Fewer but Better: How Superstar Firms Affect New Firm Creation

Jack Liebersohn Victor Lyonnet*

Abstract

This paper offers a novel explanation for the decline in new business formation: the rise of the largest and most productive "superstar firms." We show that new firm creation has decreased in industries experiencing larger increases in the market share of superstar firms. The rise of superstar firms has discouraged low-ability entrepreneurs from starting a business, but not high-ability ones, resulting in a higher average quality of new firms. Superstar firms have also displaced low-productivity incumbent firms from the market. The rise of superstar firms is linked to an increase in the productivity gap between them and the rest of the economy.

*We thank Evan Dudley, Kristoph Kleiner, Yueran Ma, Isabelle Méjean, Alexandre Moulin, Andrei Shleifer, Horng Chern Wong, and conference participants at UC Irvine, UC Santa Barbara, the Ohio State University, the TFPI, for helpful comments and suggestions. We are grateful to Boyu Wang and Daisy Wang for excellent research assistance. All errors are our own. Liebersohn is with UC Irvine (Email: cjlieber@uci.edu) and Lyonnet is with the Ohio State University (Email: lyonnet.1@osu.edu).

1 Introduction

The pace of new business formation in developed economies has been on a downward trend since the 1980s.[1] As young firms are essential for job creation and productivity growth, this decline is often regarded as a sign of reduced economic dynamism. However, it is not just the *quantity* of new businesses that matters for economic dynamism, but also their *quality*. To understand the consequences of this decline, it is crucial to uncover its causes and whether new firms of different quality have been affected equally.

We propose an explanation for the decline in new business formation that is based on the rise of the largest and most productive "superstar firms" (Autor et al., 2020). We propose a simple model of superstar firms and firm entry and show support for our model's predictions using a novel administrative dataset on the universe of French firms. First, we find that new firm creation has decreased in industries experiencing a larger rise of superstar firms – i.e., industries with larger increases in product market concentration. Second, the rise of superstar firms has discouraged low-ability entrepreneurs from starting a business, but not high-ability ones, resulting in a higher average quality of new businesses. Third, superstar firms have displaced low-productivity incumbent firms from the market.

We start by establishing stylized facts about the long-term evolution of entrepreneurship in France. As in the United States and other developed economies, we find that the number of new firms and the share of aggregate employment accounted for by young firms have fallen over our 1994-2017 sample period. We also document that the contribution of young firms to aggregate sales and value-added have declined by roughly half during this period. A second set of stylized facts documents the rise of superstar firms in France. We provide evidence of a steadily growing market share of superstar firms starting around the year 1996. This growing concentration does not depend on whether we focus on the market share of the largest 20, 8, or 4 firms per industry, or the industry's Herfindahl-Hirschman index (HHI).

We formalize the link between the rise of superstar firms and business formation in a simple model based on Autor et al. (2020). The model studies the effects of a shift in the distribution of firm productivity that benefits the most productive firms relatively more ("fatter" tail) in a standard Melitz model. The model generates three key predictions for the effects of rising superstar

[1] The decline in entrepreneurship has been documented both for the United States (e.g., Decker et al., 2014; Gourio et al., 2014; Hathaway and Litan, 2014a,b; Decker et al., 2016a,b; Akcigit and Ates, 2021), and internationally (e.g., Criscuolo, Gal, and Menon, 2014; Bijnens and Konings, 2020). For discussions of the general decline in economic dynamism, see e.g. Decker et al. (2014, 2016a,b); Pugsley, Sedlacek, and Sterk (2019); Akcigit and Ates (2021).

firm productivity: a decrease in the number of new firms being created, an increase in the *ex ante* quality of new entrepreneurs, and an exit of the least productive incumbent firms. Our analysis tests these predictions using long-difference specifications in the cross-section of French industries. Our main explanatory variable measures the increase in market concentration due to superstar firms as the change from 1994 to 2015 in the market share of the largest 4, 8, and 20 firms per industry at the 3-digit level.

