growth in the covariates seems to move size proportionately as the fit of alternative specifications are worse.

I consider an alternative formulation that does not rely on any specific proxy for the business cycle at inception. In this alternative framework, I follow Deaton (1997) and estimate the cohort effects non-parametrically:

$$\ln Y_{j,ct} = \alpha + \gamma_a + \lambda_t + \sum_{c=1}^{C} \tilde{\beta}_c \tilde{D}_c + \text{controls} + u_{j,ct} \tag{4}$$

where $\sum_{c=1}^{C} \tilde{\beta}_c = 0$ and $\sum_{c=1}^{C} c\tilde{\beta}_c = 0$. I apply the Deaton normalization to cohort effects rather than to period effects because the objective of this study is to capture the effect of entry year at business cycle frequencies. This normalization thus allows age and time effects to capture long-run trends in the evolution of the outcome of interest. I evaluate the robustness of the results obtained with the baseline specification by comparing the series of cohort fixed-effects estimated using this alternative method $\{\tilde{\beta}_1, ..., \tilde{\beta}_C\}$ with the series of business cycle indicators $\{\ln Z_1, ..., \ln Z_C\}$.

2.2 Data Description

2.2.1 The Life Cycle Outcomes of Businesses

In this paper, I study the evolution of business size using confidential data from the U.S. Census Bureau Longitudinal Business Database (LBD) and the U.S. Census Bureau Business Register (BR). The LBD is a longitudinally linked dataset that covers the entire universe of non-farm establishments and firms in the U.S. private sector that have at least one paid employee (Jarmin and Miranda, 2002).[4] This administrative database covers nearly every employer business in the U.S. and is an exceptional source of information on employment dynamics.[5] I use the BR to obtain data from annual business tax returns and compute a measure of revenue defined as the "value of shipments, sales, receipts or revenue." The LBD includes annual observations beginning in 1976 and currently runs through 2013, while the BR revenue data are only available starting in the mid-1990s.

The unit of observation in the LBD is the establishment, which is a single physical location where business is conducted or where services or industrial operations are performed. By using the appropriate firm identifiers it is possible to aggregate the establishment data into firm-level measures. A firm is a business organization consisting of one or more establishments that are under common ownership or control. Firms with multiple establishments often span multiple geographic areas and industries. I focus my empirical analysis at the establishment level and, as such, I avoid addressing questions

[4]The LBD employment and number of establishments tracks closely those of the County Business Patterns (CBP) and Statistics of U.S. Business(SUSB) programs of the U.S. Census Bureau (Haltiwanger et al., 2013).The CBP excludes crop and animal production; rail transportation; National Postal Service; pension, health, welfare, and vacation funds; trusts, estates, and agency accounts; employees of private households; and public administration. It also excludes most establishments reporting government employees.

[5]The LBD underlies the publicly available statistics provided by the Business Dynamic Statistics (BDS). Many of the patterns we discuss in this paper can be readily seen in the BDS. However, the LBD microdata has several advantages over the public domain data. First, it has a panel structure allowing to follow individual establishments over time. Second, it has more detailed information on industry and location, as well as firms ownership. Finally, it permits the construction of alternative measures of business birth and exit.

about the definition of ownership and control. The empirical evidence on decentralization among U.S. firms suggests that establishment managers have substantial hiring and investment discretion (Bloom, Sadun, and Van Reenen, 2010). In addition, changes in establishment size are independent of reallocation across firms due to mergers and acquisitions.[6] Moreover, an establishment has a precise location and sector, which are essential factors in the empirical analysis. In the 1978–2012 period, around twenty percent of establishments were part of multi-unit firms. To assuage concerns that my empirical findings are substantially different at the firm level, I also provide robustness results using the firm as the unit of analysis.

The identification of entries and exits by establishments is critical to my analysis.[7] I identify each establishment in the LBD as a birth, death, or continuer. A birth is defined as the period during which an establishment makes its first appearance in the universe of employer businesses.[8] A death is defined as the period when the activity of an establishment ceases.[9]

I use the level of employment of an establishment as the main measure of its size. The LBD defines employment as the number of full- and part-time employees on the payroll as of March 12th of each calendar year. Accordingly, I measure employment in year t as the number of employees as of March 12th of that year.[10] I complement this information on employment with revenue and productivity measures. The previous literature on revenue and productivity focuses on the manufacturing sector due to data limitations. I develop new data that tracks annual revenue at the establishment level (for single-unit firms) and at the firm level for the entire U.S. private sector. To my knowledge, this is the only paper other than Haltiwanger et al. (2015) to track revenue and employment outcomes for the entire universe of businesses. I use this new data to compute measures of labor productivity for all sectors.[11]

The universe of businesses encompasses substantial heterogeneity. Using detailed information from the LBD on business characteristics, I partially account for that heterogeneity. In my empirical analysis, I condition on the sector (mostly using 4-digit NAICS codes), geographical location, whether the establishment belongs to a firm with multiple establishments (single or multi-unit firm), and the

[6] Establishments can be acquired, divested, or spun off into new firms, so the ownership structure of firms can be very complex.

[7] Appendix A.1 provides extensive details on the longitudinal construction of the LBD.

[8] The moment in which an establishment enters the LBD may not necessarily coincide with the moment the establishment was effectively created. Some establishments may have started operating as a non-employer business, and then evolved into an employer business.

[9] Note that establishments that change ownership (sold to another owner) are not considered an establishment exit.

[10] I consider an alternative measure of employment in year t, defined as average number of employees of March 12th of year t and March 12th of year $t+1$.

[11] I was not able to obtain information on revenue for approximately 30 percent of establishments. Real revenue is defined as the value of shipments, sales, receipts, or revenue (in thousands), deflated by industry-specific value added deflators. BR's revenue measure is based on administrative data from annual business tax returns. Unlike payroll and employment, which are measured at the establishment level going back to 1976, the nominal revenue data are available at the tax reporting or employer identification number (EIN) level starting in the mid-1990s. Thus, in the SSEL, revenue is only measured at the establishment-level for single location firms. For multi-unit firms, we cannot obtain establishment-level measures, only firm-level revenues. Appendix A.2 provides details on the algorithm to obtain nominal revenue and the deflation method.

type of legal ownership of the business (e.g. corporation, self-proprietorship, partnership, or other organizational form). Appendices A.1 and A.2 offer a complete description of these variables.

The main sample includes all establishments born between 1978 and 2001 and their outcomes over ten years of activity (*sample 1*).[12] This means that the first cohort under analysis is the set of establishments created in 1978, which I track from 1979 through 1988, and the last cohort under analysis is the set of establishments born in 2001, which I track from 2002 through 2011. I exclude the outcomes of each cohort during its year of inception because of the difficulty in comparing employment levels for firms that start their operations before and after March 12th of each year. I include establishments that exit before they complete ten years of activity. Hence, this sample follows the evolution of establishment size conditional on survival.

To ensure that the results are robust, I employ other sample selection criteria. The frequent exit of firms from the sample raises the possibility of non-random attrition. To assuage these concerns, I consider an alternative sample (*sample 1.1*), which is restricted to establishments that were active for at least 10 years, and thus measuring the effect on employment conditional on being active for more than 10 years. Samples 1 and 1.1 are balanced in that I do not use more years for older cohorts. I also consider a sample consisting of all establishments born between 1978 and 2001 and their outcomes until 2011 regardless of the cohort under consideration (*sample 1.2*). Finally, the prior samples do not include the cohorts born from 2002 to 2010. These cohorts may have characteristics that are unlike those of any other cohort. Therefore, I also use an unbalanced sample of cohorts from 1978 until 2010 (*sample 1.3*), excluding the cohort born in 2002. In 2002 the U.S. Census made changes to the BR that resulted in an abnormal number of businesses classified as entrants during that year.

The longitudinal analysis of revenue and revenue per employee is restricted to a sample of establishments for which both revenue and employment information is available (*sample 2*). Revenue data is only available from 1994 on. Moreover, I cannot obtain revenue data for approximately 30 percent of the single-unit establishments for which I have employment information. To obtain a balanced sample with a sufficiently large number of cohorts, I restrict the panel to five years of activity. As a result, the main sample used to study revenue and revenue per employee includes the first five years of activity for establishments born between 1994 and 2007 (excluding the cohort born in 2002).

Table 1 reports summary statistics for the samples used in the empirical analysis. The main employment sample has around 12.8 million establishments and 78.3 million establishment-years. The sample that contains both employment and revenue information has approximately 3.8 million establishments and 14.5 million establishment-years. Importantly, the descriptive statistics suggest that there are not significant differences across samples regarding the summary statistics of the main dependent variables.

[12] The LBD does not allow us to determine the age of establishments that already existed in 1976. At the same time, there is some concern with larger measurement error in the first years of the LBD, therefore I consider only cohorts born after 1978 (Moscarini and Postel-Vinay, 2012).

2.2.2 Business Cycle Conditions at Inception

In my baseline specification, I treat cohort effects as a proxy for economic conditions at inception, which are measured by business cycle variables. The primary business cycle indicator in this paper is the annual real Gross Domestic Product (GDP). To capture the cyclicality of entry conditions, I apply a Hodrick-Prescott filter,[13] or compute the demeaned deviations of log differences. Hence the business cycle conditions for each entrant are expressed either in terms of the log deviation of the real GDP from its trend or the demeaned log differences in real GDP during the entry year. I also consider other indicators of business cycle conditions such as annual total employment, annual *per capita* personal income, annual unemployment rate, yearly average of the monthly Newspaper-based Index of Economic Policy Uncertainty, and the yearly average of the monthly National Financial Conditions Index from the Federal Reserve. Appendix A provides details on the sources of these variables. The Panel A of Table 2 reports summary statistics of these indicators between 1978–2012 and their correlations with the annual real GDP.

I also conduct additional empirical tests in which I use industry-specific and local indicators of economic conditions such as annual nominal GDP, aggregate employment and other composite proxies.[14] I take the industry output and employment measures directly from the BEA's industry accounts. I compute a composite proxy of downstream fluctuations by industry using the input-output tables and information on output by industry. This proxy measures the magnitude of demand shocks faced by businesses in a given industry. I provide details on the computation of this proxy in the appendix A. I also compute proxies of local economic conditions during the entry year. I define local markets at multiple levels, specifically state, county, and core-based statistical areas (CBSA) and I compute *per capita* annual personal income, and annual total employment at those levels.[15] Panels B and C of Table 2 report summary statistics for these indicators between 1978 and 2012.

2.3 Main Results

I begin my empirical analysis by evaluating the relationship between the size of businesses over their lifecycle and the business cycle conditions at inception. I estimate the baseline specification of equation (1) using the level of employment as the dependent variable and the log deviation of aggregate real GDP as the proxy for cohort effects. Table 3 reports the results of this analysis. The results presented in column 1 show that the coefficient of the proxy for business cycle conditions at inception is positive and statistically significant, i.e. recessionary startups are smaller, on average, than expansionary ones. The results are also economically significant: the elasticity of employment relative to fluctuations in aggregate real GDP at inception is around 1.2. This means that a one

[13] Using a smoothing parameter of 6.25, as suggested by Ravn and Uhlig (2002) for annual data.

[14] There is no information on real GDP at the level of the industry starting in 1978.

[15] I mostly use employment and per capita personal income instead of output because for counties and CBSAs there is not data on GDP for the entire period of analysis.

percent deviation of aggregate output from its non-linear trend affects the average size of establishments born in that year by more than one percent. Alternatively, I use a dummy variable that takes the value of one if the economy grew above its trend. The results reported in column 4 reveal similar findings using this indicator. Businesses that start their operations in years with above-trend economic growth are, on average, 2.1 percent larger.

Recent studies on the career effects of graduating during a recession suggest that there is a strong correlation between starting wages and economic conditions when individuals enter the labor market. This literature, however, documents that the effect of initial business cycle conditions on labor market outcomes fades over time (Oreopoulos et al., 2012; Kahn, 2010). Similarly, low demand, high uncertainty, and tight credit market conditions might constrain the size of businesses at inception. Yet, it is not clear whether the initial size differences across cohorts will decay over time once the negative effects of an economic downturn subside.

To examine whether the effect of business cycle conditions at inception disappears over time, I interact the detrended log real GDP with the age of the establishment as in equation (2). The interaction is positive and significant, which suggests that the cohort effects on size do not disappear as establishments become older. One possible concern with this result is that the cohort effects do not change monotonically as establishments age. To avoid imposing parametric assumptions, I allow the effect of each additional year to vary by interacting the business cycle variable with a set of indicator variables for the age of the business. These results, reported in column 3, further suggest that the elasticity of size with respect to aggregate business cycle conditions at inception does not decrease with age. In fact, there is no sign of reversion of the cohort effect: a one percent increase in the aggregate real GDP at inception is associated with a one percent increase in the initial average business size and a 1.2 percent increase in average size after ten years. To evaluate the robustness of this result I estimate the baseline specification using employment growth as the dependent variable. The results are reported in Column 5 of Table 3. The estimate is statistically significant and shows that a one percent deviation of aggregate output from its non-linear trend increases the average growth of establishments by 0.04 percentage points, which is consistent with the results presented above. These results challenge the notion that the effect of business cycle conditions at inception are not persistent over time.

One potential explanation for the lack of size convergence between the expansionary and recessionary cohorts is that ten years is a relatively short time horizon. I examine whether the cohort effects persist over longer horizons. I estimate equations (1)–(3) with an unbalanced sample of all establishments born between 1978 and 2001 and their outcomes for ages 1 or more. Using this sample I estimate the average elasticity of size with respect to business cycle conditions at inception over a 34-year horizon. I present this analysis in panel A of Table C.1 in the appendix. The results of the balanced and unbalanced sample are very similar, and there is no indication that the cohort effects disappear over the long run. Hence, these results further reinforce the notion that recessionary startups are smaller than comparable expansionary startups and that this gap does not close even over long horizons.

I illustrate the relationship between business cycle conditions at inception and the average size of establishments in Figure 1. I plot the estimated (log) size over the lifecycle of businesses in three simulated scenarios, "good," "neutral," and "bad" economic conditions at inception. I define a "bad" ("good") entry year as a two standard deviation negative (positive) difference from the detrended real GDP, whereas a "neutral" entry year is a year in which the real GDP grew according to trend. The simulated lifecycle patterns are computed using the estimated coefficients of specification (1) after normalizing the initial level of employment to the average size of a business with one year of activity. The vertical distance between the lines depends heavily on the elasticity of size with respect to business cycle conditions at inception $\hat{\beta}$. Panel A plots the simulated paths using the balanced sample (*sample 1*) which follows businesses for 10 years. Panel B plots the simulated paths using the unbalanced sample (*sample 1.2*) which follows businesses over longer periods. Figure 1 further illustrates that the cohort effects are economically significant. On average, a three-year-old establishment born in a "good" year employs the same number of employees as a four-year-old establishment born in a "bad" year. Moreover, the plots suggest that the effect is stable, since the vertical distance between the "bad" and "good" cohorts does not fade over time.

Next, I formally examine whether the previous results remain economically and statistically significant if I use alternative business cycle indicators. In column 1 of Table 4, I replicate the estimation of equation (1) using the demeaned log change in real GDP as an alternative method to capture the cyclical component of a series. The results reported in column 1 of Table 4 suggest that the prior results are robust to this alternative detrending method. I also use the change in total aggregate employment and unemployment rate during the inception year as business cycle indicators, thereby focusing the analysis on labor market conditions. The results are reported in columns 2 and 3 and suggest that using indicators of aggregate labor market conditions does not significantly alter the prior results. Finally, I also use business cycle measures of personal income, aggregate uncertainty (the HP-filtered natural logarithm of the newspaper-based index of economic policy uncertainty), and financial constraints such as the Chicago Federal Reserve's National Financial Condition Index (NFCI). The only indicator for which I could not establish a robust empirical relation with establishment size was the uncertainty proxy. Overall, these results support the idea that the effects of business cycles on the average size of cohorts could stem from many channels including financial constraints, labor market conditions, or the level of aggregate demand at inception.

Another possible concern with the baseline results is that the empirical patterns that I document are not specific to business cycle conditions in the year of inception. Prior literature (e.g. Fort et al. (2013)) suggests that the current state of the economy affects older businesses as well. To examine this possibility, I assess whether the effects of business cycle conditions at inception have greater effect on the average size of cohorts than the economic conditions that these cohorts experience later in life. I re-estimate equation (1) using business cycle conditions when the establishments are five years old rather than at inception as the main variable of interest. The results, reported in Table 5, suggest that the coefficient associated with business cycle conditions when the establishment is five

years old is neither statistically nor economically significant. This result suggest that the economic conditions that businesses experience once they are more established do not place an imprint on their future the way conditions at inception do. This empirical pattern remains robust even when I consider the business cycle conditions at other ages as the main variable of interest.

A related issue is that the previous estimates capture not only the economic conditions at inception, but also the weighted sum of the economic conditions that a certain cohort faces over its lifecycle. In other words, because a bad year is likely to be followed by another bad year, the main effect could capture the fact that recessionary startups face more bad years than good years over their lifecycle. Year fixed-effects may not fully control for this effect if young businesses are more sensitive to the business cycle (e.g. Fort et al. (2013)). To isolate the effect of the business cycle at inception, net of the effects on size from exposure to a possibly prolonged recession, I estimate equation (1) using the log deviation of real GDP from its trend in the inception year while further controlling for log deviations of real GDP from its trend for years one through five. The results in columns (2) and (4) of Table 5 show that the magnitude of the effect is larger in the entry year than in any other year. This approach offers further support for the notion that the persistent effects of aggregate economic conditions at time of entry are substantially greater than the effect of business cycle conditions later in a business's life.

The set of results presented above uses an age-period-cohort model, in which I attain identification by proxying for cohort effects using underlying variables that are not themselves linearly dependent. An alternative approach is to apply Deaton's normalization to the cohort effects as indicated in specification (4). Figure 2 plots the estimated effects for 24 cohorts using this alternative procedure overlaid on shaded areas indicating NBER recession periods. This plot strongly suggests that NBER recession periods coincide with low cohort-effect values. I correlate the series of estimated cohort fixed-effects with the detrended real GDP and other proxies for economic and financial conditions. Table 6 reports these correlations. The correlation between the deviation of real GDP from its trend and the change in non-parametric cohort effects is 0.52. I interpret these results as supplementary evidence that the cohort effects exist and that they are significantly positively correlated with the business cycle conditions.

I conclude this section of the empirical analysis by conducting additional robustness tests. I start by re-estimating my baseline model using a sample of establishments that survived for at least 10 years. I use this alternative sample to assess whether attrition bias related to the exit of firms from the sample distorts the empirical results. I find that the effect of business cycle conditions at inception on businesses that are active for at least 10 years,is only marginally larger than that of the main sample, and the effects of business cycle conditions also persist through time as businesses become older.

Finally, I also use revenue as an alternative measure of business size. Employment level is a convenient measure of size because the LBD has recorded it for the entire universe of non-farm U.S. employer establishments since the late 1970s. Hence, using employment instead of revenue allows me to analyze

a larger number of cohorts over a longer period. However, using employment as a measure of size could bias the empirical analysis. For instance, to the extent that wages are more cyclical than interest rates (King and Rebelo, 1999), new businesses may be relatively more labor intensive in downturns, and therefore employment level would tend to overestimate business size. As an alternative, I use revenue, which is more directly related to the concept of output, to measure business size. The availability of revenue data is, however, more limited: Revenue data are not available at the establishment level and are only available for recent periods.[16] I estimate the baseline model using log real revenue as the dependent variable for a sample of cohorts born between 1994 until 2007 and follow them for 5 years (*sample 2*). Table C.4 of the Appendix reports the results. The results indicate that (single-location) recessionary startups have 0.9 percent less revenue than expansionary ones.

2.4 Local and industry-specific business cycles

The previous set of results uses national-level business cycle indicators. However, the most relevant indicator for entrepreneurial choice may be the business cycle conditions at the local or industry level. This would be the case if, for instance, the market for entrepreneurial activity is segmented within the United States or if entrepreneurs have industry-specific expertise that cannot be transferred across industries. In those cases, the relevant business cycle indicators for entrepreneurial choice may be defined more precisely.

I start by re-estimating equation (1) using the demeaned log changes in *per capita* personal income and demeaned log changes in total employment as the indicator of local business cycle conditions at inception.[17] Table 7 presents the results. The results suggest that business size is positively correlated with the business cycle conditions of the local economy, regardless of the local indicator used. The magnitude of the effect is similar across national and regional specifications. A one standard deviation change in the within state/county/CBSA-level demeaned total-employment growth rate is associated with a 0.8/1.4/1.0 percent variation in the average business size. These elasticities are comparable to the elasticity of 1.2 obtained using a similar national-level indicator. [18]

I also follow the strategy of several recent papers (e.g. (Mian and Sufi, 2012; Adelino et al., 2013)) and use the differential impact of local demand shocks on tradable and non-trabable industries to analyze the effects of demand fluctuations on average cohort size. The idea is that tradable industries are more sensitive to aggregate business cycle conditions than non-tradable and construction sectors, and should therefore be less responsive to aggregate economic conditions. The results reported in Table 8 provide partial support for the hypothesis: the national-level business cycle has a larger impact

[16]Receipts data varies substantially by type of activity (industry) and the legal structure of the firm according to different tax treatments. These problems of comparability across sectors are less relevant in the context of this paper because I include a set of industry fixed-effects at a very detailed level.

[17]I use employment and *per capita* personal income instead of output because for counties and CBSA's there is not data on GDP for the whole period under analysis.

[18]Note that because I control for local fixed-effects, I only use time-series variation within geographies.

on the average initial size of businesses that produce tradable goods. By contrast, I find that the size elasticity of firms with respect to state-level economic conditions at inception is not significantly different between tradable and non-tradable sectors. Nevertheless, the relevant geographical market for non-tradable and construction industries might be defined at a much more localized level and, as such, these results could simply reflect the fact that the state-level GDP is a poor proxy for local economic conditions. Overall these results suggest that aggregate demand shocks are an important determinant of differences in initial size of the cohorts.

Next, I exploit industry variation in demand conditions at inception to further assess the role of demand fluctuations in determining average business size by cohort. I do so by using the BEA industry accounts to compute two different indicators of industry demand. The first uses nominal added value by industry, and the second is a composite indicator that uses supply chain information to compute the industry-specific variation in demand from downstream buyers.[19] This second indicator accounts for the direct (first-order) effect on output demand of an industry i that stems from demand shocks to its immediate downstream buyers, whether they are industries or final consumers.

The results of the analysis are presented in Table 9. In column (1), results are computed using the nominal added value indicator of industry demand conditions. The coefficient is positive and statistically significant, which indicates that positive demand conditions at the industry level have a positive impact on the average size of startup cohorts in that industry. However, the economic magnitude of the elasticity of average cohort size with respect to nominal added value in the industry is significantly lower than that obtained in the main analysis. A plausible explanation is that the nominal added value by industry captures not only business cycle fluctuation in demand, but also variation in the industrial cost structure. I attempt to isolate the effect of industry demand by using the composite indicator to evaluate cohort effects at the industry level. The results, reported in column (2) of Table 9, indicate that the elasticity of size with respect to the composite industry demand proxy is statistically and economically significant. A one percent increase in the deviation of industry demand from its trend results in a one percent increase in the average cohort size. This result is consistent with the notion that demand fluctuations are a very important determinant of the size wedge expansionary and recessionary startups. Moreover, the results in columns (3)–(5) suggest that these inferences are robust to other proxies of industry demand.