Our main result is that the quantity of new businesses has decreased in industries with larger increases in market concentration. We show this in several ways. First, we regress the changes in the long-term share of superstar firms on the long-term rate of new firm creation from 1994 to 2015. We find that a one percentage point increase in the share of the largest 20 firms implies a 2.7% decrease in the long-term rate of new firm creation. This coefficient is virtually the same when we use the market share of the largest 8 or 4 firms in each industry, and the results remain after controlling for sector-specific fixed effects, industry growth rates, and the initial level of value-added of each industry. They are also robust to measuring concentration using the share of the largest firms in aggregate value added or employment excluding firms created after 1994. Second, we find qualitatively and quantitatively similar results when studying 10-year changes in the share of superstar firms and new firm creation. Third, we show that the share of total sales coming from young firms is lower in industries with larger increases in market concentration. To the best of our knowledge, this evidence is the first to relate the decline in business formation to the rise of superstar firms.

Although our main result explains the aggregate pattern of declining business dynamism, it masks some heterogeneity over time and across industries. Over time, our analysis reveals that the concentration changes between 1994 and 2008 have the most significant impact on the long-term decline in new firm creations, while changes after 2008 do not significantly affect new firm creations. These findings suggest that the consequences of changes in concentration take time to manifest, implying a lag between the initial shift and its influence on new businesses. Across industries, we observe that the number of new firm creations has increased in certain industries where the market share of superstar firms has decreased. This heterogeneity echoes the evidence from the US that the decline in business dynamism did not affect all sectors of the economy in the same manner (Decker et al., 2016b).

Although the number of new businesses has declined, our research highlights a concurrent rise in the quality of these new firms. Using administrative survey data on entrepreneur characteristics, we explore various proxies for the ex-ante quality of new firms. Our findings imply that a one percentage point increase in the market share of the largest 20 firms is associated with a 26 p.p increase in the share of entrepreneurs who previously held executive or CEO positions, an 11 p.p. decrease in the share of entrepreneurs with prior blue-collar employment, and a 37 p.p. increase in the share of innovative entrepreneurs. These results imply that the rise of superstar firms has discouraged low-ability entrepreneurs from starting a business, but not high-ability ones, resulting in a higher average quality of new firms.

A firm of a given ex-ante quality is likely to perform better before than after the rise of superstar firms, so changes in ex-post performance might not capture changes in the average quality of new firms. Therefore, our preferred proxies for new firm quality use ex-ante characteristics of entrepreneurs. However, we do two things to build confidence in our ex-ante measures of new firm quality. First, we show that these measures are good predictors of ex-post firm performance. Second, we show that the relationship between superstar firms' market share and new firm quality holds even when quality is measured using ex-post measures.

Finally, we show that the increase in concentration did not only cause the decline in new firm creations; it also displaced incumbent firms from the market. We find that a percentage point increase in the market share of the largest four firms is associated with a 25 p.p decrease in the survival probability of (employment-weighted) incumbent firms. This excess of firm exits is driven by low-productivity firms, as predicted by the model.

Our evidence on the increase in the market share of superstar firms is consistent with a "winner takes most" mechanism enabling superstar firms to benefit from scale advantages. Around the year 2000, superstar firms experienced an increase in both their value added per worker and total factor productivity, whereas other firms did not show any significant change in their productivity. We underscore that several causes can explain this increase in the productivity gap between superstar firms and other firms: network effects, the toughness of product market competition, the development of IT and the internet (Autor et al., 2020; Kwon, Ma, and Zimmermann, 2021; Hsieh and Rossi-Hansberg, 2021). Our paper does not aim to explain the emergence of superstar firms. Instead, we accept this phenomenon as a given and show that it can explain the observed decrease in new firm creation. We support this argument by presenting new empirical evidence on the evolution of entrepreneurship.

We document a long-term increase in the quality of new businesses in France, which we reconcile with the decline in the number of new businesses and their share of economic activity. The apparent contradiction between those facts has led researchers to reach different conclusions on the state of entrepreneurship in the United States. When looking at the changes in the quantity of new businesses (e.g., Decker et al., 2014; Gourio et al., 2014; Decker et al., 2016a; Karahan, Pugsley, and Şahin, 2019; Bijnens and Konings, 2020) or in the contribution of young firms to aggregate employment (e.g., Decker et al., 2014, 2016b; Furman and Orszag, 2018; Pugsley, Sedlacek, and Sterk, 2019), the evolution of entrepreneurship seems coherent with a decline in business dynamism. However, when capturing the heterogeneity in growth aspirations among entrepreneurs (Schoar, 2010; Pugsley and Hurst, 2011) by measuring the quality of new businesses (Guzman and Stern, 2016; Fazio et al., 2016; Guzman and Stern, 2020), a more optimistic picture of entrepreneurship emerges because the average quality of new businesses has increased. We show that are all happening in France at the same time. In particular, we provide evidence that while the skewness of the distribution of employment growth rates for young firms is *decreasing*, the skewness of their sales growth rate is *increasing*. These findings suggest that high-quality new firms have not disappeared but that their business model has evolved towards less labor-intensive processes so that they contribute less to job creation. This interpretation is consistent with Barkai and Panageas (2021), who focus on publicly listed U.S. firms and show that young public firms' share of aggregate output has not significantly declined from 1985 to 2014.

Our paper also relates to the literature on the increase in market concentration and the rise of superstar firms. This literature has studied the implications of the rise of superstar firms for the labor share (Autor et al., 2020), the average markup of price over marginal cost (Decker et al., 2018), and competition (Philippon, 2019; Baker, 2019). Given our findings on the increased productivity of superstar firms around the year 2000, our analysis suggests that the decline in business formation is a sign of economic dynamism – not sclerosis. This interpretation aligns with Lashkari, Bauer, and Boussard (2018), who argue that falling IT prices explain around half of the changes in concentration in France. This timing is also consistent with recent papers that link the rise of superstar firms to improvements in technology (Hsieh and Rossi-Hansberg, 2019; De Loecker, Eeckhout, and Mongey, 2021), in particular IT (Autor et al., 2020; Bessen, 2020; Kwon, Ma, and Zimmermann, 2021), and international evidence that the decline in business formation is driven by the most IT-intensive industries (Bijnens and Konings, 2020).

This paper contributes to the literature on the causes of the decline in new business formation. Recent studies have suggested that this decline may be driven by the slowdown in population growth (Hopenhayn, Neira, and Singhania, 2018; Karahan, Pugsley, and Şahin, 2019; Engbom, 2019; Bornstein, 2018) or skill-biased technical change (Salgado, 2020; Kozeniauskas, 2018; Jiang and Sohail, 2022). These factors are compatible with our proposed explanation and we verify that they do not explain our findings. First, we show that our main result is robust to controlling for the ex-ante age distribution of entrepreneurs. Second, we find no significant relationship between the ex-ante education level of entrepreneurs and changes in business formation in the cross-section of industries. Our main results are also robust to controlling for a broader set of variables measuring the ex-ante skill composition of entrepreneurs in each industry. Overall, these results suggest that the rise of superstar firms explains the decline in new business formation over and beyond existing theories. Our evidence also suggests that explanations that cannot account for the disappearance of low-ability entrepreneurs cannot fully explain the decline in new business formation.

2 Framework

To provide intuition for why the increase in superstar firms' productivity affects new business formation, we consider a standard Melitz (2003) model and study the effects of greater superstar productivity on the entry of new firms. We build on the model in Autor et al. (2020) which studies the effects of superstars on the labor share of output. This section describes the model's main predictions and Appendix B develops the framework in full.

There is a continuum of firms, each choosing to produce a different variety ω by paying a fixed cost f and a marginal cost $1/\varphi(\omega)$ to produce goods, where productivity φ is the realization of a random variable. Labor is the sole factor of production with wages normalized to unity. The firm's profit is

$$\pi(\omega) = p(\omega)q(\omega) - \left(f + \frac{q(\omega)}{\varphi(\omega)}\right), \tag{1}$$

where $p(\omega)$ is the price, and $q(\omega)$ the demand, for each product variety ω. Appendix B shows that in equilibrium, there exists a unique productivity level φ^* such that firms only stay in the market if $\varphi > \varphi^*$. Free entry requires the total expected value of profits to be equal to the fixed cost of entry, which together with the zero cutoff profit condition $\pi(\varphi^*) = 0$, pins down the value of φ^* in equilibrium.