3 Selection at Business Cycle Frequencies

The set of results presented in the previous section document a strong positive correlation between economic conditions at inception and the average size of cohorts born in that period. One potential explanation for this finding is that businesses entering at different stages of the business cycle are

[19] The industry-specific demand proxy is computed as $s_{it}^D = \sum_j \omega_{ij}^D \tilde{y}_{jt}$, in which $\tilde{y}_{it}$ is defined as the HP-filtered output of industry j in year t, and ω_{ij}^D is purchaser j's share of industry i's total sales.

different in terms of unobservable characteristics that affect their size over the lifecycle. In particular, a positive correlation between average size and the state of the economy at entry could simply indicate that better entrepreneurs are more likely to start businesses during economic booms.

I implement two strategies to investigate the relation between the average quality of new entrepreneurs and the state of the business cycle. First, I study a measurable dimension of business quality – labor productivity – and examine how it fluctuates with business cycle conditions at inception.[20] Second, I exploit the relationship between economic conditions and the likelihood of entry to make inferences about the variation in the average quality of cohorts across the business cycle. The empirical evidence suggests that the average quality of entrants is countercyclical, suggesting that selection mechanisms do not fully account for the differences in average size of businesses across cohorts.

3.1 Average Productivity of Entrants

I begin this section by documenting the relationship between a direct measure of productivity and business cycle conditions at inception. Labor productivity at the establishment level is a convenient measure of efficiency and entrepreneurial talent. The prior literature (Syverson, 2011) suggests that high productivity is associated with several positive outcomes such as higher growth and greater likelihood of survival. I evaluate the evolution of labor productivity for different cohorts by estimating the following equation:

$$\ln LP_{j,ct} = \alpha + \gamma_a + \lambda_t + \beta^{LP} \ln Z_c + \text{controls} + \varepsilon_{j,ct} \tag{5}$$

in which $LP_{j,ct}$ is the revenue per employee and all other terms are defined as in prior equations. As in my previous estimations, standard errors are clustered at the state-level.

The results indicate that expansionary startups are, on average, less productive. Column 1 of Table 10 reveals that the aggregate real GDP at inception is negatively associated with labor productivity. The effects are economically significant: a one percent increase in the cyclical component of output is associated with a 0.3 percent decrease in labor productivity. This result is consistent across alternative proxies of business cycle conditions.

Next, I evaluate whether the differences in productivity across cohorts are short-lived or persist over the first five years of activity.[21] The results in columns 2 and 3 indicate that the interaction terms

[20] Another important measure of quality of cohorts can be inferred from the survival patterns of the cohorts. I conduct some preliminary empirical exercises in which I investigate if firms born in bad times are more or less likely to exit in the first years of activity. In the appendix C (Figure C.1), I plot the Kaplan-Meier estimator of the survival functions for cohorts born in years that saw above and below trend growth in output, respectively. The estimated survival functions are very similar. In Section 5, I show that my model generates a similar outcome, even though there are systematic differences in the quality of entrants.

[21] As mentioned in the data section, I use the *sample 2* to conduct the longitudinal analysis of revenue per employee. I consider alternative samples to check if the results are robust to this particular selection. The results are available upon request. I also reestimate equation 1 that use employment as the main dependent variable in this sample. The

between age and business cycle conditions are not significant, thereby suggesting that the differences in productivity across cohorts do not revert in the first years of activity. In columns 4–6, I repeat the analysis used for columns 1–3 after weighting each observation by the respective level of employment. The results are broadly similar but not statistically significant. In figure 3, I plot the estimated productivity profile for expansionary and recessionary cohorts during their first five years of activity. The figure further illustrates that recessionary startups are more productive than expansionary ones.

The above findings are likely to be robust to alternative measures of productivity. The LBD and the BR do not have data on capital and other inputs that would allow the computation of measures of total factor productivity such as those used in studies that focus on the manufacturing sector. Nevertheless, I do not expect the countercyclicality of productivity to be significantly affected by the nature of the measure for two reasons. First, even in capital-intensive sectors such as manufacturing, the correlation between total factor productivity and labor productivity is approximately 0.4. (Foster et al., 2001). Second, Lee and Toshihiko (2015) also reports that the total factor productivity of new businesses is countercyclical.[22]

Overall, this evidence supports the idea that the productivity distribution of entrants varies predictably across booms and busts. This empirical pattern is consistent with a basic framework in which entrepreneurs have some private information about their productivity and decide to start operations when their expected productivity is above a certain threshold.

3.2 Number and Composition of Entrants across Cohorts

In this section, I attempt to provide further evidence for the role of selection mechanisms in mediating the relation between business cycle conditions at inception and the persistent differences in average size across cohorts. I do so by studying changes in the composition of entrants over the business cycle.

The unit of observation in the ensuing analysis is a market segment, which is defined as a set of businesses that share the same U.S. state, industry classification, multi-unit status, and type of legal ownership.[23] I analyze the relation between economic conditions and the number of entrants in a given period within each market segment to gauge if this relation is predictably affected by segment characteristics. I start by computing the correlation between the number of entrants and business cycle conditions for each segment. The distribution of the correlation coefficients suggests that there

significance and magnitude of the effect is equivalent to the ones presented above. The results are presented in the appendix.

[22] Lee and Toshihiko (2015) compute measures of total factor productivity of new plants (as well as continuing and exiting plants) in the manufacturing sector, and report a negative correlation with the annual growth rates of manufacturing output.

[23] Results are robust to several alternative specifications of market segments. Alternative definitions of segments are available upon request.

is substantial variation in the cyclicality of entry across segments.[24] To examine this heterogeneity I estimate the following equation:

$$\Delta E_{k,c} = \eta^E \Delta Z_c + \zeta_x^E \Delta Z_c X_k + v_{k,c} \tag{6}$$

in which $\Delta E_{k,c}$ is the demeaned log change in the number of entrants in market segment k in year c, ΔZ_c is the business cycle indicator and X_k represents the characteristics X of market k. The coefficient ζ_x^E indicates how the relation between the entry margin and business cycle conditions varies with characteristic X. A positive estimate indicates that an increase in that characteristic increases the cyclicality of entry.

I report the results of this analysis in Table 12. In columns 1 and 2, I explore how the entry decisions of established businesses and corporations vary with the state of the business cycle. The results suggest that the entry rate of single-unit businesses, sole-proprietorships, and partnerships is significantly more procyclical than that of other businesses. These results are consistent with the earlier finding that the average quality of entry cohorts is countercyclical: established businesses and corporations, which are often larger and more productive, are more likely to start new establishments during economic downturns. These findings also go against the idea that a large number of relatively low-quality entrepreneurs start businesses during economic downturns because it is their second-best option once they become unemployed (Ghatak et al., 2007).[25]

I also examine whether the relation between entry decisions and business cycle conditions depends on the type of technology employed in each sector. The idea is to evaluate whether entry is more or less procyclical in sectors that employ more complex technologies, and therefore may require greater entrepreneurial skill. I collect information on the input mix and innovative activity of each sector and merge those sectoral characteristics with the LBD. In columns 3 and 4, I evaluate whether entry is more procyclical in skill- and capital-intensive sectors. Using the Current Population Survey (CPS), I obtain the joint distribution of industries and educational levels to compute measures of skill intensity by industry. As a proxy for capital intensity, I use the average log ratio of capital stock to total employment, which I obtain from the NBER-CES Manufacturing Industry Database.[26] The results reveal that the entry rates of skill- and capital-intensive industries are less procyclical. I also collect information on the importance of innovation in each sector, which I define as the proportion of businesses in that sector that developed or introduced a new or significantly improved product or

[24] Figure C.2 in the Appendix plots the density of the correlation between $\Delta E_{k,c}$ and the demeaned log differences in real/nominal output (aggregated and industry-specific), for k defined by the sector (Panel A) or k defined based on the state (Panel B). The correlations range from values near zero to correlations above 0.5 and the variance of the distribution is larger when I explore differences across sectors than across states. This distribution of coefficients suggests that there is a large amount of heterogeneity that I can exploit to learn about changes in the composition of entrants at business cycle frequencies.

[25] Moreira (2014) uses the Current Population Survey to examine the flows in and out self-employment. The paper offers evidence that the likelihood that someone becomes an entrepreneur out of necessity does not vary substantially with the business cycle.

[26] Detailed information on the construction of these variables can be found in appendix A.

process. The results presented in column 5 suggest that the entry rate in innovative industries is less procyclical than in less innovative industries.[27] Overall, these results suggest that sectors that are likely to require sophisticated and well-educated entrepreneurs are less procyclical, thereby further supporting the notion that average quality of entrepreneurs is countercyclical.

Finally, I explore the relationship between entrepreneurship and financing constraints to make further inferences about the cyclicality of entrepreneurial quality. Using a strategy similar to that of Adelino et al. (2013) and Hurst and Lusardi (2004), I take advantage of the importance of collateral for small business financing and look at variation across industries in the amount of startup capital needed to set up a firm. The notion is that some businesses, like cleaning services, can be started with small amounts of capital that may reasonably be financed by the personal wealth of the entrepreneurs or by debt financed through their personal balance sheets. By contrast, other sectors require large amounts of startup capital that can only be financed via external debt sources that cannot be financed via individual loans. Hence, entry in these sectors could potentially be more affected by a credit crunch during business cycle downturns. I investigate the cyclicality of entry in industries with different levels of average startup capital. The empirical findings suggest that low average startup capital industries are more procyclical than other industries. This result is consistent with the idea that the credit channel may disproportionately affect industries with a higher prevalence of small-scale entrepreneurs who are more likely to depend on bank financing through their personal balance sheets to begin operations. This result is consistent with those of Siemer (2014) who argues that a reduction in bank lending affected small and young firms most directly during the 2007–2009 recession. Also, Saffie and Ates (2013) studies the impact of the Russian sovereign default on the Chilean manufacturing firms and finds that the resulting credit shortage was associated with a lower number of entering firms.

Overall the results on productivity and the composition of entry across the business cycle suggest that the average quality of entrepreneurs entering during an economic downturn is higher than that of entrepreneurs entering during economic booms.

4 Economic Mechanisms of Persistence

Next I assess the role of certain economic mechanisms in explaining the observed differences in initial investment size across cohorts and the persistence in these size differences. As described in the previous section, the average productivity of the pool of startup entrepreneurs is countercyclical. This evidence suggests that in the absence of the observed countercyclicality of entrant productivity, the relation between the initial average business size and business cycle conditions would be even more procyclical. A potential empirical test of this hypothesis is to explore the cross-sectional

[27] All specifications seem to point to a negative coefficient of the interaction between the business cycle indicator and the average innovative rate. However, depending on the proxies for aggregate business cycle, the coefficient may not be significant.

heterogeneity in the data to investigate whether industries that have a less procyclical entry margin are more procyclical in terms of their average initial size. I then examine the economic factors behind the persistent size differences across cohorts. I also exploit cross-sectional differences in industrial characteristics to shed light on the mechanisms that explain the lack of reversion of size differences across cohorts.

4.1 Entry Size

I begin by formally investigating whether the correlation between average initial business size and the business cycle indicator increases when I hold the observed sectoral composition of cohorts constant. The results reported in Table 13 provide evidence that it does. The correlation between average initial size and aggregate real GDP increases from 0.40 to 0.58 when I account for differences in the composition of cohorts. This result offers suggestive evidence that entry size is more procyclical once I condition on the sectoral composition of cohorts and focus on variation within business types.

Next, I use information about the characteristics of each business establishment to investigate the economic forces that amplify or attenuate the differences in average initial size across cohorts. To carry out this analysis, I employ the following framework:

$$\tilde{y}_{k,c} = \eta^y \tilde{Z}_c + \zeta_x^y \tilde{Z}_c X_k + \varepsilon_{k,c} \tag{7}$$

in which k represents a market segment, defined as the interaction between the location, industry, multi-unit status, and legal ownership of the business, $\tilde{y}_{k,c}$ is the demeaned log change in the average size of entrants in market segment k and cohort c, $\tilde{Z}_c$ is the business cycle indicator and X_k represents the characteristics X of group k. A positive estimate for ζ_x^y indicates that an increase in characteristic X increases the cyclicality of the entrant size during a given year. This empirical framework is designed to provide descriptive evidence about which characteristics X_k amplify the relation between average entrant size and aggregate real GDP.

The results are presented in Table 14. Column 1 reports that the initial size of new establishments of multi-unit businesses fluctuates more with the business cycle than the initial size of new standalone businesses. Similarly, the results presented in column 2 suggest that the average initial investment decisions of corporations are more cyclical than those of sole proprietorships and partnerships. Next, I examine the effects of certain technological characteristics on the cyclical properties of the average business size. The results presented in columns 3–7, suggest that the average entry size in capital-intensive sectors, innovative industries, and sectors with greater minimum efficient scale is more procyclical than that of other sectors. By contrast, I do not find any statistically significant differences in average initial size of skill-intensive sectors. Overall, these results indicate that market segments with less procyclical extensive margin have are more procyclical intensive margins of adjustment. That is, entry decisions by businesses in these segments are less responsive to business cycle conditions,

but their average initial size is more responsive to current market conditions. These results provide suggestive evidence that the effects of business cycle conditions at inception on average initial size are stronger when the selection mechanisms are muted.

4.2 Persistence Mechanisms

In the previous sections, I document that businesses born during recessionary periods start smaller and remain smaller over their lifecycle. Here, I empirically assess which economic mechanisms are associated with the persistence of these size differences.

I consider two broad economic channels that could explain the observed persistence of cohort effects in the average size of businesses over time. First, technological constraints and capital adjustment costs could hinder the ability of businesses to quickly adjust their size and scale to the optimal level. Standard investment models (e.g. Caballero (1999), Cooper and Haltiwanger (2006)) suggest that under the assumption of convex adjustment costs businesses are slow to adjust their capital stock. These adjustment costs could reflect a variety of interrelated factors such as disruption costs during installation of new equipment, firing and hiring costs, and the irreversibility of projects due to the lack of liquidity in the secondary markets for capital goods.

To test the relation between capital adjustment costs and the persistence of cohort effects, I explore variation in capital intensity across industries. The idea underlying this empirical test is that the magnitude of capital adjustment costs is greater than that of labor adjustment costs (e.g. Hall (2004)). Therefore, capital-intensive industries are likely to face greater adjustment costs

The second mechanism that could generate persistence across cohorts is demand-side constraints. Foster et al. (2015) shows that in certain industries the difference in size between incumbents and new establishments do not reflect productivity gaps, but rather differences in the individual demand for their products. New businesses enter small because initial demand for their product is low due to informational, reputational, or other frictions (Perla (2015), Gourio and Rudanko (2014), Luttmer (2006), Ravn et al. (2006)). Overcoming these frictions is a slow process that depends on repeated customer interactions and customer learning through 'word of mouth' and other media. Hence, recessionary startups in these industries may experience difficulties converging toward the size of their expansionary counterparts.

I investigate the relation between demand-side frictions and the persistence of cohort effects by taking advantage of the fact that demand frictions are more important in sectors with low product-market substitutability. In these markets, consumers cannot easily switch between producers and, as a result, reputation and repeated interactions are more important. (e.g. Gourio and Rudanko (2014)). I employ two proxies for the importance of reputation and product-market substitutability across sectors: the share of advertising in total inputs and the product differentiation index of Gollop and Monahan (1991). The share of advertising in total inputs directly measures variation in the importance of brand recognition and reputation across sectors. I also use the the generalized differentiation

index of Gollop and Monahan (1991) to proxy for differences in product-market differentiation across sectors. This index is a composite measure that accounts for the number different products in an industry, the inequality of the production shares of specific products within an industry, and the dissimilarity of products as measured by the input shares of various intermediate products used to make them.

To formally examine whether capital adjustment costs and demand-side frictions increase the persistence of cohort effects, I estimate the following equation for each four-digit NAICS industry:[28]

$$\ln Y_{ct}^{i} = \alpha^{i} + \gamma_{a}^{i} + \lambda_{t}^{i} + b^{i} \ln Z_{c} + \kappa^{i} \ln Z_{c} \times Age + \text{controls}^{i} + u_{j,ct}^{i} \tag{8}$$

in which $i = 1, ..., I$ represents the ith four-digit NAICS industry, $controls^{i}$ include U.S. state, multi-unit status, and legal ownership type fixed effects, and all other variables are defined as above. Next, I retrieve the industry-specific coefficients κ_{a}^{i} that measure the persistence of the size elasticity with respect to aggregate business cycle conditions at inception in each industry i. I then examine whether κ_{a}^{i} is associated with the aforementioned industry characteristics.

I report the results in Table 15. The results in columns 1 and 2 suggest that there is no statistically significant association between capital adjustment costs and the persistence of cohort effects. In fact, industries with greater capital adjustment costs, as measured by their physical capital intensity tend to have lower persistence of cohort effects. The results presented in columns 3–6, however, indicate that demand-side frictions are a significant determinant of cohort size persistence. The share of advertising in total inputs and the product differentiation index are significantly associated with greater persistence of cohort size differences. In columns 7–10, I include proxies for both capital adjustment costs and demand-side frictions. The sign and magnitude of the coefficients are similar but they are no longer statistically significant at conventional levels due to a lower number of observations in these tests. These results can also be seen in Figure 4, which plots the coefficients κ_{a}^{i} against the proxies for capital intensity and the differentiation index.

The results above provide evidence in favor of the hypothesis that the persistent effects of initial economic conditions is correlated with demand frictions. Recessionary startups enter with a smaller size than expansionary ones, which means that their initial customer base is relatively small. As businesses grow older, they are able to overcome these frictions and generate more demand. However, their growth is constrained by their initially small customer base. In the next section, I develop a model framework that formally illustrates how these demand frictions affect the post-entry dynamics of cohorts that start operations at different stages of the business cycle.

[28] For less than 5 per cent of industries, there were not enough observations to allow for such a detailed specification.

5 Model of Firm Dynamics with Dynamic Demand

In this section I develop an industry model of firm dynamics in an environment with exogenous aggregate demand shocks that affect the profitability of businesses. Motivated by the empirical facts documented in the previous sections, the model features endogenous entry and exit and a demand-side mechanism that hinders the ability of firms to adjust their size. I quantitatively explore the model with two main purposes: First, I isolate the effect of an aggregate shock by holding constant the composition of each cohort to identify the selection-free elasticity of size with respect to the business cycle conditions at entry. Second, I quantify the role of persistence of initial economic conditions on the propagation of aggregate shocks in the short and long run.

The model framework largely follows that of Hopenhayn (1992), but departs from it in three fundamental aspects: I add aggregate shocks to study the business cycle implications of the model; I allow for the productivity distribution of entering and exiting business to vary endogenously with the aggregate shock;[29] and the product market is characterized by a monopolistic competitive environment, while the input market is characterized by a perfectly elastic supply of labor. Some of these new elements such as aggregate shocks and endogenous productivity are also found in the models of Clementi and Palazzo (2013) and Lee and Mukoyama (2008). Yet, these papers do not analyze all these aspects at once and are not interested in the sort of analysis that I develop here.

The monopolistic competitive environment is essential for allowing demand-side factors to play a role in hindering the ability of businesses to adjust their size. The idea is that the economy is characterized by informational frictions concerning product characteristics, so businesses invest in customer relationships and product reputation to overcome these frictions. I formalize this intuition in the model by defining customer capital as a state variable. The demand for a firm's differentiated product depends on the stock of idiosyncratic customer capital b_{it} which increases with quantities sold in the past. Thus, firms can invest in customer acquisition by increasing their sales effort today.

5.1 Setup

In this model, time (indexed by $t = 1, 2, ...$) is discrete and agents face an infinite horizon. There are two types of agents: incumbent firms and potential entrants. Firms, indexed by i, are heterogeneous and their time-varying idiosyncratic productivities s_{it} evolve according to a persistent first-order Markov processes that is independent among firms. The process is defined as:

$$\log s_{it} = \rho_s \log s_{it-1} + \sigma_s \epsilon_{it} \tag{9}$$

[29] In Hopenhayn (1992) potential entrants receive a productivity draw after the decision to enter, so that the productivity distribution of entrants in the economy is always the same and is given exogenously. The empirical results of Section 3 show that the productivity distribution of entrants vary endogenously with the business cycle, which seems to indicate that entrepreneurs must have some sort of signal on their productivity and use this information to make entry decisions.

where $\epsilon_{i,t} \sim \mathcal{N}(0,1)$ for all $t \geq 0$. The conditional distribution of s_{it} is denoted by $H(s_{it+1}|s_{it})$. The demand for a firm's differentiated product depends on a common aggregate demand shock z_t that also evolves as a persistent first-order Markov process, defined as:

$$\log z_t = \rho_z \log z_{t-1} + \sigma_z \varepsilon_t \tag{10}$$

where $\varepsilon_t \sim \mathcal{N}(0,1)$ for all $t \geq 0$. The conditional distribution of z_t is denoted by $F(z_{t+1}|z_t)$.

Demand

Each firm produces a differentiated product and faces a dynamic demand due to the presence of demand accumulation processes (e.g. building a customer base). At time t, the isoelastic contemporaneous demand function of firm i depends on its price p_{it}, stock of previous sales b_{it}, and a common aggregate shock:

$$y_{it} = \alpha p_{it}^{-\rho} b_{it}^{\eta} z_t \tag{11}$$

in which $\eta \in (0,1)$ measures the elasticity of demand with respect to customer capital and $\rho > 1$ is the demand price elasticity. The stock of customer capital evolves according to:

$$b_{it} = \begin{cases} (1-\delta)b_{it-1} + (1-\delta)p_{it-1}y_{it-1} & \text{if } i \text{ is produced in } t-1 \\ b_0 & \text{otherwise} \end{cases} \tag{12}$$

in which $\delta \in (0,1)$ is the depreciation rate of customer capital. This mechanism creates persistence in the dynamics of production and employment. By selling more today, businesses acquire customer capital and expand their future demand. The optimal production level in this case is higher than in a purely static problem in which current sales are not tied to future demand.

Technology

The production function is linear in labor (n_{it}) and is given by:

$$y_{it} = s_{it} n_{it} \tag{13}$$

in which n_{it} is the quantity of labor, whose spot market price is w_t and is taken as given. There is an i.i.d. fixed operational cost $f \sim G(f)$ that all firms must incur to produce every period. The cost is stochastic and unknown in period $t-1$. I assume that the distribution of the operating cost is log-normal with parameters μ_f and σ_f.

Incumbent's Problem

At the beginning of each period, the industry has a mass of incumbent firms Ω_{t-1}. Each incumbent firm has a stock of customer capital b_{it} and observes the aggregate demand shock z_t and its own idiosyncratic productivity s_{it}. Using this information set, it decides its optimal production level and whether to continue operating and pay the fixed cost f. Firms discount future profits at the

time-invariant factor β. Even if a firm decides to continue it may be hit by a random exit shock with probability $\gamma \in (0,1)$. Firms that exit cannot return to the market at a later stage with their capital stock intact. Once a firm exits, it loses its reputation capital. I normalize the outside option to zero. The incumbent firm solves the following functional equation:

$$\begin{aligned} V^I(b,s,z) = &\max_{y,p,b'} \Big(p - \frac{w}{s}\Big)y + \int \max\Big\{0, -f + \beta(1-\gamma) \iint V^I(b',s',z')\, dH(s'|s)\, dF(z'|z)\Big\} dG(f) \\ &\text{s.t. } b' = (1-\delta)(b+y) \\ &p = \Big(\frac{\alpha b^{\eta} z}{y}\Big)^{\frac{1}{\rho}} \end{aligned} \tag{14}$$

Entrant's Problem

The industry has a constant mass Γ of potential entrants,[30] each of whom receive a signal q about their initial idiosyncratic productivity, with $q \sim W(q)$. Potential entrants observe the state of the economy and their quality signal. Conditional on that information, they decide whether or not to enter. Potential entrants that decide to enter pay an entry cost c and observe the idiosyncratic shock. The distribution of idiosyncratic productivity in the first period of operation is $H_e(s|q)$. Once the potential entrant pays the entry cost and receives its baseline reputation, it behaves like an incumbent. A potential entrant solves the following problem:

$$V^P(b_0,q,z) = \max\Big\{0, -c + \int V^I(b_0,s,z)\, dH_e(s|q)\Big\} \tag{15}$$

I assume that the distribution of signals for the entrants is Pareto with location parameter $\underline{q}$ and scale $\varkappa$. The idiosyncratic shock in the first period of operation follows the process $\log(s_{it}) = \rho_s \log(q_i) + \sigma_s \epsilon_{it}$.