We make two simplifying assumptions. First, we assume a standard CES utility function so that markups are the same across all firms regardless of size and productivity:

$$U = \left[\int_{\omega\in\Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where σ is the elasticity of substitution between varieties and Ω is the mass of available goods. Second, we make the standard assumption that the productivity distribution follows a Pareto distribution (as in Melitz and Ottaviano, 2008):

$$G(\varphi) \equiv 1 - \left(\frac{\underline{\varphi}}{\varphi}\right)^{\theta} \text{ for } \varphi \geq \underline{\varphi}. \tag{3}$$

Making these assumptions allows us to derive closed form solutions to clearly show the effects of an increase in superstar firms' productivity and that our results do not depend on changes in markups across firms or in the time series. Furthermore, the Pareto distribution has been shown to be a reasonably good approximation to capture the upper tail of firms' productivity in the data (e.g., Del Gatto, Ottaviano, and Mion, 2006), which is the focus of this paper.

We model the increase in the productivity of superstar firms as a decrease in θ, which is the parameter governing the shape of the tail of the Pareto distribution in (3). Such a change implies a shift in the distribution of firm productivity that benefits the most productive firms relatively more ("fatter" tail). The literature has studied various reasons for the occurrence of a "winner takes most" mechanism whereby economic activity reallocates towards the most productive firms: consumers who become more sensitive to quality-adjusted prices (e.g., through greater competition due to globalization or improved search technologies), the growth of platform competition, or scale advantages related to the growth of intangible capital and advances in information technology.

We derive four predictions of the effects an increase in the productivity of superstar firms in Appendix B. This paper's contribution is to show support for these predictions in the data.

Prediction 1. *An increase in the productivity of superstar firms leads to a decrease in the number of new firms.*

Prediction 1 proposes an explanation for the decline in business formation. Intuitively, an increase in the productivity of superstar firms increases the productivity threshold φ^* below which firms do not enter the market. Therefore, less firms find it profitable to pay the fixed cost f so that a smaller number of firms enter the market, and conditionally on paying this fixed cost, a smaller fraction of entrants survive. As a result, the number of new firms decreases. This effect is illustrated in Appendix Figure B3.

Prediction 2. *An increase in the productivity of superstar firms leads to an increase in the average productivity of new firms.*

An increase in the productivity of superstar firms increases the productivity threshold φ^* below which firms do not enter the market. Therefore, the productivity of entrants that survive increases, so that the average productivity of new firms increases.

Prediction 3. *An increase in the productivity of superstar firms leads to an exit of the least productive incumbent firms.*

The productivity of the least productive incumbent firms is just above the threshold φ^*. When the productivity of superstar firms increases, the threshold φ^* increases so that the least productive incumbent firms are no longer above this new threshold value and they exit the market.

Prediction 4. *An increase in the productivity of superstar firms leads to an increase in market concentration.*

Expenditures per variety, hence per firm, increase with firm productivity. Following an increase in the productivity of superstar firms, the average productivity of surviving firms increases, so that expenditures per variety are more concentrated. Whether measured using the Herfindhal index (HHI), which is the sum of squared market shares of all firms, or the share of the largest (superstar) firms' market share, market concentration increases following an increase in the productivity of superstar firms.

3 Data and stylized facts

3.1 Empirical design

We test our model's prediction of the effects of industry concentration on new business formation.

Our empirical approach consists of long-difference estimates that compare the characteristics of new firms in different industries. Our main independent variable is the change in concentration of the largest firms in each industry, which we refer to as superstars. We study changes in this variable over the sample period we can measure, 1994-2015. Our dependent variables are changes in industry-level aggregates. As such, our main regression specification is:

$$\Delta Y_i = \Delta Concentration_i + \varepsilon_i \tag{4}$$

where $\Delta Concentration_i$ measures the change in concentration of the largest firms in each industry and ΔY_i is one of several dependent variables. We measure concentration using sales, focusing on the sales share of the top 20, 8, and 4 largest firms; and on the log of the sales HHI by industry.

We study independent variables from each of our data sources described below. We make two important assumptions here. First, our industry groupings are all constructed on a *national* basis. In some industries, it may be more appropriate to group firms by region or even by neighborhood — for example, the appropriate measure of concentration may be highly local in the restaurants or retail sectors. Insofar as we group firms at the wrong level, this will bias our estimates towards not finding an effect. Second, by considering industry-level differences, we ignore spillovers that occur across industries. Superstar firms may produce products in a variety of industries, especially as they expand. We will not be able to measure this effect.