Recursive Equilibrium

Let the distribution of operating firms over the two dimensions of heterogeneity be denoted by $\Omega_t(b,s)$, and let Γ_t denote the vector of aggregate state variables. For a given Γ_0, a recursive equilibrium consists of: (i) value functions $V^I(b,s,z), V^P(b_0,s,z)$, (ii) policy functions $y(b,s,z)$, $p(b,s,z)$, $n(b,s,z)$, and $b'(b,s,z)$, (iii) distribution of operating firms $\{\Omega_t\}_{t=1}^{\infty}$ such that:

1. $V^I(b_0,s,z)$, $y(b,s,z)$, $p(b,s,z)$, $n(b,s,z)$, and $b'(b,s,z)$ solve the incumbent's problem;
2. $V^P(b,s,z)$ solves the entrant's problem;
3. Labor market clears.

[30] The pool of potential entrants is assumed to have a constant mass. A natural extension of this model is to allow to have depletion effects of potential entrants.

5.2 Calibration

I list the parameter values that I use to calibrate the model in Table 16.[31] For consistency with the empirical analysis, I assume that a period is equivalent to one year and that the unit of analysis is the establishment. The time preference parameter, β, is assumed to be 0.95. I calibrate the aggregate demand shock based on the autoregressive properties of my main business cycle indicator: the log deviation from trend of aggregate real GDP between 1978 and 2012.[32] Accordingly, I set ρ_Z to 0.56 which is very close to the value estimated in Foster et al. (2015) and implies a correlation of around 0.80 at quarterly frequencies, which is in the low range of the parametrizations typically used in RBC models (King and Rebelo, 1999). I choose the parameters of the process driving the idiosyncratic productivity shocks externally. I employ estimates from Foster et al. (2008) on the physical total factor productivity process and I set the persistence parameter to 0.814 and the standard deviation to 0.26. These values are also within the range of estimates used in Lee and Mukoyama (2008) and Clementi and Palazzo (2013).

To calibrate the parameters of the demand function, I rely on Foster et al. (2015), which uses a similar specification of the demand structure. The price elasticity of demand is set to 1.6, which is within the range of estimates in Foster et al. (2008). Nevertheless, this elasticity is smaller than that of Ravn et al. (2006) and Hall (2013). For instance, Hall (2013) uses a elasticity of 6 which implies a markup of 1.2. A price elasticity of demand of 1.62, like the one estimated in Foster et al. (2015), implies a markup of around 2.5 if the firm sets prices according to the monopoly markup, i.e. in the case of no reputation capital. Nevertheless, the average markup in the model is about 1.5, which is closer to the reasonable range for the parameters that describe the average markup in the U.S. economy.[33] I also use the depreciation rate of reputation estimated in Foster et al. (2015). This depreciation rate is close to the range of intangible capital depreciation in Corrado et al. (2009), and smaller than the annual rate of depreciation of advertising, which is around 60 percent. Yet, it is within the range of observed customer turnover rates, which vary between 10 and 25 percent, in a broad range of industries as discussed in Gourio and Rudanko (2014).

The parameters governing the distribution of potential entrants, the distribution of operational costs, the exogenous exit shock, the entry cost, and the initial level of customer capital are all chosen to match important moments in the size distribution and survival function. I use a panel of simulated data from the steady state version of the model (by setting z to 1) to compute the model counterparts of the empirical moments. I chose to match the average size, the ratio of average size of entrants to average size of active businesses, the ratio of average size of exiters to active businesses, the entry rate (defined as number of entrants divided by total number of businesses), the survival rate

[31] Details on the numerical approximation are available in Appendix B.

[32] The filtering process is a HP with smooth parameter of 100. I use the estimated parameters after fitting an autoregressive process to the data series. I obtain a coefficient in the lagged variable of 0.56 and standard deviation of 0.015.

[33] The wage rate was normalized to $(\rho - 1)/\rho$. Because labor supply is perfectly elastic at this level, the size of the industry is determined by the constant α in the demand function and by the mass Γ of potential entrants.

until age 5, and average survival rate from ages 21 to 26. The targets are the statistics computed using the Business Dynamics Statistics for the period 1978 and 2012. The entry cost is essential to determining the cutoff quality of entrants. Conditional on the quality of entrants, the exit shock and mean operational cost parameters are very important in determining the shape of the survival function. The initial level of reputation determines the initial size of entrants and their growth as measured by the average size of entrants relative to active businesses. Table 17 presents the values of the simulated moments and their empirical counterparts.

5.3 Model Fit

The model yields strong predictions for the lifecyle size dynamics and survival function. I use these predictions to evaluate the fit resulting from this calibration by comparing the patterns implied by the model with the corresponding moments observed in the U.S. economy. Panel A of Figure 8 plots the log estimated average business size over the lifecycle of businesses in the model and in the data. The plot suggests that the model fits the observed data well. In particular, the model generates a very similar growth pattern in the average size of the cohort as it becomes older. Panel B of Figure 8 plots the survival function implied by the model and that observed in the data. Overall, the figure suggests that the survival function generated by the model closely mimics that of the data.

Prior empirical literature (e.g. Haltiwanger et al. (2015)) documents that firm growth rates decline with age and size. Models of idiosyncratic productivity in the spirit of Hopenhayn (1992) generate a negative relation between size and growth but cannot generate a negative relation between age and growth conditional on size. My model generates this relationship endogenously: small businesses with low idiosyncratic productivity are more likely to receive better productivity draws in the following periods because the process governing idiosyncratic productivities is mean reverting. As a result, smaller businesses grow more quickly than businesses that are already large and highly productive. To the extent that businesses entering the market have a below-average productivity, this mechanism generates an unconditional negative correlation between age and growth.

In addition to the mean-reversion mechanism, the model also features an additional state variable, customer capital, that also generates a negative relationship between age and growth. Two businesses with the same productivity level may grow at different rates because of differences in their accumulated stock of customer capital. Businesses with a high level of customer capital grow less quickly because they are closer to their steady-state level and do not invest as much in accumulating future demand. Thus, this feature of the model also creates a direct negative relation between age and growth conditional on size Because of the above features, my model does not require younger businesses to have lower average productivity in order to generate a negative correlation between size and growth. In fact, the calibrated model generates a minor difference between the average productivity of entrants and incumbent businesses , which is consistent with the empirical evidence in Foster et al. (2015). To my knowledge, this is the first model of firm dynamics with this feature.

The model also features both exogenous and endogenous exit. A business endogenously decides to exit if its continuation value, which is increasing in its stock of customer capital and on its level of idiosyncratic productivity, is greater than the operational cost that the business incurs to produce. This operational fixed cost is invariant to size and the outside option of each entrepreneur are also held constant. Hence, endogenous exit decreases with size and age because older and larger businesses with higher stocks of customer capital are less likely to exit. I also include an exogenous shock such that the model generates exit among older businesses. As illustrated in Panel B of Figure 8 these two elements of the model combine to generate a survival function that closely matches the observed survival function in the data.

Overall, the results of this section suggest that the model delivers a very good fit of the empirical patterns of firm growth, entry, and exit. This will be important in the later sections, as I construct counterfactuals to evaluate the role of selection mechanisms and demand-side channels in explaining the persistent effect of initial business cycle conditions.

5.4 Selection and Persistence of Initial Economic Conditions

Composition of Cohorts

The productivity distribution of entrants varies with the state of the economy. Potential Entrants follow a cut-off strategy and decide to enter if their signal q is above a threshold $q^\star$, which varies countercyclically. The value of starting a business is given by $\int V^I(b_0, s, z)\, dH_e(s|q)$ which is a strictly increasing function of the signal q. This relation follows from the fact that $V^I(.)$ is weakly increasing in the idiosyncratic productivity shock s and from the fact that the distribution of idiosyncratic shocks is stochastically increasing in the value of the signal. Figure 6 plots $V^I(.)$ as a function of q for two alternative values of the aggregate shock. The value of the cutoff $q^\star$ is countercyclical because $\int V^I(b_0, s, z)\, dH_e(s|q)$ is increasing in the aggregate shock z and the entry cost and initial customer capital are constant. These features imply that when the economy receives a positive aggregate shock, more businesses decide to enter but their average productivity is lower. The histogram of idiosyncratic productivities at time t=1 plotted in Figure 7 illustrates this empirical pattern: the mass of entrants is procyclical but the average productivity is countercyclical. This is also consistent with the empirical facts documented in section 3.

The survival patterns of new businesses also vary with the state of the economy at inception. Expansionary startups produce more in their first years of activity and accumulate customer capital at a faster rate. As a result, the continuation value of expansionary startups grows faster, and they are therefore less likely to exit in response to a negative idiosyncratic shock. On the other hand, differences in the productivity composition of businesses increase the exit rate of expansionary cohorts. As Figure 7 illustrates, there are relatively more low-productivity startups during good economic periods. Because these less productive businesses are more likely to exit, differences in composition

increase the exit rate of cohorts starting during favorable aggregate demand periods. The plot presented in Figure 9 shows that these two forces offset each other and result in a simulated survival function that does not vary substantially with differences in initial economic conditions.

The differences in the idiosyncratic productivity distribution of each cohort cohort subside as they grow older, regardless of how different these cohorts were initially. There are two forces that explain the convergence in the productivity distribution over time. First, the exit patterns documented above suggest that the differences in composition tend to disappear over time: the disproportionately large share of low-productivity businesses entering in economic booms decreases over time as these low-productivity businesses exit. Second, the idiosyncratic productivity process is mean-reverting, which also contributes to the dissipation of initial differences in the distribution of idiosyncratic productivities.

Differences in Size among Cohorts

The simulated data obtained from this theoretical framework mimics the main empirical fact documented in this paper. Figure 10 clearly shows that the model generates a lifecycle dynamics in which recessionary startups are smaller than expansionary ones, and, most importantly, these differences persist long after these businesses enter the market.

I also document in section 4 that average business size at size is procyclical. The theoretical framework does not necessarily generate this feature of the data. Two countervailing forces affect the relationship between entry size and the business cycle in the model. On one hand, the optimal entry size is increasing in the level of the aggregate demand shock. On the other hand, the cutoff signal above which potential entrants start operations is countercyclical. Therefore, the average entry size may be procyclical or countercyclical depending on the sensitivity of the initial distribution of productivities to business cycles. As Figure 10 clearly illustrates, the model does mimic this feature of the data: the initial average size of expansionary cohorts is larger than that of recessionary cohorts.[34] This result suggests that, given the calibration described above, the selection mechanisms are not strong enough to overcome the impact of aggregate demand shocks on average entry size.

Next, I use the theoretical framework to create a counterfactual scenario that allows me to quantify the role of selection at entry and demand-side factors in explaining the size differences between recessionary and expansionary cohorts. Figure 7 shows that the distribution of idiosyncratic productivities at entry depends on the initial aggregate demand conditions. I consider a counterfactual scenario in which I isolate the effect of customer capital on firm dynamics by holding constant the productivity composition of entrants across cohorts. This exercise compares the average size over time of two cohorts with the same productivity distribution at entry, expect that the counterfactual cohort begins operating during a positive demand shock period with the productivity distribution of a recessionary cohort and the recessionary cohort begins operating during a negative aggregate demand shock period.

[34] This is at odds with Clementi and Palazzo (2013), whose model generates that entrants are larger during recessions.

Figure 10 presents the lifecycle evolution of the counterfactual cohorts.[35] The plot uncovers two main insights. First, the wedge between the initial sizes of the counterfactual and recessionary cohorts is more than two times larger than that observed in the data. Second, as can be clearly seen from Figure 11, the average size difference between the counterfactual and recessionary cohort decays over time whereas the observed average size difference between expansionary and recessionary cohorts remains constant. In fact, the initial difference between the average sizes of the counterfactual and recessionary cohorts is 6.2 percent, but it declines to 4.1 percent after 10 years and to approximately 2 percent after more than 30 years. This decay contrasts with the observed difference between cohorts, which is constant at around four percent over time.

The above exercise sheds light on the economic forces that create persistence in the average size differences of cohorts that start during different stages of the business cycle. First, it shows that endogenous entry and exit strengthen the persistence of the entry-condition effects because the productivity distribution of each cohort becomes more similar over time. The average business size of expansionary cohorts increases as the smaller, less productive businesses, which would not have entered during downturns exit or become relatively more productive. This process slows the rate of size convergence between recessionary and expansionary cohorts. The second force that the model uncovers is that in the absence of differences in the productivity distributions across cohorts, the average business size of recessionary cohorts converges with that of other cohorts, albeit slowly. Businesses that start during recessions are constrained in their ability to rapidly accumulate demand, which hinders their ability to catch up to other businesses. Yet as they approach their steady-state level of customer capital the differences in size across cohorts tend to disappear.

5.5 Aggregate Implications of Persistent Entry Conditions

This subsection analyzes the aggregate implications of the persistent effects of entry conditions documented above. I simulate an aggregate demand shock and analyze its propagation over time through its effect on startup businesses. In the short run, the effect of the shock on the size and employment level of startups accounts for a relatively small share of the shock's total effect in the economy. Over the long run, however, this effect accounts for a much larger share of the total economic response because the effects of initial conditions on startup businesses are larger and persist for longer than the effects on other businesses. I also use the model to evaluate the consequences of the 2008–2009 recession on the size and composition of startup cohorts during this period. I conjecture that the 2008 and 2009 cohorts will employ approximately less 1.5 million workers than cohorts entering under regular economic conditions.

Amplification and persistence

[35] I apply to the simulated data, the specification 2.

I study the amplification and propagation mechanisms in my model by computing impulse responses of industry output to a positive aggregate demand shock at time t = 1. More specifically, I shock the level of aggregate demand by two standard deviations of the corresponding distribution. Figure 12 plots the evolution of this economy over 30 periods after repeating this exercise 500 times. The figure suggests that there is a substantial amplification of the effects of the shock: aggregate output in the year of the shock increases by approximately seven percent, which is twice as large as the induced increase in aggregate demand. In addition, the increase in industrial output is substantially more persistent than would be implied by the aggregate demand process. The implied autocorrelation of aggregate output is 22 percent higher than the implied autocorrelation of the aggregate demand shock z.[36] These results suggest that the process of demand accumulation and shifts in the composition of active businesses are powerful determinants of internal shock propagation. In fact, these mechanisms may be critically important for understanding the low-frequency fluctuations of aggregate output.

To isolate the component of the shock's aggregate effect that is attributable to its impact on startup businesses, I compare the impulse response described above with the impulse response of an alternative economy in which all cohorts except that entering at time t=1 are impacted by the same aggregate demand shock.[37] The only difference between these two scenarios is that, in this alternative scenario, new entrants do not benefit from the positive aggregate demand shock. As a result, fewer businesses start operations, while those that do start on a smaller scale. This effect naturally affects the output in the year of the shock and in the years thereafter. Figure 12 plots the (log) impulse response of output under both scenarios.

The decisions of incumbents are similar in both scenarios.[38] Thus, the difference between the two plots in Figure 12 reflects differences in the entry decisions and post-entry dynamics of potential entrants. In the alternative scenario, fewer businesses enter at time t=1, and the average size of businesses that do enter is smaller. The differences in total output between these two scenarios is sizable because the entrants' investment decisions are more cyclical than those of incumbent firms due to their lower level of customer capital relative to the industry average. The difference between the lines plotted in Figure 12 suggests that the entering cohort accounts for a significant portion, approximately 7.5 percent, of the response of the economy to a positive aggregate demand shock. Moreover, the difference between the impulse-response functions of the baseline and alternative economies does not decrease significantly over time. As a result, the impact of the aggregate demand shock on entering firms accounts for an even larger share of the total impact of the aggregate demand shock over time. After 30 years, total output is one percent greater than the unconditional mean output, and approximately half of this effect is attributable to the impact of the aggregate demand

[36] This means that if I internally calibrate the model to generate the aggregate fluctuations of the aggregate output, then I could use a autocorrelation parameter for z below the level currently used. Results with alternative calibration for ρ_z and σ_z are available upon request.

[37] I.e. $\ln(z_t|\text{entrant}) = 0$, and $\ln(z_t|\text{incumbent}) = 0.037$. Note that this only affects the cohort that was born in $t = 1$, all cohorts born after that receive the exact same shocks in both scenarios.

[38] This results, partially, from the lack of general equilibrium effects and the type of labor supply faced by this industry. In future work, I plan to extend my model to a general equilibrium setting.

shock on the entering cohort. The results of this exercise are consistent with the hypothesis that temporary shocks that affect firms started during those periods have long-lasting implications for the economy as a whole.

Long-run implications of the 2008–2009 Recession

The number of startups saw an unprecedented decline during the 2008–2009 recession. During this period the number of new establishments declined by about 22 percent relative to the 2004–2006 period. In addition, the businesses started during the recession employed 23 percent fewer workers in their first year of activity than their pre-recession counterparts. As a result, startup's share of total employment declined from a pre-crisis level of 5.6 percent to 4.2 percent.[39]

Because the Great Recession ended only a few years ago, I am unable to conduct an empirical analysis of the long-term employment losses caused by the recession's effect on startups. Nevertheless, the model allows me to simulate the propagation effects of the 2008–2009 that are attributable to the mechanisms of persistence documented in this paper. First, I choose the size of the aggregate demand shock to mimic the decline in the number of entrants that was observed during 2008–2009.[40] Next, I simulate the lifecycle path of the cohort of businesses that started operating in that period and repeat the experiment 500 times. Figure 13 plots the estimated effect of aggregate demand shock on the number of firms that compose that cohort, their average size, and their total employment over the next 30 years.

Overall the plots of Figure 13 suggest that the number of businesses in the Great Recession cohort, their average size, and their employment will remain on a relatively low plateau over time. In particular, the total employment of this cohort relative to non-recessionary cohorts is approximately 25 percent lower in the first year of activity and remains 20 to 25 percent lower over time because of the persistent effects of entry conditions on the average size and number of businesses. Figure 14 plots the propagation of the shock on the total employment in the economy. As can be clearly seen from the figure, the model suggests that the Great Recession implied a loss of 1 million startup jobs in the short run and a permanent loss of approximately 600,000 thousand jobs in the economy.

6 Conclusion

This paper has explored the role of business cycle conditions at the time businesses begin their operations on the evolution of their average size and productivity. Using a sample that comprises the entire universe of establishments in the U.S., I find that business cycle conditions at inception play a strong role in determining the evolution of size and productivity throughout the lifecycle. Businesses born during economic downturns start smaller and stay smaller than businesses born under more favorable conditions.

[39] These statistics are computed using Business Dynamics Statistics.

[40] It corresponds to 2.5 standard deviations of the aggregate shock distribution.

I evaluate the economic mechanisms that mediate the above relationship. The first is selection at entry. Businesses born during economic upswings might begin larger and remain larger because they are more productive than businesses born during recessions. The empirical findings suggest otherwise: businesses born in economic downturns are more productive and more likely to belong to industries that require high entrepreneurial skill. Therefore, selection at entry cannot explain the procyclicality of business size at entry. Next, I evaluate the role of capital adjustment costs and demand-side factors in explaining the persistent differences in average size throughout the businesses' lifecycle. The empirical findings suggest that the average initial size differences among cohorts are not more persistent in industries with high adjustment costs. On the other hand, I find that these size differences are more persistent in industries in which the demand accumulation process, such as building a customer base or developing a reputation, plays a larger role.

I interpret these empirical findings using the analytical framework that I develop in the paper. A positive shock to aggregate demand allows potential entrants to make a larger initial investment. At the same time, the positive aggregate demand shock also encourages less-productive and smaller-scale entrepreneurs to enter the market, thereby mitigating the average size difference between recessionary and expansionary cohorts. The difference in average size across cohorts persists because of two mutually reinforcing mechanisms. First, expansionary startups take advantage of initially high aggregate demand to accumulate a customer base and develop a reputation more rapidly, thereby expanding their future demand. On the other hand, the small, low-productivity businesses started during economic expansions gradually exit the market as economic conditions deteriorate. The exit of these low-productivity businesses increases the average business size of expansionary cohorts, which reinforces the initial size differences across cohorts. Overall, these two mechanisms seem to explain the very slow convergence of average business size among expansionary and recessionary cohorts.

I use the model to quantitatively evaluate the role of the persistent effects of entry conditions in the propagation of the Great Recession. The results of this analysis reinforce the policy relevance of the empirical fact that I document in this paper. I find that the persistently smaller size of businesses started during the Great Recession accounts for a permanent one percentage point decrease in the level of aggregate employment.

References

Adelino, M., Schoar, A., Severino, F., 2013. House prices, collateral and self-employment. Tech. rep., National Bureau of Economic Research.

Baker, G., Gibbs, M., Holmstrom, B., 1994. The wage policy of a firm. The Quarterly Journal of Economics 109 (4), 921–955.

Bernanke, B., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. american economic review 79 (1), 14–31.

Bloom, N., Sadun, R., Van Reenen, J., 2010. Does product market competition lead firms to decentralize? The American Economic Review, 434–438.

Boeker, W., 1989. Strategic change: The effects of founding and history. Academy of Management Journal 32 (3), 489–515.

Caballero, R. J., 1999. Aggregate investment. Handbook of macroeconomics 1, 813–862.

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

Clementi, G. L., Palazzo, B., 2013. Entry, exit, firm dynamics, and aggregate fluctuations. Tech. rep., National Bureau of Economic Research.

Cogley, T., Nason, J. M., 1995. Output dynamics in real-business-cycle models. The American Economic Review, 492–511.

Cooper, A. C., Gimeno-Gascon, F. J., Woo, C. Y., 1994. Initial human and financial capital as predictors of new venture performance. Journal of business venturing 9 (5), 371–395.

Cooper, R. W., Haltiwanger, J. C., 2006. On the nature of capital adjustment costs. The Review of Economic Studies 73 (3), 611–633.

Corrado, C., Hulten, C., Sichel, D., 2009. Intangible capital and us economic growth. Review of income and wealth 55 (3), 661–685.

Davis, S. J., Haltiwanger, J., 2014. Labor market fluidity and economic performance. Tech. rep., National Bureau of Economic Research.

Davis, S. J., Haltiwanger, J., Jarmin, R., Miranda, J., 2007. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. In: NBER Macroeconomics Annual 2006, Volume 21. MIT Press, pp. 107–180.

Deaton, A., 1997. The analysis of household surveys: a microeconometric approach to development policy. World Bank Publications.

Decker, R., Haltiwanger, J., Jarmin, R., Miranda, J., 2014. The role of entrepreneurship in us job creation and economic dynamism. The Journal of Economic Perspectives, 3–24.

Fort, T. C., Haltiwanger, J., Jarmin, R. S., Miranda, J., 2013. How firms respond to business cycles: The role of firm age and firm size. IMF Economic Review 61 (3), 520–559.

Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review 98 (1), 394–425.

Foster, L., Haltiwanger, J. C., Krizan, C. J., 2001. Aggregate productivity growth. lessons from microeconomic evidence. In: New developments in productivity analysis. University of Chicago Press, pp. 303–372.

Foster, L., Haltiwanger, J. C., Syverson, C., 2015. The slow growth of new plants: Learning about demand. Economica, *forthcoming*.

Ghatak, M., Morelli, M., Sjöström, T., 2007. Entrepreneurial talent, occupational choice, and trickle up policies. Journal of Economic Theory 137 (1), 27–48.

Gollop, F. M., Monahan, J. L., 1991. A generalized index of diversification: trends in us manufacturing. The Review of Economics and statistics, 318–330.

Gourio, F., Rudanko, L., 2014. Customer capital. The Review of Economic Studies, rdu007.

Hall, R., 2013. What the cyclical response of advertising reveals about markups and other macroeconomic wedges? Tech. rep., Hoover Institution Stanford University.

Hall, R. E., 2004. Measuring factor adjustment costs. The Quarterly Journal of Economics, 899–927.

Haltiwanger, J., Jarmin, R., Kulick, R., Miranda, J., 2015. High growth young firms: Contribution to job growth, revenue growth and productivity. In: Measuring Entrepreneurial Businesses: Current Knowledge and Challenges. University of Chicago Press.

Haltiwanger, J., Jarmin, R. S., Miranda, J., 2013. Who creates jobs? small versus large versus young. Review of Economics and Statistics 95 (2), 347–361.

Heckman, J., Robb, R., 1985. Using longitudinal data to estimate age, period and cohort effects in earnings equations. In: Cohort analysis in social research. Springer, pp. 137–150.

Hopenhayn, H. A., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica: Journal of the Econometric Society, 1127–1150.

Jarmin, R. S., Miranda, J., 2002. The longitudinal business database. Center for Economic Studies, US Census Bureau.

Kahn, L. B., 2010. The long-term labor market consequences of graduating from college in a bad economy. Labour Economics 17 (2), 303–316.

King, R. G., Rebelo, S. T., 1999. Resuscitating real business cycles. Handbook of macroeconomics 1, 927–1007.

Lee, Y., Mukoyama, T., 2008. Entry, exit and plant-level dynamics over the business cycle.

Lee, Y., Toshihiko, M., 2015. Entry and exit of manufacturing plants over the business cycle. European Economic Review 77, 20 – 27.

Luttmer, E. G. J., 2006. Consumer search and firm growth. Federal Reserve Bank of Minneapolis, Research Department.

Mian, A. R., Sufi, A., 2012. What explains high unemployment? the aggregate demand channel. Tech. rep., National Bureau of Economic Research.

Moreira, S., 2014. Self-employment and the business cycle. Tech. rep.

Moscarini, G., Postel-Vinay, F., 2012. The contribution of large and small employers to job creation in times of high and low unemployment. The American Economic Review 102 (6), 2509–2539.

Oreopoulos, P., von Wachter, T., Heisz, A., 2012. The short-and long-term career effects of graduating in a recession. American Economic Journal: Applied Economics 4 (1), 1–29.

Ouimet, P., Zarutskie, R., 2014. Who works for startups? the relation between firm age, employee age, and growth. Journal of financial Economics 112 (3), 386–407.

Perla, J., 2015. Product awareness, industry life cycles, and aggregate profits. Working paper.

Rampini, A. A., 2004. Entrepreneurial activity, risk, and the business cycle. Journal of Monetary Economics 51 (3), 555–573.

Ravn, M., Schmitt-Grohé, S., Uribe, M., 2006. Deep habits. The Review of Economic Studies 73 (1), 195–218.

Ravn, M. O., Uhlig, H., 2002. On adjusting the hodrick-prescott filter for the frequency of observations. Review of economics and statistics 84 (2), 371–376.

Saffie, F. E., Ates, S. T., 2013. Fewer but better: Sudden stops, firm entry, and financial selection.

Schoar, A., Zuo, L., 2013. Shaped by booms and busts: How the economy impacts ceo careers and management styles.

Sedlácek, P., Sterk, V., 2014. The growth potential of startups over the business cycle.

Siemer, M., 2014. Firm entry and employment dynamics in the great recession.

Syverson, C., 2011. What determines productivity? Journal of Economic Literature 49 (2), 326–365.

Table 1: Summary Statistics

The table reports summary statistics for the samples used in the empirical analysis. Panel A presents the statistics for the main samples used throughout the paper. Sample 1 is a balanced sample that includes all establishments born between 1978 and 2001 and their level of employment for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Sample 2 includes establishments with information on revenue and employment born between 1994 and 2007 (excluding 2002) and their outcomes for up to 5 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Panel B presents the statistics for the samples used in the robustness checks exercises. Sample 1.2 is an unbalanced sample that differs from Sample 1 by including observations for up to 34 years. This means that the oldest cohort has establishments of up to 34 years old and the youngest cohort has observations until 10 years old. Sample 1.3 is an unbalanced sample that differs from Sample 1 by including cohorts from 1978 to 2010 and observations for up to 34 years. In this sample the oldest cohort has establishments of up to 34 years old and the youngest cohort has observations in their first year of activity. Sample 1.1 is a balanced sample that differs from Sample 1 because it only includes establishments that survived at least 10 years. Both panels reports the mean, standard deviation (defined within a cell in the intersection between age, period, cohort, state, industry at the four digit level, multi-unit status of the establishment, and legal type of ownership), number of establishments (in thousands), and number of establishment-years (in thousands) for age and the dependent variables. *Age* is the number of years since the establishment was born. *Ln (Emp)* is the natural logarithm of the number of employees at the establishment-level. *Ln (Rev)* is the natural logarithm of the value of shipments, sales, receipts, or revenue (in thousands). *Ln (LP)* is the measure of labor productivity defined as revenue per employee (in thousands). Source: LBD and author's calculations.

Panel A — Main samples

	SAMPLE 1		SAMPLE 2			
	Cohorts 1978-2001		Cohorts 1994-2007			
	Period 1979-2011, Ages 1-10		Period 1995-2011, Ages 1-5			
	Age	Ln(Emp)	Age	Ln(Emp)	Ln(Rev)	Ln(LP)
Mean	4.638	1.518	2.711	1.188	-1.0765	-2.2644
St. Dev.	2.839	0.766	1.399	0.563	0.749	0.629
n (in '000s)	12,800	12,800	3,808	3,808	3,808	3,808
N (in '000s)	78,300	78,300	14,500	14,500	14,500	14,500

Panel B — Alternative samples

	SAMPLE 1.1		SAMPLE 1.2		SAMPLE 1.3	
	Cohorts 1978-2001		Cohorts 1978-2001		Cohorts 1978-2010	
	Period 1978-2011, Ages 1-10		Period 1979-2011, Ages 1-34		Period 1978-2011, Ages 1-34	
	Age	Ln(Emp)	Age	Ln(Emp)	Age	Ln(Emp)
Mean	5.477	1.635	8.554	1.608	7.820	1.577
St. Dev.	2.851	0.815	6.788	0.809	6.611	0.807
n (in '000s)	5,245	5,245	12,800	12,800	17,600	17,600
N (in '000s)	52,900	134,000	115,000	115,000	134,000	52,900

Table 2: Descriptive Statistics of Business Cycle Indicators (1978–2012)

The table reports summary statistics of the business cycle indicators used in the main analysis. Panel A reports summary statistics of business cycle indicators aggregated at the national level. *Ln(GDP real detrended)*$_t$ is the Hodrick and Prescott (HP) filtered natural logarithm of the annual aggregate real gross domestic product. *Ln(GDP nom. detrended)*$_t$ is the HP-filtered natural logarithm of the annual aggregate nominal gross domestic product. *Ln(NBHEPU detrended)*$_t$ is the HP-filtered natural logarithm of the newspaper-based index of economic policy uncertainty (NBHEPU) averaged over the 12 months of each year. The uncertainty index was obtained from `http://policyuncertainty.com/us_historical.html`. *NFCI*$_t$ is the Federal Reserve's National Financial Condition Index, which measures the risk, liquidity, and leverage in money, debt, and equity markets. The index is averaged over the quarters of the year without seasonal adjustment. Δ*ln(GDP Real)*$_t$ is the demeaned difference in logarithms of the annual aggregate real gross domestic product. Δ*ln(GDP Nom.)*$_t$ is the demeaned difference in logarithms of the annual aggregate nominal gross domestic product. Δ*ln(Emp.)*$_t$ is the demeaned difference in logarithms of the civilian non-institutional employed individuals aged 14 and over obtained from the BLS. Δ*ln(PIpc)*$_t$ is the demeaned difference in logarithms of the personal income per capita in dollars obtained from the BEA. Δ*ln(Unemp. Rate)*$_t$ is the demeaned difference in logarithms of the annual unemployment rates (%) from the BLS-LAUS. Panel B of the table reports summary statistics of business cycle indicators aggregated at the industry level. *Ln(GDP nom. detrended)*$_{it}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the annual nominal value added by Industry obtained from the BEA Industry Accounts and treated for the concordance between SIC-based measurements and NAICS-based measurements. *ln(Emp. detrended)*$_{it}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the total number people engaged in production (in thousands) from the BEA industry accounts. *ln(Demand detrended)*$_{it}$ is the Hodrick and Prescott (HP) filtered natural logarithm of a proxy for the demand shocks of an industry. The proxy is the demand shock in year t across all downstream buyers of industry i, weighted by the downstream industry's share of the industry i total sales. Δ*ln(GDP Nom.)*$_{it}$ is the demeaned difference in logarithms of the annual nominal value added by Industry obtained from the BEA Industry Accounts and treated for the concordance between SIC-based measurements and NAICS-based measurements. Δ*ln(Emp.)*$_{it}$ is the demeaned difference in logarithms of the total number people engaged in production (in thousands) from the BEA industry accounts. Panel C of Table 2 reports summary statistics of business cycle indicators aggregated at the regional level. *Ln(GDP nom. detrended)*$_{st}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the annual aggregate nominal gross domestic product by state obtained from the BEA's regional economic accounts. *ln(Emp. detrended)*$_{st}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the civilian non-institutional employed individuals aged 14 and over by state obtained from the BLS. *ln(PIpc detrended)*$_{st}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the personal income per capita by state in dollars obtained from the BEA. Δ*ln(GDP Nom.)*$_{st}$ is the demeaned difference in logarithms of the annual aggregate nominal gross domestic product by state obtained from the BEA's regional economic accounts.. Δ*ln(Emp.)*$_{st}$ is the demeaned difference in logarithms of the civilian non-institutional employed individuals aged 14 and over by state obtained from the BLS. Δ*ln(PIpc)*$_{st}$ is the demeaned difference in logarithms of the personal income per capita by state in dollars obtained from the BEA. Δ*ln(Emp.)*$_{cnty,t}$ is the demeaned difference in logarithms of the number of employees at the county-level calculated by summing up the number of employees by county in the LBD database. Δ*ln(PIpc)*$_{cnty,t}$ is the demeaned difference in logarithms of the personal income per capita by county in dollars obtained from the BEA. Δ*ln(Emp.)*$_{cbsa,t}$ is the demeaned difference in logarithms of the number of employees at the CBSA-level calculated by summing up the number of employees by county in the LBD database. Δ*ln(PIpc)*$_{cbsa,t}$ is the demeaned difference in logarithms of the personal income per capita by CBSA in dollars obtained from the BEA.

Panel A — National Indicators

	Observations		Mean*	Std. Dev.	Correlations with	
	N	n			Ln(GDP real detrended)	Δln(GDP Real)$_t$
Ln(GDP real detrended)$_t$	35	1	0.000	0.013	1.000	0.517
Ln(GDP nom. detrended)$_t$	35	1	0.000	0.013	0.860	0.325
Ln(NBHEPU detrended)$_t$	35	1	0.002	0.087	-0.232	-0.250
NFCI$_t$	35	1	-0.033	0.849	-0.039	-0.414
Δln(GDP Real)$_t$	35	1	0.000	0.020	0.517	1.000
Δln(GDP Nom.)$_t$	35	1	0.000	0.028	0.515	0.668
Δln(Emp.)$_t$	35	1	0.000	0.015	0.689	0.861
Δln(PIpc)$_t$	35	1	0.000	0.028	0.540	0.493
Δln(Unemp. Rate)$_t$	35	1	0.000	0.145	-0.568	-0.887

Panel B — Industry Indicators

	Observations		Mean	Std. Dev.	Correlations with	
	N	n			Ln(GDP real detrended)	Δln(GDP Real)$_t$
Ln(GDP nom. detrended)$_{it}$	2,065	59	0.001	0.053	0.285	0.128
ln(Emp. detrended)$_{it}$	2,065	59	0.001	0.024	0.555	0.082
ln(Demand detrended)$_{it}$	2,065	59	0.001	0.027	0.543	0.236
Δln(GDP Nom.)$_{it}$	2,065	59	0.000	0.087	0.188	0.268
Δln(Emp.)$_{it}$	2,065	59	0.000	0.040	0.440	0.474

Panel C — Regional Indicators

	Observations		Mean	Std. Dev.	Correlations with	
	N	n			Ln(GDP real detrended)	Δln(GDP Real)$_t$
State						
Ln(GDP nom. detrended)$_{st}$	1,785	51	0.001	0.021	0.429	0.155
ln(Emp. detrended)$_{st}$	1,785	51	0.000	0.011	0.688	0.191
ln(PIpc detrended)$_{st}$	1,785	51	0.000	0.016	0.434	0.001
Δln(GDP Nom.)$_{st}$	1,785	51	0.000	0.041	0.322	0.381
Δln(Emp.)$_{st}$	1,785	51	0.000	0.018	0.462	0.621
Δln(PIpc)$_{st}$	1,785	51	0.000	0.032	0.452	0.380
County						
Δln(Emp.)$_{cnty,t}$	109,806	3,163	0.000	0.113	0.140	0.158
Δln(PIpc)$_{cnty,t}$	108,776	3,138	0.000	0.063	0.226	0.175
CBSA						
Δln(Emp.)$_{cbsa,t}$	32,865	939	0.000	0.071	0.224	0.233
Δln(PIpc)$_{cbsa,t}$	37,795	1,080	0.000	0.037	0.358	0.293

Table 3: Size and the Business Cycle at Entry: Average and Age-Specific Elasticities

The table reports the coefficients of OLS regressions. The dependent variable in columns 1–4 is the natural logarithm of the number of employees at the establishment-level. The dependent variable in column 5 is defined as the number of employees at the establishment-level in year t minus the number of employees at the establishment level in year $t-1$ divided by the average number of employees of the establishment in years t and $t-1$. Age is the number of years since the establishment was born. $\mathbf{1}_{\{\text{age}=i\}}$ is an indicator variable that takes the value of one if the establishment is i years of age. $1_{\{\text{Ln(GDP real detrended)}>0\}}$ is an indicator variable that takes the value of one if Ln(GDP real detrended) in the entry year is greater than zero. Other controls include age fixed-effects, year fixed-effects, state fixed-effects, industry fixed-effects, ownership type fixed-effects, and multi-Unit fixed-effect. The remaining variables are described in Table 2. The sample used in this table includes all establishments born between 1978 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	Ln(Emp)				Δ%Emp
Ln(GDP real detrended)	1.166***	1.016***			0.040***
	(0.099)	(0.103)			(0.013)
age×Ln(GDP real detrended)		0.031***			
		(0.010)			
$\mathbf{1}_{\{\text{age}=1\}}$×Ln(GDP real detrended)			0.964***		
			(0.127)		
$\mathbf{1}_{\{\text{age}=2\}}$×Ln(GDP real detrended)			1.078***		
			(0.101)		
$\mathbf{1}_{\{\text{age}=3\}}$×Ln(GDP real detrended)			1.063***		
			(0.093)		
$\mathbf{1}_{\{\text{age}=4\}}$×Ln(GDP real detrended)			1.336***		
			(0.100)		
$\mathbf{1}_{\{\text{age}=5\}}$×Ln(GDP real detrended)			1.195***		
			(0.099)		
$\mathbf{1}_{\{\text{age}=6\}}$×Ln(GDP real detrended)			1.182***		
			(0.112)		
$\mathbf{1}_{\{\text{age}=7\}}$×Ln(GDP real detrended)			1.181***		
			(0.126)		
$\mathbf{1}_{\{\text{age}=8\}}$×Ln(GDP real detrended)			1.280***		
			(0.113)		
$\mathbf{1}_{\{\text{age}=9\}}$×Ln(GDP real detrended)			1.281***		
			(0.114)		
$\mathbf{1}_{\{\text{age}=10\}}$×Ln(GDP real detrended)			1.268***		
			(0.105)		
$\mathbf{1}_{\{\text{Ln(GDP real detrended)}>0\}}$				0.021***	
				(0.002)	
R-squared	0.603	0.603	0.603	0.604	0.041
Age Fixed-Effects?	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes	Yes
Multi-Unit Fixed-Effect?	Yes	Yes	Yes	Yes	Yes
Sample	1	1	1	1	1

Table 4: Size and the Business Cycle at Entry: Alternative Indicators

The table reports the coefficients of OLS regressions. The results presented in columns (1)–(9) replicate the specification of column (1) of Table 3 using alternative indicators of business cycle conditions. The dependent variable in columns 1–10 is *Ln(Emp)* which is defined as the natural logarithm of the number of employees at the establishment-level. *Beta* is the coefficient of interest using the alternative business cycle indicators in columns (1) through (10). The alternative business cycle indicators are described in Table 2. Other controls in columns 1–9 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample used in this table includes all establishments born between 1978 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ln(GDP real detrended)	Ln(GDP nom. detrended)	Ln(NBHEPU detrended)	NFCI	Δln(GDP Real)	Δln(GDP Nom.)	Δln(Emp.)	Δln(PIpc)	Δln(Unemp. Rate)
β	1.166***	1.578***	0.010	-0.012***	0.654***	0.750***	0.734***	0.221***	-0.076***
	(0.099)	(0.069)	(0.007)	(0.002)	(0.063)	(0.045)	(0.109)	(0.044)	(0.010)
R-squared	0.603	0.604	0.603	0.603	0.603	0.603	0.603	0.603	0.603
Age Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	1	1	1	1	1	1	1	1	1

Table 5: Size and the Business Cycle at Entry: Entry versus Subsequent Years

The table reports the coefficients of OLS regressions. The dependent variable in columns 1–4 is *Ln(Emp)* which is defined as the natural logarithm of the number of employees at the establishment-level. *Ln(GDP real detrended)*$_{c+i}$ is the Hodrick and Prescott (HP) filtered natural logarithm of the annual aggregate real gross domestic product i years after the cohort was born. Other controls in columns 1–4 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample used in this table includes all establishments born between 1978 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)$_{c}$		1.805***		1.744***
		(0.095)		(0.105)
Ln(GDP real detrended)$_{c+1}$		-0.680***		-0.762***
		(0.095)		(0.093)
Ln(GDP real detrended)$_{c+2}$		0.614***		0.492***
		(0.064)		(0.068)
Ln(GDP real detrended)$_{c+3}$		0.831***	0.193***	0.669***
		(0.047)	(0.053)	(0.046)
Ln(GDP real detrended)$_{c+4}$		0.017		
		(0.080)		
Ln(GDP real detrended)$_{c+5}$	0.104	0.488***		
	(0.106)	(0.107)		
R-squared	0.603	0.604	0.603	0.604
Age Fixed-Effects?	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	Yes	Yes	Yes	Yes
Sample	1	1	1	1

Table 6: Correlation between Estimated Non-Parametric Cohort Effects on Size and Indicators of Business Cycle at Entry

The table reports correlations between estimated non-parametric cohort effects, obtained from the estimation of specification 4, and a set of business cycle indicators, which are defined in the captions of Table 2.