We study three main groups of independent variables, which we organize based on the predictions of our hypothesis.

New firm formation Our measure of firm creation is the log number of new firms per year. We only consider incorporated (i.e., non-owner-operator) firms. We measure the change in log creations by industry from 1994-2015. Since this is our most important outcome variable, we show that the results are robust to measuring new firm formation in a variety of ways, including firm formation *rates* and share of output. We also measure the relationship between firm formation and concentration in ten-year differences instead of long differences.

A possible concern with our findings is that the link between firm formation and concentration is mechanical. If there is an "exogenous" increase in the number of new firms in an industry (for example, due to new tax incentives), then the new firms could decrease industry concentration, leading to a spurious link between concentration and firm formation. To ensure that our results are not affected by this type of reverse-causality, we repeat our ten-year differences specification, but measure concentration changes using *only firms that have existed for the past ten years.* We show that the results hardly change when measuring concentration in this way.

Incumbent firms exits We measure the number of firm exits at the industry level as the number of firms in a given industry that stop reporting to the tax authorities and thus disappear from the tax files data.

New firm quality We proxy for new businesses' quality using entrepreneurs' education and past employment. Industries with the largest increase in superstar productivity should also have the greatest increase in the quality of new entrants.

3.2 Data sources

We use four sources of French administrative data provided by the French Statistical Office (INSEE): the exhaustive firm registry, accounting data from the tax files, employment data from employer payrolls, and a survey of entrepreneurs conducted every four years. Firms are uniquely identified by a 9-digit code (SIREN) that allows us to merge the different databases together. We focus on incorporated firms, dropping sole-proprieterships and non-employer startups allowed under special legislation (the so-called "auto-entrepreneurs").

To measure firm creation rates, we use data exclusively from employer firms, which are defined as firms having a minimum of one employee either at the time of creation or within their first year of operation. Bento and Restuccia (2019) show that nonemployers account for 82% of all US firms in 2014. We find very similar numbers in France. When we calculate statistics for the entire French economy (in particular, to calculate industry-wide concentration), we include both employer and non-employer firms. However, because non-employer firms make up a small fraction of total sales in all industries, our calculations do not change substantially if we drop firms with no employees.

Firm registry. The firm registry (*SIRENE*) contains the universe of firms registered in France from 1998 to 2017. For each newly created firm, the registry contains the industry the firm operates in based on a four-digit classification system similar to the four-digit SIC. It also provides the firm's legal status (e.g., Sole Proprietorship, Limited Liability Corporation, Corporation), the official creation date and geographical location.

Accounting data. Accounting data (balance sheet and income statements) is extracted from the tax files used by the Ministry of Finance for corporate tax collection purposes. The accounting information is therefore available for all French firms, public or private, whose annual sales exceed €32,600 (€81,500 in retail and wholesale trade).[2] We retrieve total sales and value added from the tax files. Industries are defined using the French classification (*Nomenclature des activités Francaises, NAF*). We use the 3-digit industry level and end up with 162 different industries given our sample construction.

To ensure a consistent industry code during our sample period, we identify firms that exist both before and after the 2001 change from NAF1 to NAF2 industry classification change to calculate the

[2]Small firms with annual sales below this threshold can opt out and choose a special micro-business tax regime (*micro-entreprise*). These firms are not growth oriented. Income falling into this category is taxed at the individual level, hence they do not appear in the corporate tax files (Aghion et al., 2017). We exclude from our sample firms in the financial, agricultural and public sectors as they use different accounting systems.

fraction of firms in each NAF1 sector that belong to each NAF2 sector. For this sample, we keep the NAF1 codes for the firms' entire existence. For firms that are newly created after the NAF2 switch, we use two methods to create a panel of firms at the NAF1 level. When we calculate aggregate statistics, we use the calculated probabilities to allocate newly-created firms in each NAF2 sector to the corresponding NAF1 sector. For individual-level estimates, such as those using the entrepreneur survey (SINE) described below, we assign each firm the most likely NAF1 sector (i.e., the sector with the highest probability).