	Correlations		Correlations
Ln(GDP real detrended)	0.52	Δln(GDP Real)$_t$	0.39
Ln(GDP nom. detrended)	0.66	Δln(GDP Nom.)$_t$	0.35
Ln(NBHEPU detrended)	-0.27	Δln(Emp.)	0.23
NFCI	-0.04	Δln(PIpc)	0.20
		Δln(Unemp. Rate)	-0.29

Table 7: Size and the Local Business Cycle at Entry

The table reports the coefficients of OLS regressions. The results presented in columns 1–8 replicate the specification of column 1 of Table 3 using local business cycle indicators measured at the national, state, county, and CBSA levels. The dependent variable in columns 1–8 is *Ln(Emp)* which is defined as the natural logarithm of the number of employees at the establishment-level. *Beta* is the coefficient of interest using the alternative local business cycle indicators in columns 1 through 8. The local business cycle indicators are defined in Table 2. Other controls in columns 1–8 include Age Fixed-Effects, Year Fixed-Effects, Location Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample used in this table includes all establishments born between 1994 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at the state-level in columns 1–4, at the county-level in columns 5 and 6, and at the CBSA-level in columns 7 and 8. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(Emp)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta\ln(\text{Emp.})_t$	$\Delta\ln(\text{PIpc})_t$	$\Delta\ln(\text{Emp.})_{st}$	$\Delta\ln(\text{PIpc})_{st}$	$\Delta\ln(\text{Emp.})_{cnty,t}$	$\Delta\ln(\text{PIpc})_{cnty,t}$	$\Delta\ln(\text{Emp.})_{cbsa,t}$	$\Delta\ln(\text{PIpc})_{cbsa,t}$
β	0.221***	0.734***	0.287***	0.478***	0.167***	0.124***	0.222***	0.136***
	(0.044)	(0.109)	(0.080)	(0.155)	(0.022)	(0.010)	(0.041)	(0.019)
Age Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type of location Fixed-Effects?	State	State	State	State	County	County	CBSA	CBSA
Type of industry Fixed-Effects?	Ind4	Ind4	Ind4	Ind4	naicsoic	naicsoic	naicsoic	naicsoic
Cluster-Type?	State	State	State	State	County	County	CBSA	CBSA
Sample	1	1	1	1	1	1	1	1

Table 8: Size and the Local Business Cycle at Entry: Tradable versus Non-Tradable Sectors

The table reports coefficients from OLS regressions. The dependent variable in columns 1–5 is *Ln(Emp)* which is defined as the natural logarithm of the number of employees at the establishment-level. *Tradables*, *Non-Tradables*, *Construction*, and *Other* are indicator variables that take the value of one if the industry belongs one of these categories according to the classification in Mian and Sufi (2012). All other variables are defined as in Table 2. Other controls in columns 1–3 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample used in this table includes all establishments born between 1994 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(Emp)	(1)	(2)	(3)
Ln(GDP nom. detrended)$_t$ × Tradables	2.856***		3.667***
	(0.278)		(0.282)
Ln(GDP nom. detrended)$_t$ × Non-Tradables	1.901***		2.313***
	(0.084)		(0.167)
Ln(GDP nom. detrended)$_t$ × Construction	0.676***		1.477***
	(0.149)		(0.155)
Ln(GDP nom. detrended)$_t$ × Other	1.536***		2.040***
	(0.089)		(0.135)
Ln(GDP nom. detrended)$_{st}$ × Tradables		0.400*	-0.903***
		(0.229)	(0.247)
Ln(GDP nom. detrended)$_{st}$ × Non-Tradables		0.287	-0.468***
		(0.171)	(0.169)
Ln(GDP nom. detrended)$_{st}$ × Construction		-0.458**	-0.913***
		(0.199)	(0.269)
Ln(GDP nom. detrended)$_{st}$ × Other		0.092	-0.565***
		(0.113)	(0.117)
Age Fixed-Effects?	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	Yes	Yes	Yes
Sample	1	1	1

Table 9: Size and the Sectoral Business Cycle at Entry

The table reports coefficients from OLS regressions. The dependent variable in columns 1–5 is *Ln(Emp)* which is defined as the natural logarithm of the number of employees at the establishment-level. All other variables are defined in Table 2. Other controls in columns 1–3 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample used in this table includes all establishments born between 1994 and 2001 and their outcomes for up to 10 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP nom. detrended)$_{it}$	0.099***			
	(0.024)			
ln(Demand detrended)$_{it}$		1.084***		
		(0.051)		
Δln(GDP Nom.)$_{it}$			-0.021	
			(0.015)	
Δln(Demand)$_{it}$				0.581***
				(0.040)
R-squared	0.604	0.604	0.604	0.604
Age Fixed-Effects?	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	Yes	Yes	Yes	Yes
Sample	1	1	1	1

Table 10: Labor Productivity and Business Cycle at Entry: Average and Age-Specific Elasticities

The table reports the coefficients of OLS regressions. The dependent variable in columns 1–6 is *Ln(LP)* which is defined as the natural logarithm of the ratio between revenues (defined as value of shipments, sales, receipts, or revenue (in thousands)) and the number of employees at the establishment-level. All other variables are defined in Tables 2 and 3. Other controls in columns 1–6 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, and Ownership Type Fixed-Effects. The observations in columns 4–6 are weighted by the number of employees in the establishment. The sample used in this table includes all establishments born between 1994 and 2011 and their outcomes for up to 5 years. The outcomes of establishments that did not survive over the period are included (except when inactive). Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(LP)	(1)	(2)	(3)	(4)	(5)	(6)
Ln(GDP real detrended)	-0.256**	-0.159		-0.381	-0.459	
	(0.103)	(0.351)		(0.412)	(0.630)	
age × Ln(GDP real detrended)		-0.036			0.028	
		(0.107)			(0.175)	
$\mathbf{1}_{\{age=1\}}$ × Ln(GDP real detrended)			-0.314			-0.977*
			(0.244)			(0.550)
$\mathbf{1}_{\{age=2\}}$ × Ln(GDP real detrended)			0.000			-0.200
			(0.170)			(0.314)
$\mathbf{1}_{\{age=3\}}$ × Ln(GDP real detrended)			-0.183			1.022
			(0.124)			(1.458)
$\mathbf{1}_{\{age=4\}}$ × Ln(GDP real detrended)			-0.416***			-0.121
			(0.138)			(0.436)
$\mathbf{1}_{\{age=5\}}$ × Ln(GDP real detrended)			-0.395*			-1.234**
			(0.226)			(0.573)
R-squared	0.600	0.600	0.600	0.535	0.535	0.535
Age Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Ownership Type Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	SU only	SU only	SU only	SU only	SU only	SU only
Employment-Weighted?	No	No	No	Yes	Yes	Yes
Sample	2	2	2	2	2	2

Table 11: Descriptive Statistics of Industry Characteristics

The table reports the mean and standard deviation of the variables used to characterize the industries, defined by 4-digit NAICS codes. The number of observations reports the number of industries that have information on a particular variable. *Capital Intensity* is defined as the average log ratio of capital stock to total employment (obtained from the NBER-CES Manufacturing Industry Database). *Skill Intensity* is the average educational level of the industry calculated using data on the distribution of educational levels and wage income across industries from the Current Population Survey (author calculations using Current Population Survey from January 2000 to December 2012). *Innovation Rate* is the proportion of businesses that undertake activities that result in the development or introduction of new or significantly improved product or process (obtained from NSF). *Minimum Efficient Scale* is the size of the median establishment within a four-digit NAICS industry category (author calculations using LBD data). *Average Startup Capital* is defined as the average amount of startup or acquisition capital at the two-digit NAICS industry level. This measure is calculated using data from Survey of Business Owners Public Use Microdata Sample for firms established in 2007. *Durability Index* measures the expected life of the industry's products. The index is based on the work of Bils and Klenow (1998). *Differentiation Index* is a product differentiation index based on Gollop and Monahan (1991). The index captures the number of industry products, the inequality in production shares of product lines within an industry, and the dissimilarity of products as measured by the input shares of various intermediate products. *Advertising Index* is the weight of advertising sector in the inputs of each industry (author calculation using BEA detailed IO tables).

	Units	Mean	St. Dev	Obs
Capital Intensity	ln	4.43	0.73	86
Skill Intensity	ln	0.56	0.15	316
Innovation Rate	%	0.23	0.13	199
Minimum Efficient Scale	emp	13.28	28.91	309
Average Startup Capital	$1,000	240.71	116.19	309
Durability Index		7.42	5.20	86
Differentiation Index		0.07	0.03	102
Advertising Intensity	ln	-5.11	1.56	272

Table 12: Cyclicality of Entrants across Different Segments

The table reports the coefficients from OLS regressions. The dependent variable is defined as $\Delta ln(Entrants)$ which is defined as the demeaned difference in logarithms of the number of firms entering in each segment bin during a certain period. A segment bin is a group of firms that share the same 4-digit NAICS industry code, state, ownership type, and establishment-type. $\Delta ln(GDP\ Real)$ is the demeaned difference in logarithms of the annual aggregate real gross domestic product. $\mathbf{1}_{\{MU=1\}}$ is an indicator variable that takes the value of one if the establishment is part of a multi-unit (MU) establishment firm. $\mathbf{1}_{\{Sole-Proprietorship=1\}}$ is an indicator variable that takes the value of one if the establishment is organized as a sole-proprietorship. $\mathbf{1}_{\{Partnership=1\}}$ is an indicator variable that takes the value of one if the establishment is organized as a partnership. $\mathbf{1}_{\{Other\ Type=1\}}$ is an indicator variable that takes the value of one if the establishment is organized in a Ownership form other than a sole-proprietorship, partnership, or corporation. The remaining variables are all standardized and defined as in Table 11. The sample used in this table includes all establishments born between 1978 and 2011. Years 2002 and 2003 are excluded from the sample due to data problems associated with computing the number of entrants in 2002. The number of observations of each regression varies depending on the number of missings in the corresponding independent variable. Standard errors are presented in parentheses. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: $\Delta ln(Entrants)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δln(GDP Real)	2.451***	2.121***	2.039***	2.027***	1.910***	1.966***	1.934***	2.087***	2.204***	2.192***
	(0.041)	(0.057)	(0.102)	(0.036)	(0.052)	(0.038)	(0.038)	(0.098)	(0.075)	(0.042)
Δln(GDP Real) $\times\ \mathbf{1}_{\{MU=1\}}$	-1.946***									
	(0.088)									
Δln(GDP Real) $\times\ \mathbf{1}_{\{Sole-Proprietorship=1\}}$		0.655***								
		(0.102)								
Δln(GDP Real) $\times\ \mathbf{1}_{\{Partnership=1\}}$		0.467***								
		(0.102)								
Δln(GDP Real) $\times\ \mathbf{1}_{\{Other\ Type=1\}}$		-2.812***								
		(0.111)								
Δln(GDP Real)$\times$ Capital Intensity			-0.535***							
			(0.107)							
Δln(GDP Real)$\times$ Skill Intensity				-0.254***						
				(0.038)						
Δln(GDP Real)$\times$ Innovation Rate					-0.091					
					(0.061)					
Δln(GDP Real)$\times$ Minimum Efficient Scale						-0.220***				
						(0.059)				
Δln(GDP Real)$\times$ Average Startup Capital							-0.293***			
							(0.043)			
Δln(GDP Real)$\times$ Durability Index								0.594***		
								(0.090)		
Δln(GDP Real)$\times$ Differentiation Index									-0.556***	
									(0.079)	
Δln(GDP Real)$\times$ Advertising Intensity										-0.549***
										(0.049)
R-Squared	0.003	0.004	0.004	0.003	0.003	0.003	0.003	0.004	0.004	0.003
Sample	1	1	1	1	1	1	1	1	1	1

Table 13: Cyclical Properties of Cohort Size and Initial Entry Size: Contemporaneous Correlation with Business Cycle Indicators

The table reports the correlation between indicators of business cycles and number of entrants, entry rates, and statistics of initial size. The time series of number of entrants, entry rates, and statistics of initial size are HP filtered in the first and second columns and are a demeaned difference of logarithms in the columns 3 and 4. Entry for establishments is defined using the information on the year of birth. The counterfactual average entry size is computed by regressing initial size on time fixed-effects, controlling for industry fixed-effects (naics4), state fixed-effects, single-/multi-unit Fixed-Effects, and Ownership type Fixed-Effects. Entry rate is defined as number of newborns divided by active business (entrants plus continuers). *Ln(GDP real detrended)* and Δ*ln(GDP Real)* are defined as in Table 2. P-values are presented in parentheses. The sample used in this table includes all establishments born between 1978 and 2001, or 1978 and 2010. The period 1978-2001 corresponds to the cohorts used in the main estimations.

	Ln(GDP real detrended)		Δln(GDP Real)	
	1978-01	1978-11	78-01	78-11
Average entry size	0.40	0.37	0.43	0.37
	(0.06)	(0.03)	(0.04)	(0.03)
Average entry size (counterfactual)	0.58	0.41	0.55	0.41
	(0.00)	(0.02)	(0.01)	(0.02)
Standard deviation entry size	0.03	0.05	0.08	0.13
	(0.89)	(0.76)	(0.72)	(0.48)
Median entry size	0.11	0.24	(*)	0.09
	(0.61)	(0.18)	(*)	(0.62)
P90 entry size	0.45	0.42	0.48	0.38
	(0.03)	(0.01)	(0.02)	(0.03)
Entrants	0.17	0.24	0.33	0.40
	(0.43)	(0.17)	(0.13)	(0.02)
Entry rate	0.10	0.17	0.28	0.02
	(0.64)	(0.35)	(0.20)	(0.06)

Table 14: Cyclicality of Initial Size across Different Segments

The table reports the coefficients of OLS regressions. The dependent variable $\Delta ln(Average\ Entry\ Size)$ is defined as the demeaned difference in logarithms of the average initial size, measured as employment in the first year of age, in each market segment bin and period. A market segment bin is a set of firms that share the same four-digit NAICS industry code, state, ownership type, and establishment-type. $\Delta ln(GDP\ Real)$ is the demeaned difference in logarithms of the annual aggregate real gross domestic product. $\mathbf{1}_{\{MU=1\}}$ is an indicator variable that takes the value of one if the establishment is part of a multi-unit (MU) establishment firm. $\mathbf{1}_{\{Sole-Proprietorship=1\}}$ is an indicator variable that takes the value of one if the establishment is organized as a sole-proprietorship. $\mathbf{1}_{\{Partnership=1\}}$ is an indicator variable that takes the value of one if the establishment is organized as a partnership. $\mathbf{1}_{\{Other\ Type=1\}}$ is an indicator variable that takes the value of one if the establishment is organized in a Ownership form other than a sole-proprietorship, partnership, or corporation. The remaining variables are all standardized and defined as in Table 11. The sample used in this table includes all establishments born between 1978 and 2011. Years 2002 and 2003 are excluded from the sample due to data problems associated with computing the number of entrants in 2002. Standard errors are presented in parentheses. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: $\Delta ln(Average\ Entry\ Size)$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δln(GDP Real)	0.765***	1.289***	1.678***	0.851***	0.925***	0.941***	0.891***	1.590***	1.207***	0.875***
	(0.058)	(0.081)	(0.205)	(0.052)	(0.082)	(0.055)	(0.055)	(0.195)	(0.135)	(0.061)
Δln(GDP Real) $\times\ \mathbf{1}_{\{MU=1\}}$	0.427***									
	(0.127)									
Δln(GDP Real) $\times\ \mathbf{1}_{\{Sole-Proprietorship=1\}}$		-0.926***								
		(0.147)								
Δln(GDP Real) $\times\ \mathbf{1}_{\{Partnership=1\}}$		-0.140								
		(0.161)								
Δln(GDP Real) $\times\ \mathbf{1}_{\{Other\ Type=1\}}$		-1.621***								
		(0.193)								
Δln(GDP Real)$\times$ Capital Intensity			0.269							
			(0.215)							
Δln(GDP Real)$\times$ Skill Intensity				-0.053***						
				(0.055)						
Δln(GDP Real)$\times$ Innovation Rate					0.319***					
					(0.096)					
Δln(GDP Real)$\times$ Minimum Efficient Scale						0.460***				
						(0.088)				
Δln(GDP Real)$\times$ Average Startup Capital							0.161***			
							(0.062)			
Δln(GDP Real)$\times$ Durability Index								-0.035		
								(0.179)		
Δln(GDP Real)$\times$ Differentiation Index									0.160	
									(0.143)	
Δln(GDP Real)$\times$ Advertising Intensity										-0.107
										(0.072)
R-Squared	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000
Sample	1	1	1	1	1	1	1	1	1	1

Table 15: Industry-Specific Elasticities and Industry Characteristics

The table reports the coefficients of the second stage OLS regression of a two step procedure. The first stage of the procedure is to estimate the model of equation 8, for each four-digit NAICS industry-level. In the second stage, the dependent variable is defined as κ^i: the coefficient of age×Ln(GDP real detrended) in the first stage equation. The coefficient on age×Ln(GDP real detrended) is a proxy for the persistence of cohort effects, and the regression has the purpose of evaluating how the degree of persistence varies with the characteristics of the industries. *Capital Intensity* is defined as the average log ratio of capital stock to total employment (obtained from the NBER-CES Manufacturing Industry Database). *Diff. Index* is a product differentiation index based on Gollop and Monahan (1991). The index captures the number of industry products, the inequality in production shares of product lines within an industry, and the dissimilarity of products as measured by the input shares of various intermediate products. *Ln (Advertising)* is the weight of advertising sector in the inputs of each industry. All variables are standardized. b^i is the coefficient on the variable Ln(GDP real detrended) in the first stage regression. The unit of observation in the first state regression is an establishment. The unit of observation in the second-stage analysis is a four-digit NAICS code industry group. Each observation in the second-stage regression is weighted by the number of establishments in each four-digit NAICS category. The sample used in this table includes all establishments born between 1978 and 2011. Years 2002 and 2003 are excluded from the sample due to data problems associated with computing the number of entrants in 2002. Standard errors are presented in parentheses. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: κ^i	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Capital Intensity	-0.041	-0.025					-0.056**	-0.032	-0.029	-0.013
	(0.028)	(0.027)					(0.027)	(0.026)	(0.038)	(0.035)
Diff. Index			0.058**	0.058*			0.048	0.029		
			(0.025)	(0.026)			(0.034)	(0.037)		
Ln(Advertising)					0.027*	0.037*			0.033	0.030
					(0.015)	(0.017)			(0.072)	(0.067)
b^i		-0.031**		-0.033**		-0.021		-0.033*		-0.031*
		(0.013)		(0.016)		(0.014)		(0.013)		(0.013)
R-Squared	0.021	0.080	0.082	0.150	0.015	0.045	0.040	0.095	0.023	0.082
N	86	86	101	101	260	260	83	83	86	86

Table 16: Parameter values

	Parameter	Value
Elasticity of demand to capital	η	0.919
Depreciation rate of reputation	δ	0.188
Price elasticity of demand	ρ	1.622
Time preference	β	0.95
Persistence of idiosyncratic shock	ρ_s	0.814
Standard deviation idiosyncratic shock	σ_s	0.26
Persistence of aggregate shock	ρ_z	0.559
Standard deviation aggregate shock	σ_z	0.015
Pareto Location	$\underline{q}$	0.550
Pareto Exponent	ξ	4.410
Mean log operational cost	μ_f	0.561
Std.Dev. log operational cost	σ_f	0.373
Exit shock	γ	0.035
Entry cost	c	1.305
Entry reputation	b_0	12.650

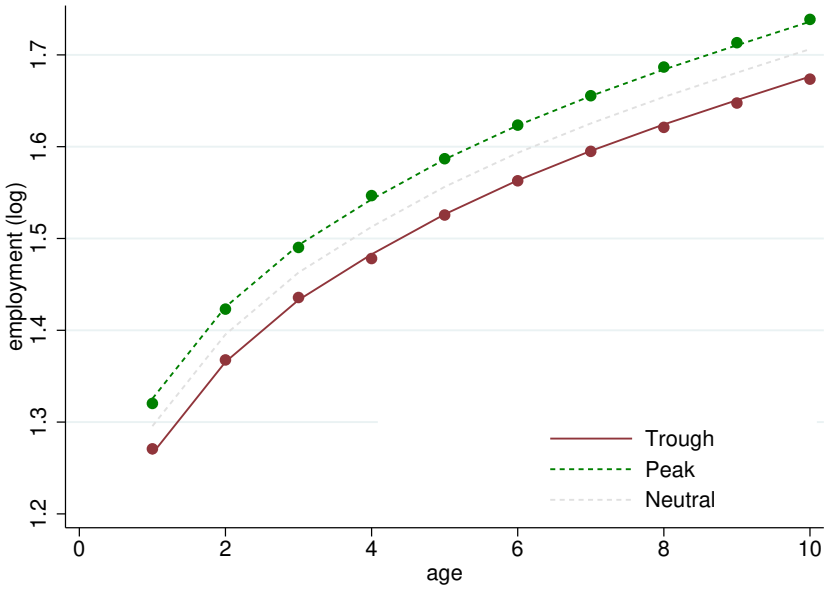
Table 17: Calibrated moments

	Data	Model
Average size	17.0	16.4
Relative size of entrants	0.52	0.48
Relative size of exits	0.49	0.44
Entry rate	11.2	11.2
Survival until 5 years old	0.46	0.49
Survival until between 21 and 25 years old	0.14	0.21

Figure 1: Effect of Economic Conditions at Inception on the Lifecycle Employment of Businesses

The figure plots the estimated effect of detrended real GDP (at inception) on employment over the lifecycle of businesses. The lines are computed using the estimates of the age fixed effects ($\hat{\gamma}_a$) and of the elasticity of size relatively to detrended real GDP at inception ($\hat{\beta}$), computed using specification (1). More specifically, I plot $\hat{c} + \sum_{a=2}^{A} \hat{\gamma}_a D_a + \hat{\beta} Z$, where D_a are age dummies, $\hat{c}$ is the unconditional average log employment of business with one year of activity. The difference between the lines is on the value of Z, which takes three different values $\{-2\sigma, 0, 2\sigma\}$, where σ is the standard deviation of log detrended real GDP in the period 1978-2012. Thus, the solid red represents the simulated size over the life cycle of a business that entered in a bad entry year (with $Z = -2\sigma$), and the dotted green lines represent the simulated size over the life cycle of a comparable business that entered in a good entry year (with $Z = 2\sigma$). The dots are computed using the estimates of the age fixed effects ($\hat{\gamma}_a$) and of the age-specific elasticities of size relatively to detrended real GDP at inception ($\hat{\kappa}$), computed using specification (3). More specifically, they are computed using $\hat{c} + \sum_{a=2}^{A} \hat{\gamma}_a D_a + \sum_{a=2}^{A} \hat{\kappa}_a D_a Z$. Panel A plots the simulated paths using the balanced sample (*sample 1*) which allows to follow business until 10 years of activity. Panel B plots the simulated paths using the unbalanced sample (*sample 1.2*) which allows to follow business for longer periods of activity. Tables 3 and C.1 –panel A present the estimates used in the figure.

Panel A



Panel B

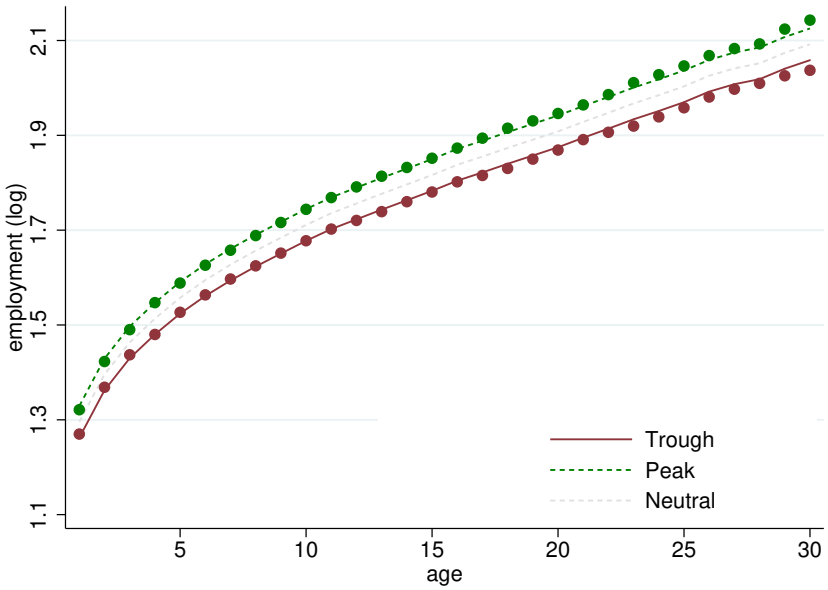


Figure 2: Size and Non-Parametric Cohort Effects

The red line plots the OLS estimates of the cohort fixed-effects on size (in logs), imposing Deaton's normalization to the cohort effects, as defined in equation 4. The red dashed lines correspond to the 95 percent confidence interval. The green line is the Hodrick-Prescott filtered natural logarithm of the annual aggregate real gross domestic product.

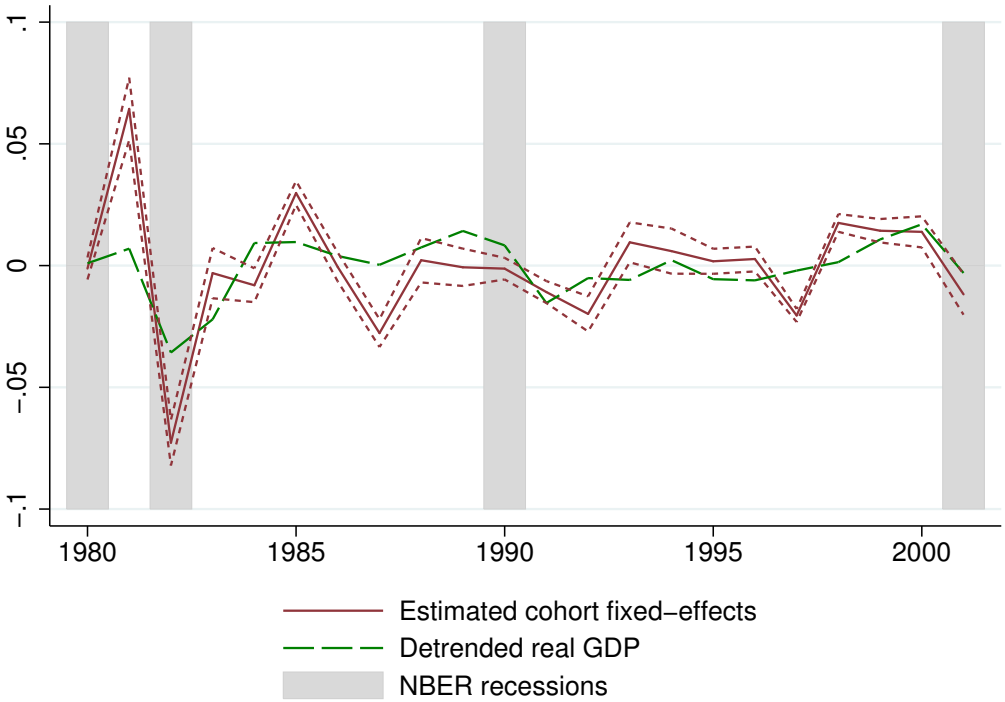


Figure 3: Effect of Economic Conditions at Birth on Labor Productivity in the First Years of Activity

The figure plots the simulated effect of detrended real GDP (at inception) on revenue per employee over the lifecycle of businesses. The lines are computed using the estimates of the age fixed effects ($\hat{\gamma_a}$) and of the elasticity of revenue per employee relatively to detrended real GDP at inception ($\hat{\beta}$), computed using specification (1). More specifically, I plot $\hat{c} + \sum_{a=2}^{A} \hat{\gamma_a} D_a + \hat{\beta} Z$, where D_a are age dummies, $\hat{c}$ is the unconditional average log employment of business with one year of activity. The difference between the lines is on the value of Z, which takes three different values $\{-2\sigma, 0, 2\sigma\}$, where σ is the standard deviation of log detrended real GDP in the period 1978-2012. Thus, the solid red represents the simulated size over the life cycle of a business that entered in a bad entry year (with $Z = -2\sigma$), and the dotted green lines represent the simulated size over the life cycle of a comparable business that entered in a good entry year (with $Z = 2\sigma$). The dots are computed using the estimates of the age fixed effects ($\hat{\gamma_a}$) and of the age-specific elasticities of size relatively to detrended real GDP at inception ($\hat{\kappa}$), computed using specification (3). More specifically, they are computed using $\hat{c} + \sum_{a=2}^{A} \hat{\gamma_a} D_a + \sum_{a=2}^{A} \hat{\kappa_a} D_a Z$. Panel A plots the simulated paths using the balanced sample (*sample 1*) which allows to follow business until 10 years of activity. Table 10 present the estimates used in the figure.

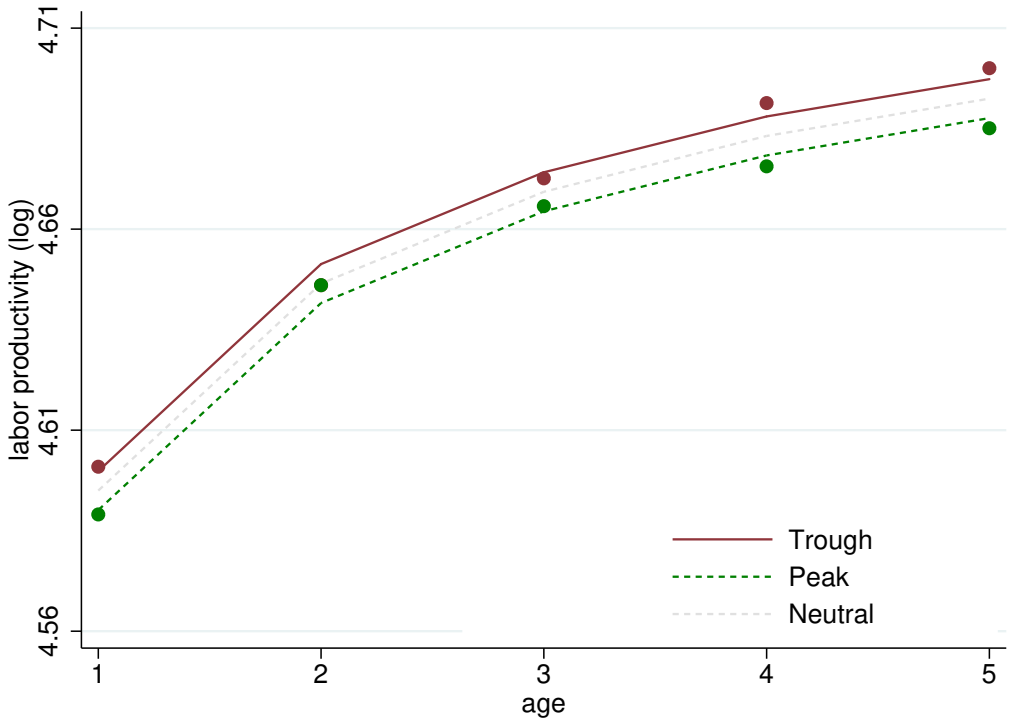
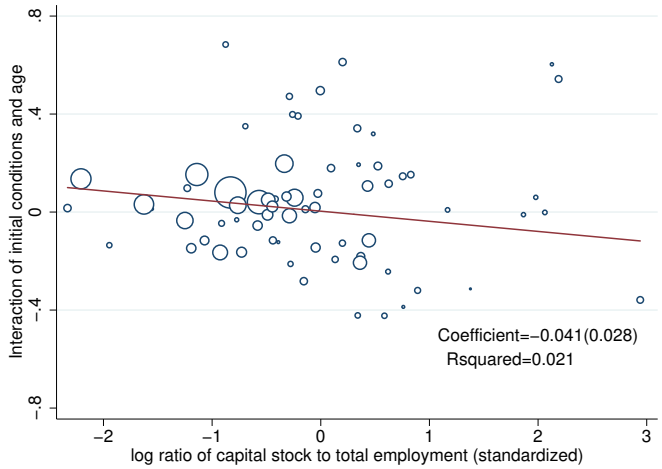


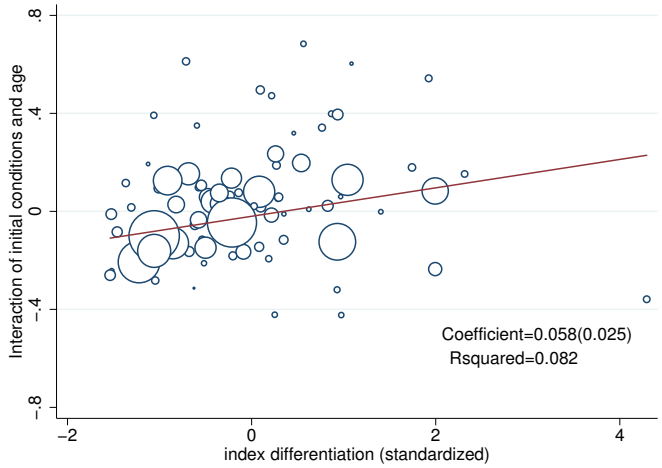
Figure 4: Industry Characteristics and Persistent Cohort Effects

The figure plots the estimated persistence in the average cohort size differences κ_a^i against the proxies for capital intensity, in Panel A, the product differentiation index of Gollop and Monahan (1991), in Panel B, and the advertising index, in Panel C. The variables are standardized and defined as in Table 11. Each circle represents one industry and length of the circle represents the number of observations in each industry.

Panel A:



Panel B:



Panel C:

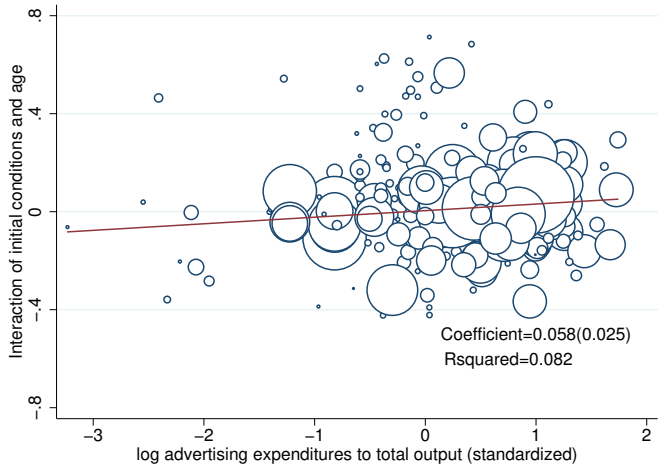


Figure 5: Timing in period t

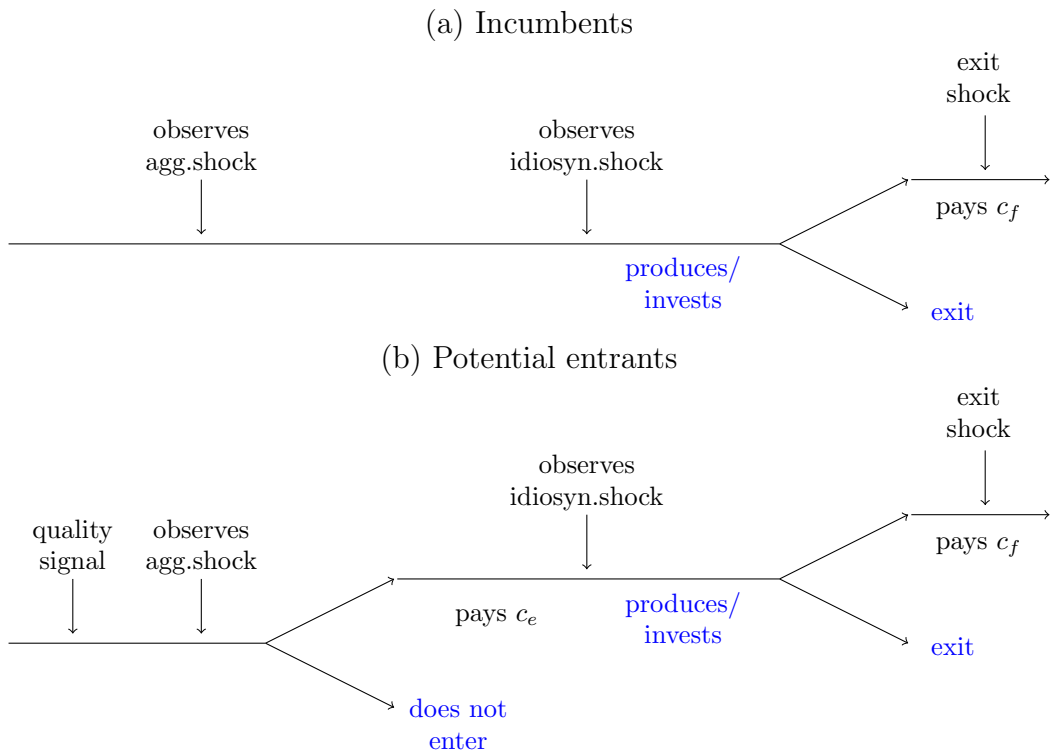


Figure 6: Value of Entering

This figure plots the value of starting operation as a function of the quality signal q received by potential entrant. The green line represents the value of entering as a function of the quality signal when the economy faces a positive aggregate demand shock. The red line represents the value of entering as a function of the quality signal when the economy faces a negative aggregate demand shock. The black dashed line represents the fixed costs of entering the market. The figure suggests that the cutoff signal $q^\star$ below which potential entrants decide not to enter is higher in economic troughs than in peaks.

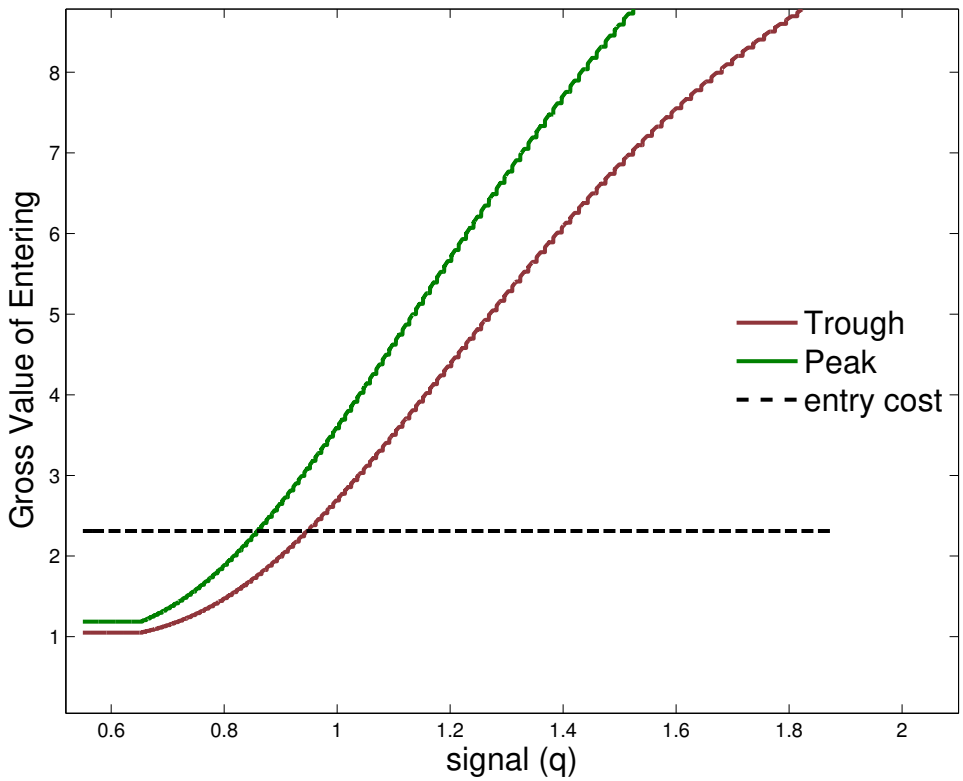


Figure 7: Distribution of Idiosyncratic Productivity at Entry

This figure plots the distribution of idiosyncratic productivities at entry conditional on the level of aggregate demand shocks at the time of entry. The green histogram represents the distribution of productivity for businesses that start operation during a peak. The red histogram represents the distribution of productivity for businesses that start operation during a trough.The peak and trough scenarios are defined as a positive or negative aggregate shock (defined by a positive or negative $\ln(z_t)$), respectively.

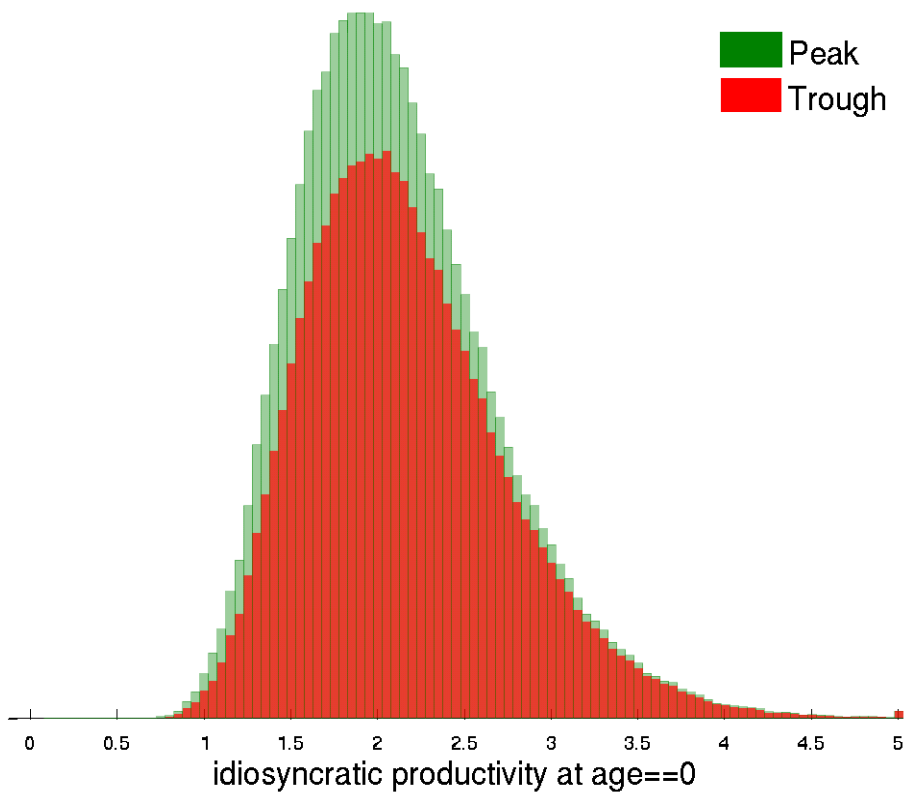
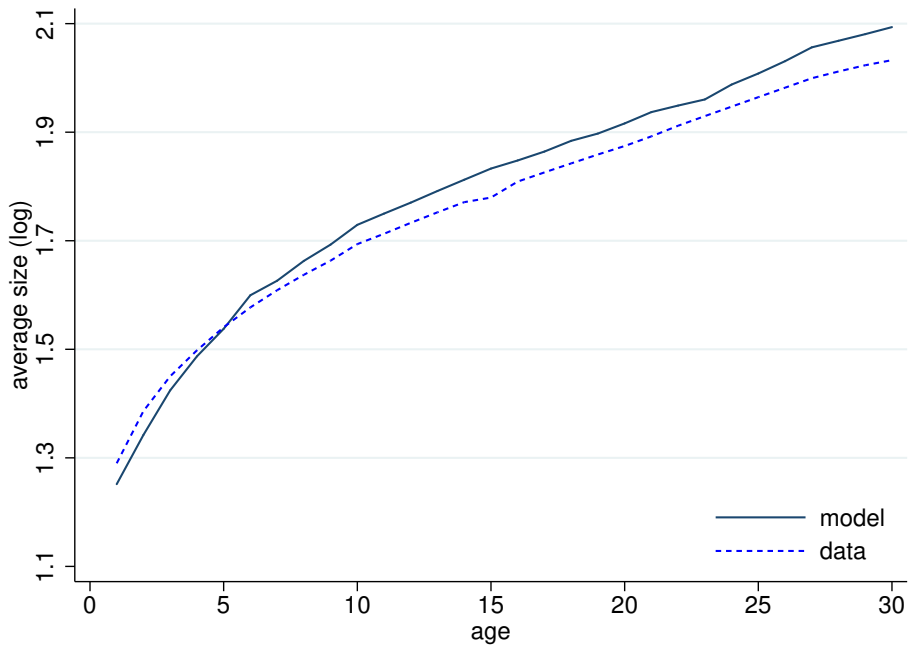


Figure 8: Simulated Paths of Average Size and Survival Functions

This figure plots the evolution of average size and the survival function observed in the data as well as evolution of average size and survival function implied by the simulation of the model. In Panel A, the dashed blue line represents the evolution of the observed average size in the data. The solid grey line represents the evolution of the average size of businesses implied by the model. In Panel B, the dashed blue line represents the evolution of the survival function observed in the data and solid grey line represents the evolution of the survival function implied by the model.

Panel A: Average Size



Panel B: Survival Function

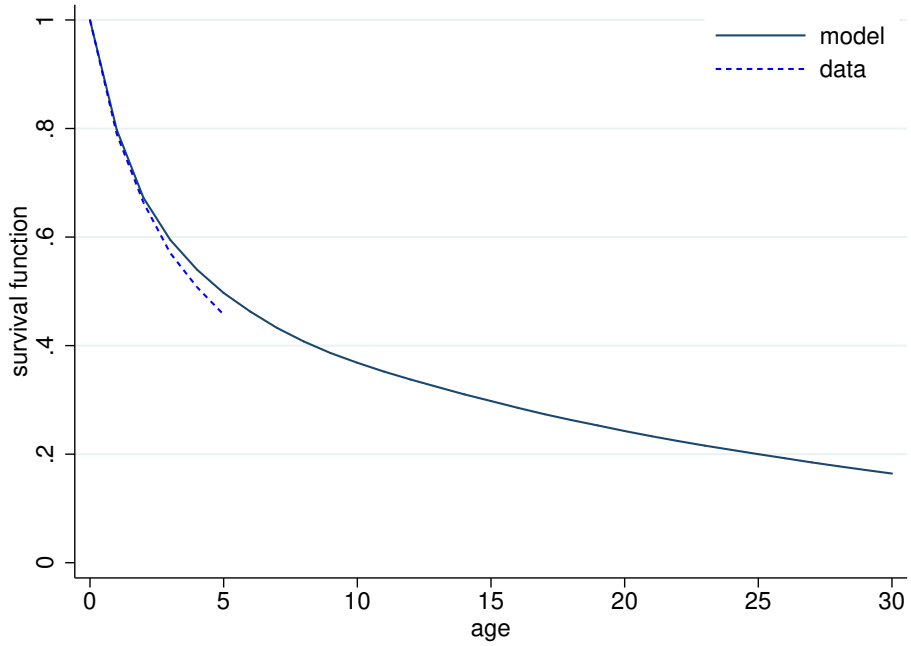


Figure 9: Simulated Survival Functions

This figure plots the survival function using simulated data from the theoretical framework in the paper. The solid red line represents the survival function of a cohort of firms that started operations during a trough. The dashed green line represents the survival function of a cohort of firms that started operations during a economic peak. The peak and trough scenarios are defined as a positive or negative aggregate shock (defined by a positive or negative $\ln(z_t)$), respectively.

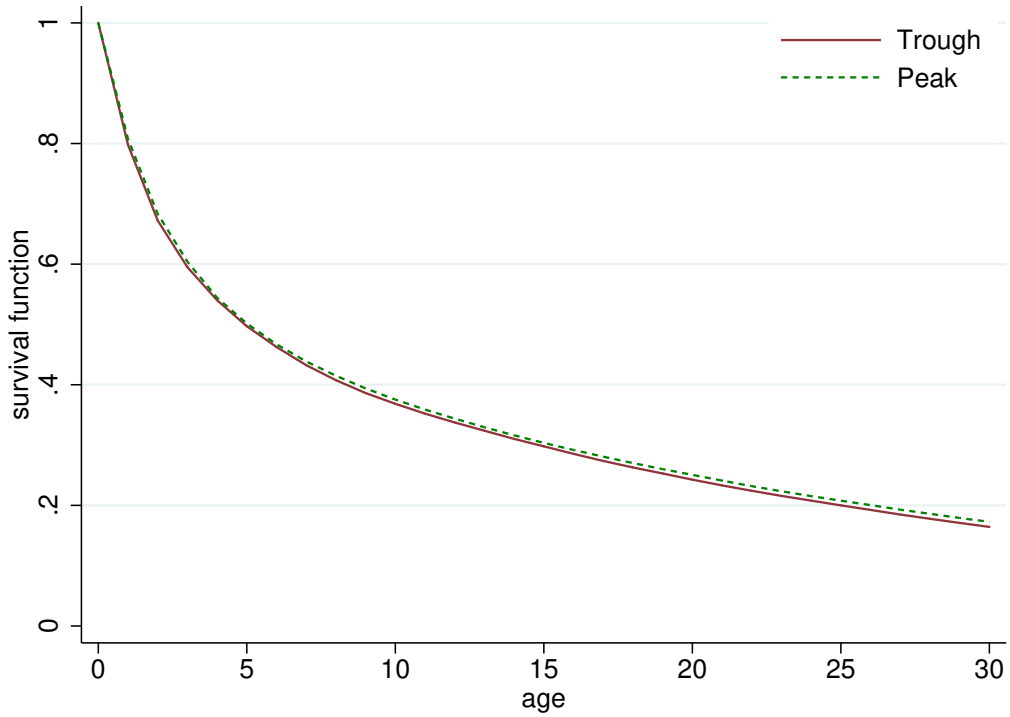


Figure 10: Simulated Effects of Economic Conditions at Inception on the Lifecycle Employment of Businesses

This figure plots the average size (measured as natural logarithm of employment) over the lifecycle of business using simulated data from the theoretical framework in the paper. The solid red line represents the average size of a cohort of firms that started operations during a trough. The solid green line represents the average size of a cohort of firms that started operations during a economic peak. The dashed blue line represents the average size of a counterfactual cohort of businesses that started operations during a peak period. The counterfactual cohort is defined as a cohort that starts during a peak period but whose distribution of entering businesses is equal to that of a cohort that enters during a trough. The peak and trough scenarios are defined as a positive or negative aggregate shock (defined by a positive or negative $\ln(z_t)$), respectively.

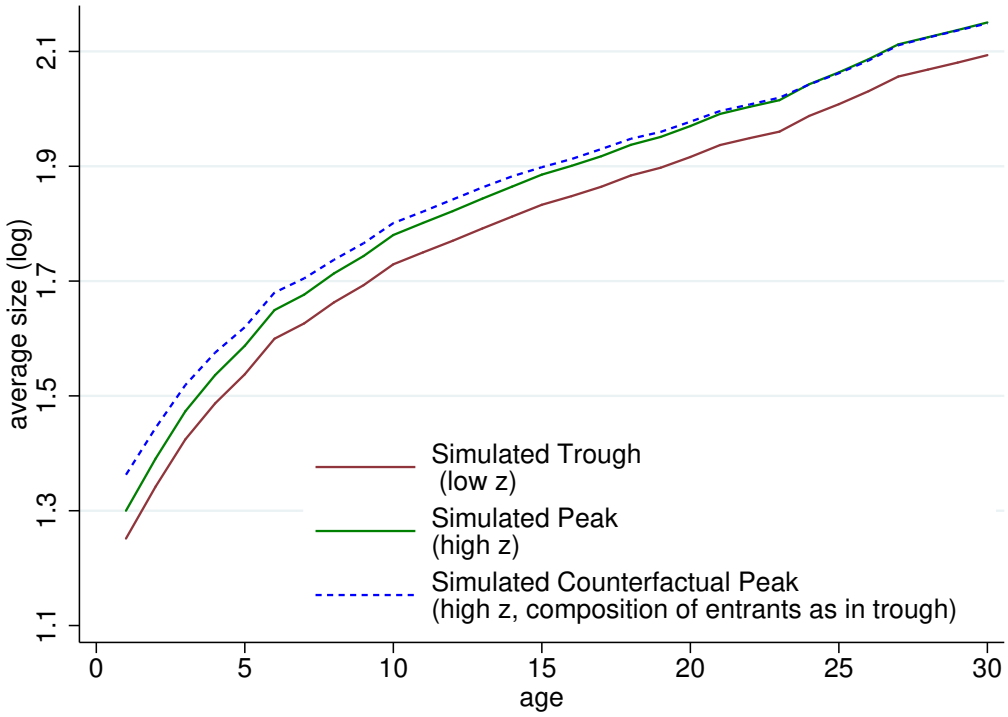


Figure 11: Simulated Differences Between Cohort through Time

This figure plots the average size differences (measured as the difference in the natural logarithm of employment) between cohorts starting in expansions and recession. The solid gray line represents the observed size difference between cohorts that start during peaks and troughs of the business cycle. The dashed orange line represents the counterfactual size difference between a counterfactual cohort that starts during a peak period and a cohort of businesses that starts during a trough. The counterfactual cohort is defined as a cohort that starts during a peak period but whose distribution of entering businesses is equal to that of a cohort that enters during a trough. The peak and trough scenarios are defined as a positive or negative aggregate shock (defined by a positive or negative $\ln(z_t)$), respectively.

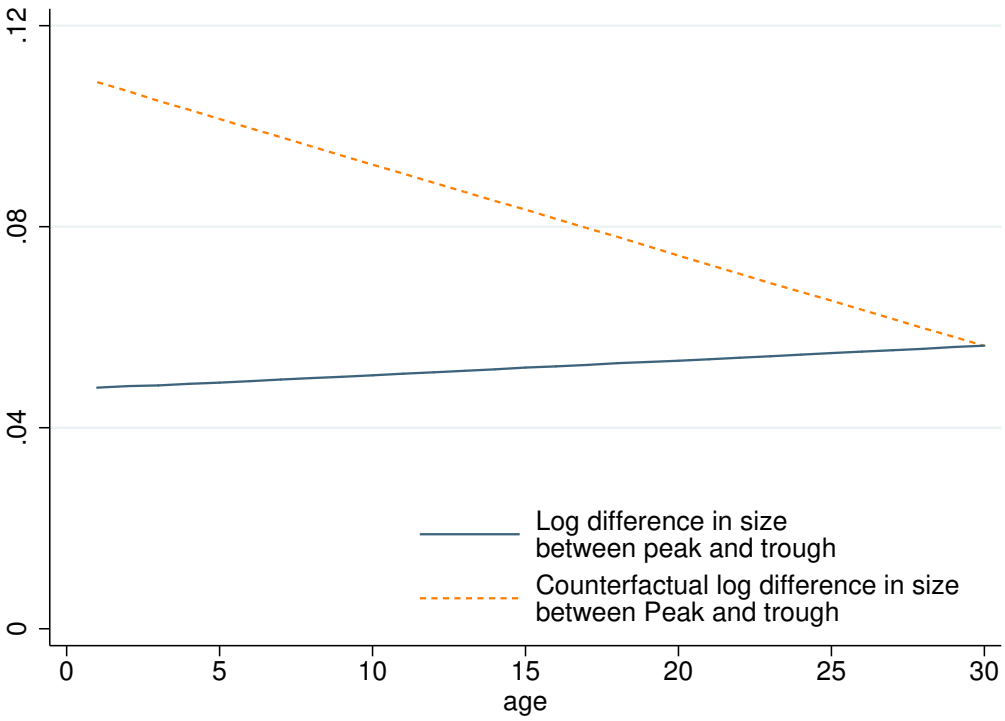


Figure 12: Impulse Response Functions

This figure plots the impulse response of industry output to a positive demand shock at time = 1 under two alternative scenarios. The solid blue line represents the impulse response of industry output when both incumbents and new entrants at time = 1 are affected by the aggregate demand shock. The dashed orange line represents the impulse response of industry output when only incumbents at time = 1 are affected by the aggregate demand shock. The aggregate demand shock is defined as a two standard deviation increase the level of aggregate demand $\ln(z_t)$.

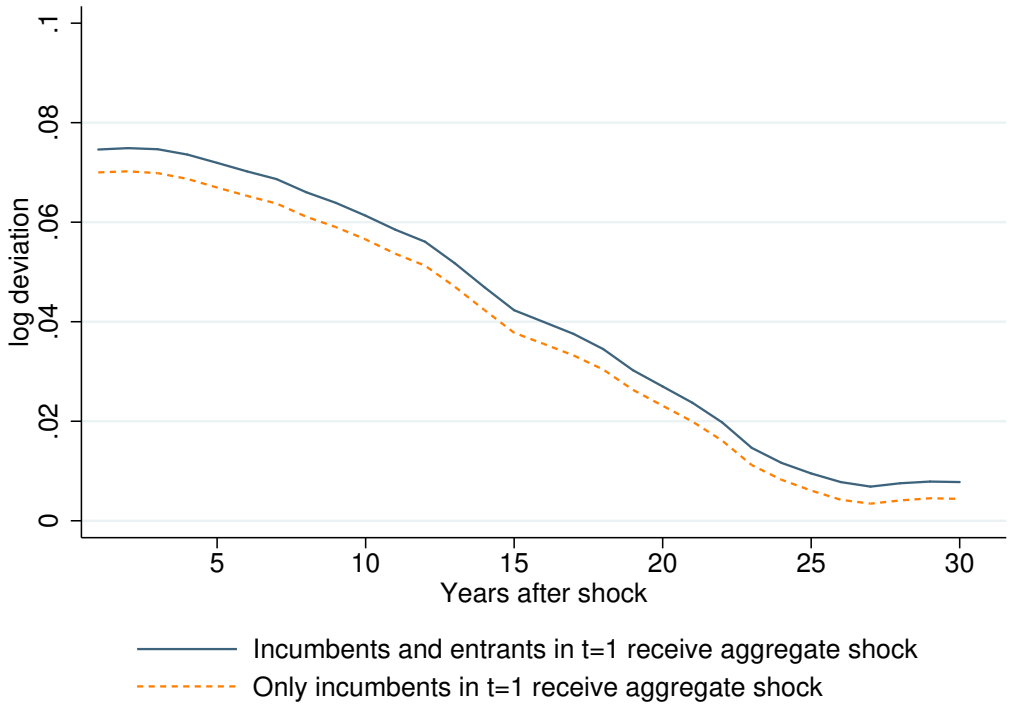


Figure 13: Simulated Propagation of the Great Recession on Number of Businesses, Average Size, and Employment

This figure plots the impulse response of the number of entrants, average size, and total employment to a negative demand shock that mimics the effect of the Great Recession. The negative demand shock is equivalent to a 2.5 standard deviations decrease in the level of the aggregate demand parameter $\ln(z_t)$

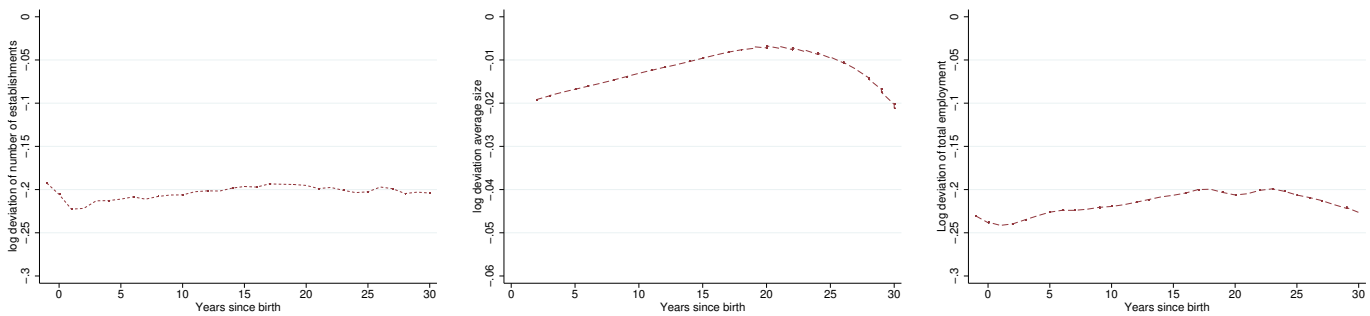
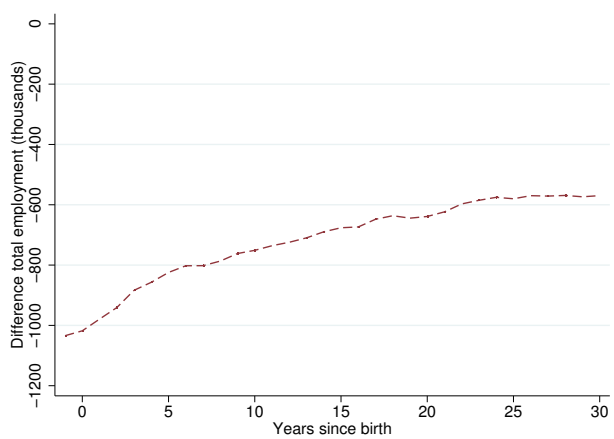


Figure 14: Simulated Propagation of the Great Recession on Total Employment

This figure plots the impulse response of the total employment (in levels) to a negative demand shock that mimics the effect of the Great Recession on the entry margin. The negative demand shock is equivalent to a 2.5 standard deviations decrease in the level of the aggregate demand parameter $\ln(z_t)$



A Data Appendix

A.1 The LBD data

Throughout the paper I use data at the establishment and firm levels. An establishment is defined as a single physical location where business is conducted or where services or industrial operations are performed. A firm is a business organization consisting of one or more establishments that are under common ownership or control. Each establishment-year record is associated with a firm identifier in the LBD. I use the firm identifier key to construct datasets at both the establishment and firm levels. Firms can own a single establishment or many establishments, which may span multiple geographic areas and industries. I compute firm-level characteristics based on the characteristics of the set of establishments of the firm.

Businesses with no paid employees are excluded from the analysis because the LBD only records the universe of non-farm businesses with at least one paid employee. As a result, the period in which an establishment enters the LBD may not necessarily coincide with the period in which the establishment was effectively created. Some establishments may have started operating as non-employer businesses, and then evolve to an employer business.[41]

A.1.1 Establishment-level data

Births, deaths and age

The LBD obtains its data from a variety of sources such as the Business Register (SSEL), Economic Censuses, and surveys. The LBD offers the most reliable and complete data on births, deaths, and age of establishments operating in the US. Nevertheless, Jarmin and Miranda (2002) identified potential problems related to the data collection process that may generate measurement error in the computation of births and deaths of establishments. First, the SSEL uses administrative sources to identify new businesses. Hence, establishments that are not administratively registered may not be identified. This problem is particularly acute for establishments of multi-unit firms. As a supplementary method to collect information on establishments, the Census Bureau directly surveys firms and collects information on their establishments. However, the Census Bureau only surveys the population of firms in Census years, whereas in non-Census years only large firms (as measured by payroll) are surveyed. As a result, the data could display a five year cycle with large swings in the number of these establishments in Census years, when, in reality the entry and exit had occurred over the previous four years. Second, the SSEL contains a substantial number of establishments that appear to become inactive for a period of time. For example, establishments active in $t-1$

[41] Establishments with no payroll are not in the LBD. Once a firm hires an employee, the LBD identifies it as a firm birth. The LBD accounts for firms that alternate between having no employees and having employees using its matching procedures.

and $t+1$ but not in period t. These gaps could also lead to possibly spurious births/deaths.[42] The LBD implements several efforts to mitigate both data issues (for details on the algorithms used, see Jarmin and Miranda (2002)).

Each establishment in the LBD is identified as either a birth, death, or continuer establishment. A birth is assigned when the establishment makes its first appearance in the universe of employer businesses. A death is assigned when the activity of the establishment ceases to exist.[43] LBD also provides information on the first and last year of inclusion in the database. I use this information to identify the cohort year of each establishment.[44] *Age* is defined as the number of years elapsed since the establishment first appeared in the database.

Employment and Payroll

Payroll and employment data are obtained from administrative records for single-unit companies and from a combination of administrative records and survey-collected data for multi-unit companies (Jarmin and Miranda (2002)). Payroll (*pay*) is defined as reported annual payroll (in $1000).[45] It includes all forms of compensation, such as salaries, wages, commissions, dismissal pay, bonuses, vacation allowances, sick-leave pay, and employee contributions to qualified pension plans paid during the year to all employees. For corporations, payroll includes amounts paid to officers and executives, while for unincorporated businesses, it does not include profit or other compensation of proprietors or partners.

Employment (*emp*) is defined as the number of full- and part-time employees (in the payroll) as of March 12th of each calendar year.[46] It includes employees on paid sick leave, holidays, and vacations. Paid employment also includes salaried officers and executives of corporations, but it excludes sole proprietors and partners of unincorporated businesses.[47] Focusing on March-to-March annual changes of employment neglects high frequency within year, firm, and establishment dynamics —for example, transitory or seasonal changes that reverse over the year. At the same time, the total number of employees as of March 12th may be very different from the number of employees that underlie the payroll information. In robustness exercises, I constructed an alternative measure of employment (*empa*) in year t that I define as the average employment between March of year t and March of year $t+1$ (and employment in $t/t+1$ if missing in $t+1/t$).

Single- and multi-unit status

The single-multi identifier indicates if an establishment belongs to a firm that owns two or more establishments, as opposed to having only one establishment. I define a firm as a single-unit firm if

[42] The gaps may be generated by transitions between employer and non-employers status or measurement error.

[43] Note that establishments that change ownership (sold to another owner) are not considered an establishment death.

[44] I consider different criteria as robustness check. For example, I also use the BDS criteria, where birth year is defined as the year an establishment first reports positive employment in the LBD.

[45] In some cases this value may be imputed due to missing or invalid data.

[46] In some cases this value may be imputed due to missing or invalid data.

[47] By definition this value should at least 1. This is not always the case for births and deaths.

it has only one establishment and as a multi-unit firm when it has more than one establishment.

Industry and geography

The LBD also includes detailed information on industry and location. The period of analysis coincides with the transition from SIC to NAICS industry classifications. This overlap creates some difficulties in the industry classification because the SIC-NAICS crosswalk is a many-to-many mapping and some establishments in the database are only classified under one of these alternative classification arrangements. A consistent classification system for the time-series is critical for the analysis. Fort (2013) developed a methodology that reclassifies all establishments in the LBD according to a consistent NAICS (2002) industry classification system (the *fk_naics02* codes). After implementing the Fort (2013) procedure, I use the industry codes at different levels of disaggregation throughout the paper.[48] I use industry information disaggregated at the 4-digit NAICS codes for most of the empirical analysis (the *naics4digit* classification contains approximately 300 industry groups).[49] Alternatively, I employ an alternative classification of establishments (the *naicsoic* contains approximately 61 industry groups and is similar to the 3-digit NAICS level) to match the establishment-level data in the LBD with aggregate data from national statistical agencies. This is the finest definition in which all categories are mapped to BEA NAICS industry data since late 1970s.

The LBD contains information on the geographic location of establishments at the zip-code level. Throughout the paper I use information on the state and county where the establishment is located (Census FIPS codes). The states and counties included in the empirical analysis completely cover the land area and population of the United States. Among the 50 states and Columbia District, there are approximately 3140 counties.[50] To match the LBD to external data, I also map the location to core-based statistical areas (CBSAs) using the 2007 counties-CBSA crosswalk. I identified 920 core-based statistical areas, in which 380 of these areas are classified as metropolitan areas and the remaining are classified as micropolitan statistical areas. In addition, I also used the New England county metropolitan areas (NECMAs) classification for the six New England States to match to external data.

Ownership Type

Firms are classified into different legal business entities, which determine their income tax treatment and limited liability regime. I use information from the LBD to classify establishments into corporations, sole proprietorships, partnerships, and a residual category for other forms. Non-incorporated businesses include sole proprietors and partnerships.

[48] I complemented with alternative information available in the LBD and map it to 2002 NAICS codes. For around 10 per cent of the establishments-year there is no information on detailed industry.

[49] There are some industry codes not covered under the universe of the CBP.

[50] Over the period under analysis the boundaries of the counties were broadly stable, with some occasional changes.

A.1.2 Firm-level data

Births, deaths and age

I am also able to identify births of firms and to follow them throughout their lifecycle using the LBD. However, the Construction of firm links presents conceptual and measurement challenges. I follow the approach adopted for the BDS and based on the work of Davis et al. (2007) and Haltiwanger et al. (2013) that present measures that are robust to ownership changes. A new firm identifier emerges in the LBD either because a new firm is born or because an existing businesses undergoes a change of ownership and control (e.g. merger and acquisition, divestitures). When a new firm identifier appears in the LBD due to a *de novo* firm birth, a firm birth is registered. By contrast, when a new firm identifier arises through a merger of two preexisting firms, it is not treated as a firm birth and is assigned the age of the oldest continuing establishment of the newly combined business. The firms are then allowed to age naturally regardless of mergers and acquisitions as long as the ownership and control does not change. Therefore, a firm birth is defined as a new firm identifier in which all the establishments at the firm are new (entering) establishments. Similarly, a firm death is determined when a firm identifier disappears and all associated establishments cease operations and exit.

Employment and Payroll

The level of employment of the firm is computed by aggregating employment across all establishments that belong to the firm. Haltiwanger et al. (2013) discuss how the evolution of employment levels may reflect organic growth within establishments and changes in the number of establishments that stem from changes in firm structure such as mergers, acquisitions, and divestitures. In the context of this paper I do not explicitly explore this distinction. Payroll is also constructed by summing payroll across all establishments that belong to the firm.

Single- and multi-unit status and number of establishments

I account for the nature of the firm by controlling for whether it has a unique location (single-unit) or multiple locations (multi-unit). I also compute the number of locations the firm has in case of multi-unit firms.

Industry and geography

Firms often own establishments that are classified in different industries. I use the industry classification of the establishments and the corresponding firm identifiers to compute the main industry of the firm. I define the main industry of the firm using two alternative methods and for two different levels of industry disaggregation (*naics4digit* and *naicsoic*). In the first procedure, I define the main industry as the industry group with the largest share of employment within the firm. Alternatively, I define the main industry of the firm as the mode industry among all establishments owned by the firm.[51] In addition, to account for the degree of diversity of the multi-sector firms, I computed the

[51] In case of multiple modes, I arbitrarily choose the industry with the smaller code.

number of different industries in which the firm operates.

Multi-unit firms often have establishments in multiple geographies, especially if these geographies are defined at a local level. I examine the location of the establishments of multi-unit firms to determine the main location of the firm. Similar to the procedures used to compute the main industry classification of the firm, I define the main location of the firm at different local levels (from state to cbsa level) as the location that employees more workers within the firm or the mode location among all establishments of the firm. I also account for the diversity by computing the number of different locations.

Ownership Type

Similar to establishments, firms are classified into different legal business entities, which determine their income tax treatment and limited liability status. Using information pertaining these categories from the establishment data and the corresponding firm identifiers, I classify firms into corporations, sole proprietorships, partnerships, and a residual category for other forms.

A.2 Revenue data

Revenue is defined as the "value of shipments, sales, receipts or revenue (in dollars)". I constructed this measure from the Census Bureau's SSEL files. Obtaining revenue data entails several challenges. First, the BR's revenue measure is based on administrative data from annual business tax returns. Unlike payroll and employment, which are measured at the establishment level going back to 1976, the nominal revenue data are only available at the tax reporting or employer identification number (EIN) level starting in the mid-1990s. Thus, in the SSEL, revenue is only measured at the establishment-level for single-unit firms. For multi-unit firms we can obtain firm-level revenues but it is not possible to obtain establishment-level measures.[52] Second, the content of the receipts data varies substantially by type of industry and legal structure of the firm depending on its tax treatment status. Third, the missing observations problem is more acute for the SSEL receipts fields than for other fields that are critical to the Census' operations such as payroll and employment. For this reason, problems with non-random missing observations are likely to be more acute for this variable. I assess the severity of this problem by comparing the main results using the full sample with those that I obtain by using the subsample of firms that have non-missing values for revenue data and find similar results. Even though, the missing observations problem is more acute in the revenue data, these results assuage concerns that this sample selection issue introduces significant bias in the regression.

[52]The SSEL data are divided into two microdata files: one of single-establishments and submasters, and the second a file of multi-establishments and submasters. Submasters are defined as the aggregation of multi-establishments to the EIN level, and thus the first microdata file contains the universe of all EIN's in the U.S. The second microdata files are the establishment-level data for those EINs with multiple establishments within the submaster. Importantly, the multi-establishments files have no sales data, and as a result, it is not possible to obtain a time series of non-imputed annual sales data at the establishment level from the SSEL.

I deflate the nominal revenue measures using a general price deflator and a industry-specific price deflator. The general price deflator is the GDP implicit price deflator. For the industry-specific price deflator, I use the BEA's Industry Accounts to compute value-added implicit price deflators at the industry level. I use the same procedure to obtain real measures of payroll. Throughout the paper, I employ the real revenue measure obtained from using industry specific deflators. This real revenue measure measures aggregated real revenue changes and accounts for changes in relative prices across industries.

B Model Appendix

B.1 Numerical Approximation

B.1.1 General

The **incumbent's value function** is approximated by value function iteration, following these steps:

1. Define the grids for the state variables s, z, and b. Denote them as I_s, I_z, and I_b, respectively. The reputational capital grid is constructed following the method suggested by Fackler and Miranda (2002). The grids and transition matrices for the two shocks are constructed following Tauchen (1986). For all pairs $(s, s\prime)$ such that $s, s\prime \in I_s$, let $P^S(s\prime|s)$ denote the probability that next period's idiosyncratic shock equals $s\prime$, conditional on being s today. For all $(z, z\prime)$ such that z, $z\prime \in I_z$, let also $P^Z(z\prime|z)$ denote the probability that next period's aggregate shock equals $z\prime$, conditional on being z today.

2. For all elements (s, z, b), guess values for the initial value function $V_0^I(s, z, b)$

3. Construct a revised guess value function $V_1^I(b, s, z)$ by solving

$$V_1^I(b,s,z) = \max_{b \in I_b} \left\{ \Pi(b,s,z) + Pr(f \leq f^\star(b,s,z))\, \beta(1-\gamma) \left[\mathbb{E}\left(V^0(b\prime, s\prime, z\prime)\right) - E(f \mid f \leq f^\star(b,s,z)) \right] \right\}$$

   subject to

$$\Pi(b,s,z) = \left(\frac{b'}{1-s} - b\right) - \frac{w}{s}\left(\frac{b'}{1-s} - b\right)^{\frac{\rho}{\rho-1}} b^{\frac{-\eta}{\rho-1}} (\alpha z)^{\frac{-1}{\rho-1}}$$

$$\mathbb{E}\left(V^0(b,s,z)\right) = \sum_i \sum_j V^0(b', s, z)\, P^Z(z_j \mid z)\, P^S(s_i \mid s)$$

   and $f^\star$ is determined by the value of fixed cost that makes the firm indifferent between continuing and exiting (note that the value of exit is equal to zero) $f^\star = E(V^0(b', s\prime, z\prime))$

4. Keep on iterating until sup $|\frac{V_{n+1}^I(s,z,b) - V_n^I(s,z,b)}{V_n^I(s,z,b)}| < 10.0^{-3}$. Denote this last function as $V^{I\star}(s, z, b)$

For the **entrant's value function**, I follow the following steps:

1. Define the grids for the signal q, denoted as I_g. The grid is constructed following the method suggested by Fackler and Miranda (2002). Let $P^G(s|q)$ denote the probability that this period's idiosyncratic shock equals s, conditional on having a signal q.

2. For all grid points (q, z), compute the gross value of entering as

$$V^{P,\text{gross}}(b_0, q, z) = \sum_i [P^G(s_i|q) V^{I\star}(b_0, s_i, z)]$$

where $V^{I\star}(b, s, z)$ is incumbents value function.

3. Obtain the value of entering as follows

$$V^P(b_0, q, z) = -c + V^{P,\text{gross}}(b_0, q, z)$$

4. Define $q^\star(z)$ as the value of the signal which makes prospective entrants indifferent between entering and not. That is $V^P(b_0, q, z) = 0$.

For the **simulation**:

1. Consider a mass M of potential entrants, and a number of periods to simulate T.

2. For each m draw a signal $\tilde{q}$ from the Pareto distribution with shape ξ and location parameter $\underline{q}$. For each potential entrant draw ϵ from normal distribution with mean zero an variance 1. Along with $P^G(s|q)$, compute for each potential entrant the sequence of idiosyncratic shocks $\tilde{s}_{it}$. For each potential entrant and for each time period draw a fixed cost from the lognormal distribution with parameters μ_f and σ_f, and a exit shock from the uniform distribution.

3. For $t = 1$, start with an arbitrary initial condition z_0. Draw subsequent shocks ε and obtain a sequence of z's.

4. Simulate the decisions of incumbents and entrants for all periods from $t = 0$ until T, using the policy functions of incumbents and $q^\star(z)$.

5. For each period compute the following aggregates: number of entrants, number of exits, number of active firms, average size of entrants, average size of exits, average size of all active firms, and total employment.

B.1.2 Internal Calibration

The algorithm used to solve the model consists of the following steps:

1. For each parameter to be internally calibrated define the initial guess, upper and lower bounds;

2. Define the target moments and the simulated counterparts;

3. Start with guess parameters and solve the model in steady state: compute the value function of the incumbents, the potential entrants, simulate the economy and compute moments.

4. Use John D'Errico matlab function *fminsearch* to find new guess

5. Repeat 3 to 4 until simulated moments aproximate target.

B.1.3 Impulse Response

C Tables and Figures Appendix

Table C.1: Size and the Business Cycle at Entry: Alternative Selection Criteria

This table replicates the analysis of Table 3 using different sample selection criteria. Sample 1.1 is an unbalanced sample that differs from Sample 1 by including observations for up to 34 years. Sample 1.2 is an unbalanced sample that differs from Sample 1 by including cohorts from 1978 to 2010 and observations for up to 34 years. In this sample the oldest cohort has establishments of up to 34 years old and the youngest cohort has observations in their first year of activity. Sample 1.3. is a balanced sample that differs from Sample 1 because it only includes establishments that survived at least 10 years. Sample 2 includes all establishments born between 1994 and 2011 and their outcomes for up to 5 years. Other controls include fixed-effects for age, year, state, industry, ownership type, single/multi-unit status. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Panel A: Sample 1.1

Dep var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	1.303***	1.018***		
	(0.087)	(0.118)		
$age \times$Ln(GDP real detrended)		0.029***		
		(0.006)		
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			1.000***	
			(0.122)	
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			1.057***	
			(0.100)	
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			1.032***	
			(0.092)	
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			1.308***	
			(0.098)	
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			1.203***	
			(0.097)	
⋮				
$\mathbf{1}_{\{age=21\}} \times$Ln(GDP real detrended)			1.427***	
			(0.093)	
$\mathbf{1}_{\{age=22\}} \times$Ln(GDP real detrended)			1.549***	
			(0.110)	
$\mathbf{1}_{\{age=23\}} \times$Ln(GDP real detrended)			1.785***	
			(0.126)	
$\mathbf{1}_{\{age=24\}} \times$Ln(GDP real detrended)			1.733***	
			(0.128)	
$\mathbf{1}_{\{age=25\}} \times$Ln(GDP real detrended)			1.719***	
			(0.120)	
⋮				
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.025***
				(0.002)
R-squared	0.582	0.582	0.582	0.582
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C1: Size and the Business Cycle at Entry: Alternative Selection Criteria (cont'd)

Panel B: Sample 1.2

Dep var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	1.177***	0.867***		
	(0.067)	(0.086)		
$age \times$Ln(GDP real detrended)		0.034***		
		(0.005)		
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			0.844***	
			(0.091)	
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			0.806***	
			(0.074)	
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			0.746***	
			(0.065)	
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			0.997***	
			(0.080)	
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			1.080***	
			(0.082)	
$\vdots$				
$\mathbf{1}_{\{age=21\}} \times$Ln(GDP real detrended)			1.524***	
			(0.090)	
$\mathbf{1}_{\{age=22\}} \times$Ln(GDP real detrended)			1.657***	
			(0.120)	
$\mathbf{1}_{\{age=23\}} \times$Ln(GDP real detrended)			1.799***	
			(0.132)	
$\mathbf{1}_{\{age=24\}} \times$Ln(GDP real detrended)			1.738***	
			(0.119)	
$\mathbf{1}_{\{age=25\}} \times$Ln(GDP real detrended)			1.679***	
			(0.107)	
$\vdots$				
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.020***
				(0.002)
R-squared	0.582	0.582	0.582	0.582
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C1: Size and the Business Cycle at Entry: Alternative Selection Criteria (cont'd)

Panel C: Sample 1.3

Dep var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	1.385***	1.352***		
	(0.104)	(0.111)		
$age \times$Ln(GDP real detrended)		0.006		
		(0.008)		
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			1.335***	
			(0.122)	
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			1.362***	
			(0.104)	
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			1.246***	
			(0.106)	
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			1.439***	
			(0.110)	
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			1.402***	
			(0.110)	
$\mathbf{1}_{\{age=6\}} \times$Ln(GDP real detrended)			1.429***	
			(0.105)	
$\mathbf{1}_{\{age=7\}} \times$Ln(GDP real detrended)			1.442***	
			(0.114)	
$\mathbf{1}_{\{age=8\}} \times$Ln(GDP real detrended)			1.462***	
			(0.100)	
$\mathbf{1}_{\{age=9\}} \times$Ln(GDP real detrended)			1.407***	
			(0.109)	
$\mathbf{1}_{\{age=10\}} \times$Ln(GDP real detrended)			1.294***	
			(0.107)	
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.024***
				(0.002)
R-squared	0.586	0.586	0.586	0.586
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C1: Size and the Business Cycle at Entry: Alternative Selection Criteria (cont'd)

Panel D: Sample 2

Dep var: Ln(Emp)	(1)	(2)	(3)
Ln(GDP real detrended)	0.838***	1.226***	
	(0.078)	(0.266)	
$age \times$Ln(GDP real detrended)		-0.145	
		(0.097)	
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			1.349***
			(0.178)
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			0.391***
			(0.117)
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			0.781***
			(0.103)
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			0.703***
			(0.190)
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			0.680***
			(0.218)
R-squared	0.579	0.579	0.579
Fixed-Effects?	Yes	Yes	Yes

Table C.2: Size and the Business Cycle at Entry for Firms: Average and Age-Specific Elasticities

This table replicates the analysis of Table 3 using the natural logarithm of the number of employees at the firm level as the dependent variable. The main variables of interest are defined as in Table 3, except for *age*, which is defined as the age of the oldest establishment within the firm. Other controls in columns 1–5 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, and Ownership Type Fixed-Effects. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: Ln(Emp)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	0.269**	0.273**		
	(0.109)	(0.105)		
$age \times$Ln(GDP real detrended)		-0.001		
		(0.011)		
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			0.527***	
			(0.133)	
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			0.218**	
			(0.100)	
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			0.065	
			(0.098)	
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			0.117	
			(0.113)	
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			0.143	
			(0.115)	
$\mathbf{1}_{\{age=6\}} \times$Ln(GDP real detrended)			0.349**	
			(0.132)	
$\mathbf{1}_{\{age=7\}} \times$Ln(GDP real detrended)			0.329**	
			(0.143)	
$\mathbf{1}_{\{age=8\}} \times$Ln(GDP real detrended)			0.425***	
			(0.130)	
$\mathbf{1}_{\{age=9\}} \times$Ln(GDP real detrended)			0.254**	
			(0.126)	
$\mathbf{1}_{\{age=10\}} \times$Ln(GDP real detrended)			0.278**	
			(0.114)	
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.008***
				(0.003)
R-squared	0.468	0.468	0.468	0.468
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C.3: Size and Business Cycle at Entry: Alternative Measure of Employment

This table replicates the analysis of Table 3 using the natural logarithm of the average between the number of employees at the establishment-level in year t and the number of employees at the establishment-level in year *t+1* as the dependent variable. All other variables are defined as in Table 3. Other controls in columns 1–5 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-Multi Unit Fixed Effects. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: Ln(Empa)	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	1.140***	0.987***		
	(0.098)	(0.100)		
$age \times$Ln(GDP real detrended)		0.031***		
		(0.009)		
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			0.889***	
			(0.120)	
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			1.049***	
			(0.100)	
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			1.089***	
			(0.094)	
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			1.313***	
			(0.100)	
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			1.197***	
			(0.095)	
$\mathbf{1}_{\{age=6\}} \times$Ln(GDP real detrended)			1.175***	
			(0.112)	
$\mathbf{1}_{\{age=7\}} \times$Ln(GDP real detrended)			1.173***	
			(0.121)	
$\mathbf{1}_{\{age=8\}} \times$Ln(GDP real detrended)			1.202***	
			(0.114)	
$\mathbf{1}_{\{age=9\}} \times$Ln(GDP real detrended)			1.240***	
			(0.112)	
$\mathbf{1}_{\{age=10\}} \times$Ln(GDP real detrended)			1.243***	
			(0.101)	
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.020***
				(0.002)
R-Squared	0.606	0.606	0.606	0.606
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C.4: Size and the Business Cycle at Entry: Revenue as a Proxy for Size

The table reports the coefficients of OLS regressions. The dependent variable in the table *Ln (Rev)* is the natural logarithm of the value of shipments, sales, receipts, or revenue (in thousands). All independent variables are defined as in the main tables Other controls in columns 1–5 include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, and Ownership Type Fixed-Effects. The observations are not weighted by the number of employees in the establishment. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: Ln(Rev)	(1)	(2)	(3)	(4)	(5	(6)	(7)
Ln(GDP real detrended)	0.893***	1.355**					
	(0.107)	(0.594)					
$age \times$Ln(GDP real detrended)		-0.173					
		(0.206)					
$\mathbf{1}_{\{age=1\}} \times$Ln(GDP real detrended)			1.318***				
			(0.376)				
$\mathbf{1}_{\{age=2\}} \times$Ln(GDP real detrended)			0.720***				
			(0.253)				
$\mathbf{1}_{\{age=3\}} \times$Ln(GDP real detrended)			0.913***				
			(0.101)				
$\mathbf{1}_{\{age=4\}} \times$Ln(GDP real detrended)			0.612*				
			(0.316)				
$\mathbf{1}_{\{age=5\}} \times$Ln(GDP real detrended)			0.601				
			(0.446)				
Ln(GDP nom. detrended)				0.705***			
				(0.135)			
Δln(GDP Real)					0.098		
					(0.240)		
$\Delta ln(Emp)$						0.520***	
						(0.115)	
Δln(PIpc)							0.074
							(0.151)
R-squared	0.581	0.581	0.581	0.581	0.581	0.581	0.581
Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	SU only	SU only	SU only	SU only	SU only	SU only	SU only
Employment-Weighted?	No	No	No	No	No	No	No
Sample	2	2	2	2	2	2	2

Table C.5: Growth and the Business Cycle at Entry: Employment Growth

This table replicates the analysis of Table 3 using Employment growth at the establishment level as the dependent variable. Employment growth is defined as number of employees at the establishment-level in year t minus the number of employees at the establishment level in year $t-1$ divided by the average number of employees of the establishment in years t and $t-1$. All other variables as defines as in the main tables. Other controls include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, Ownership Type Fixed-Effects, and Single-/Multi-Unit Fixed-Effects. The sample in this analysis uses the sample criteria of sample 1, but entails more missing observations in the dependent variable which reduces the number of observations in the analysis. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep var: Employment Growth	(1)	(2)	(3)	(4)
Ln(GDP real detrended)	0.040***	0.157***		
	(0.013)	(0.025)		
age×Ln(GDP real detrended)		-0.024***		
		(0.003)		
$\mathbf{1}_{\{age=1\}}$×Ln(GDP real detrended)			-0.126***	
			(0.045)	
$\mathbf{1}_{\{age=2\}}$×Ln(GDP real detrended)			0.225***	
			(0.034)	
$\mathbf{1}_{\{age=3\}}$×Ln(GDP real detrended)			0.317***	
			(0.023)	
$\mathbf{1}_{\{age=4\}}$×Ln(GDP real detrended)			0.061**	
			(0.028)	
$\mathbf{1}_{\{age=5\}}$×Ln(GDP real detrended)			0.202***	
			(0.030)	
$\mathbf{1}_{\{age=6\}}$×Ln(GDP real detrended)			0.005	
			(0.059)	
$\mathbf{1}_{\{age=7\}}$×Ln(GDP real detrended)			-0.188**	
			(0.071)	
$\mathbf{1}_{\{age=8\}}$×Ln(GDP real detrended)			-0.066**	
			(0.027)	
$\mathbf{1}_{\{age=9\}}$×Ln(GDP real detrended)			-0.142***	
			(0.051)	
$\mathbf{1}_{\{age=10\}}$×Ln(GDP real detrended)			-0.036	
			(0.026)	
$\mathbf{1}_{\{\textbf{Ln(GDP real detrended)}>0\}}$				0.001***
				(0.000)
R-squared	0.041	0.041	0.041	0.041
Fixed-Effects?	Yes	Yes	Yes	Yes

Table C.6: Labor Productivity and Business Cycle at Entry: Alternative Indicators

This table replicate the specification of column (1) of Table 8 using alternative indicators of business cycle conditions. The dependent variable is *Ln(LP)* which is defined as the natural logarithm of the ratio between revenues (defined as value of shipments, sales, receipts, or revenue (in thousands)) and the number of employees at the establishment-level. All other variables are defined as in the main tables. Other controls include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, and Ownership Type Fixed-Effects. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(LP)	(1)	(2)	(3)	(4)	(5)	(6)
Ln(GDP nom. detrended)	-0.197*					
	(0.099)					
$\ln(\tilde{\text{GDP}}\text{ Nom.})_{st}$		-0.364***				
		(0.102)				
$\ln(\tilde{\text{GDP}}\text{ Nom.})_{it}$			-0.500***			
			(0.054)			
$\Delta\ln(\text{GDP Real})_t$				-0.261**		
				(0.108)		
$\Delta ln(Emp)_t$					-0.485***	
					(0.092)	
$\Delta\ln(\text{PIpc})_t$						-0.265***
						(0.081)
R-squared	0.600	0.600	0.597	0.600	0.600	0.600
Fixed-Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	SU only	SU only	SU only	SU only	SU only	SU only
Employment-Weighted?	No	No	No	No	No	No
Sample	2	2	2	2	2	2

Table C.7: Labor Productivity and Business Cycle at Entry: Alternative Selection of Cohorts and Firm Evidence

The table reports the coefficients of OLS regressions. The dependent variable is *Ln(LP)* which is defined as the natural logarithm of the ratio between revenues (defined as value of shipments, sales, receipts, or revenue (in thousands)) and the number of employees at the firm-level. All other variables are defined as in the main tables Other controls include Age Fixed-Effects, Year Fixed-Effects, State Fixed-Effects, Industry Fixed-Effects, and Ownership Type Fixed-Effects. The firm-observations in column (1) are not weighted by the number of employees, whereas the firm-observations presented in column (2) are weighted by the number of employees. Standard errors are presented in parentheses, and are clustered at state-level. ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

Dep. var: Ln(LP)	(1)	(2)
Ln(GDP real detrended)	-0.285***	-0.650
	(0.067)	(0.548)
R-squared	0.597	0.421
Fixed-Effects?	Yes	Yes
Single-/Multi-Unit Fixed-Effects?	No	No
Employment-Weighted?	No	Yes
Sample	2	2

Figure C.1: Kaplan-Meier survival functions by entry conditions

This figure plot the survival functions when I aggregate the establishments in *sample 1.3* between two groups based on the the log deviations of real GDP from its trend (i.e. *Ln(GDP real detrended)*). The red line plots the estimated survival for businesses born in years in which the economy grew above the trend, and the blue line in years in which the economy grew below the trend.

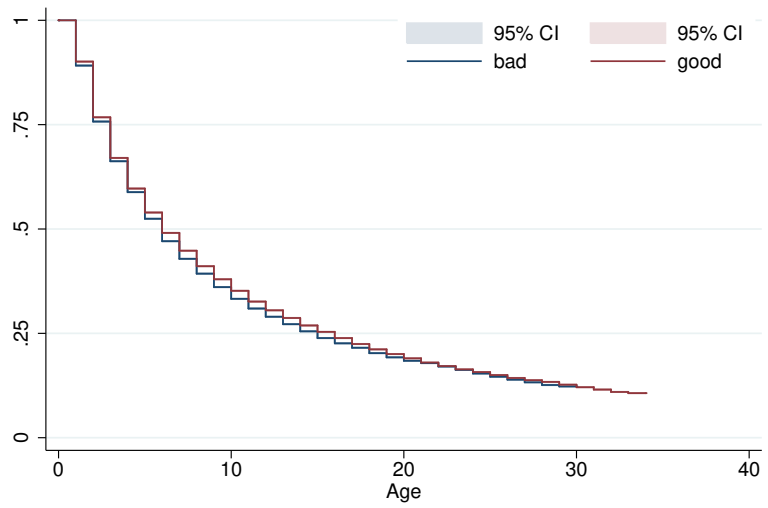
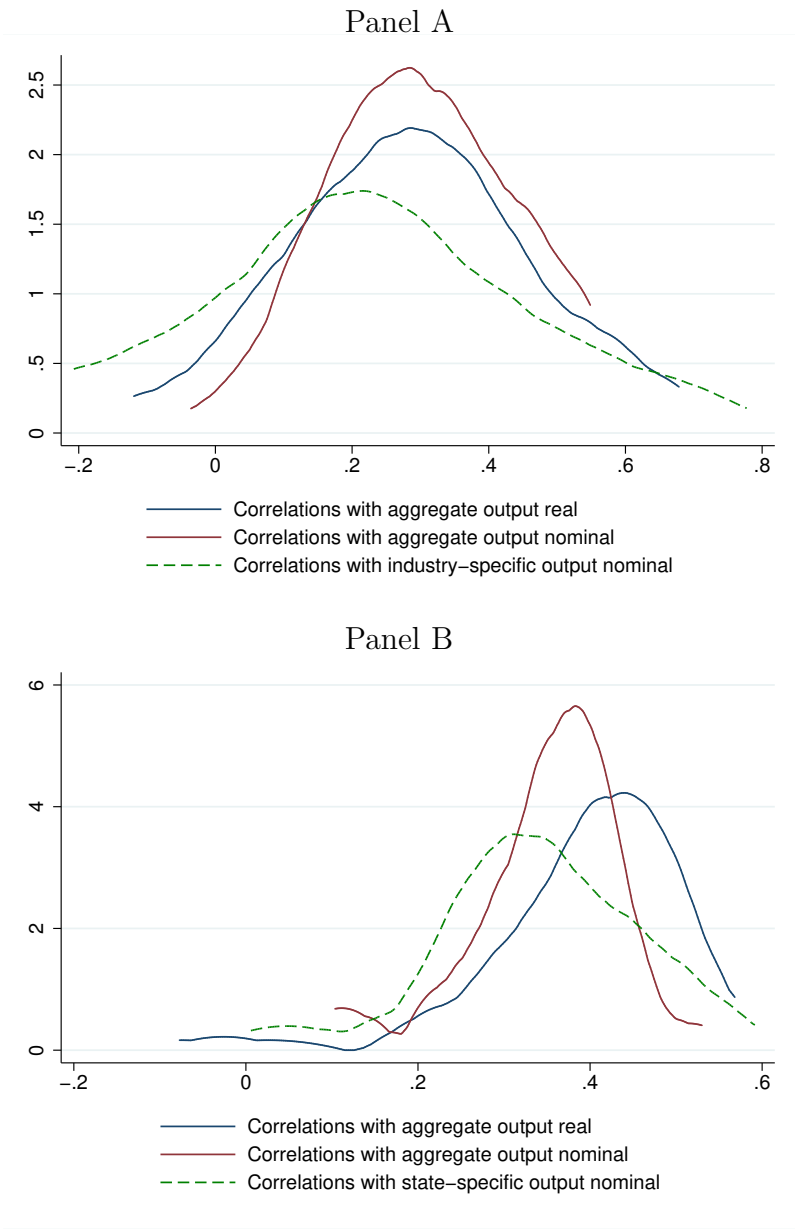


Figure C.2: Cyclicality of sectoral and geographic entry

These figures plot the kernel densities of the correlation between $\Delta E_{k,c}$ and the demeaned log differences in real/nominal output (aggregated and industry-specific), for k defined by the sector (Panel A) or k defined based on the state (Panel B). $\Delta E_{k,c}$ is defined as the demeaned difference in logarithms of the number of firms entering in each segment bin during a certain period. In Panel A, a segment bin is a group of firms that share the same 4-digit NAICS. In Panel B, a segment is defined by a U.S. state.