Employer payrolls We use the French matched employer-employee dataset (*Déclarations Annuelles des Données Sociales*, DADS) to observe firms' employment. All firms that employ at least one employee must file payroll taxes. We use the DADS data to identify firm survival so that startups that never have any employees do not "survive" even in their first year of existence.

Table 1 Panel A displays the summary statistics for our variables of interest, both in level and in long differences, arranged by industry level.

Entrepreneur survey. Our last source is the Système d'Information des Nouvelles Entreprises (SINE), which is a large-scale survey of entrepreneurs in France conducted by the French Bureau of Statistics every four years, from 1998 to 2014.[3] The main two advantages of these data is that they are not subject to any selection biases commonly encountered in the literature and that we are able to observe a large set of startups' founder characteristics. Questionnaires are sent to approximately 25% of entrepreneurs who started or took over a business in France that year. The surveyed firms are randomly selected from the exhaustive firm registry such that they are representative of the population of new businesses. The business owner is responsible for completing the documents. The response rate to the SINE survey is high (approximately 90%) because the tax authorities supervise the sending of questionnaires.

We focus on entrepreneurs who create a new startup by filtering out those who takeover an existing business (through purchase or inheritance, for instance). We exclude startups in the financial, agricultural, and public sectors from the sample. Finally, as mentioned above, we require new firms to have a least one employee in their first two years of existence. We obtain a representative sample of 37,269 new firms from the survey cohorts of 1998, 2002, 2006, 2010, and 2014. A few years after their inception, firms are resent similar questionnaires but we only focus on the initial

[3] These data have been used in existing papers (e.g., Landier and Thesmar, 2008; Hombert et al., 2020; Lyonnet and Stern, 2022). The data are available through INSEE (click here).

survey. This survey contains information on the entrepreneur's main sociodemographic characteristics, education, experience, the reasons and motivations for which the firm was started and the conditions under which it was started (e.g., financing, initial research, customer prospects).

Table 1 Panel B contains summary statistics for the variables we use in our analysis. Note that the variables on firms' employment and sales are obtained from the accounting and payroll datasets described above, matched using the unique firm identifier. Appendix A1 contains the description of these variables. In our sample, 36% of the entrepreneurs have a college degree (at least two years). 24% were working as executive or CEO, before creating their own firm. Finally, 37% are "serial" entrepreneurs, i.e., have previously founded a business.

With these data, it is possible to distinguish between true "startups," new establishments of existing businesses, and "new" firms formed by combining pre-existing establishments through merger and acquisition activity.

3.3 Stylized facts

3.3.1 Evolution of business formation and quality

It is well documented that the firm entry rate has decreased in the US (see, e.g., Decker et al., 2014; Akcigit and Ates, 2021). Does a similar picture emerge in France? To answer this question, we investigate the evolution of the rate of new business formation in France since the 1990s, we study the total number of new businesses and the share of employment and value added accounted for by young firms from 1994 to 2019.

Figure 1 plots the total number of new businesses created every year from 1987 to 2019. The green line shows that this number was on a downward trend from 1987 to 2002, then increased around 2002 (consistent with the evidence in Hombert et al., 2020),[4] and plummeted during the financial crisis of 2008. Afterwards, the total number of new businesses has continuously increased from 2012 to 2019, when it reached its all-time peak.

Comparing the orange and green lines, the first thing to notice is that most new startups are sole proprietorships without a single employee on the payroll. Therefore, the evolution of the total number of startups can be misleading. In line with the growing literature arguing that self-employment is a poor proxy of entrepreneurship,[5] we show the number of new businesses created

[4] Hombert et al. (2020) show that this increase was due to a reform of the French unemployment insurance system that provided downside insurance to unemployed workers starting a business.

[5] Pugsley and Hurst (2011) document that most self-employed workers have no intention to grow or innovate. Haltiwanger, Jarmin, and Miranda (2013) show that after controlling for firm age, small businesses do not create jobs.

with at least one employee in the orange line of Figure 1. Similar to the US, we now find that the number of startups has steadily been declining since 1989. To the best of our knowledge, we are the first ones to document the decline in business formation in France.

A consequence of the declining startup rate is that the share of young firms in the economy, and the share of activity for which they account, is declining. Figure 2 shows that while 30% of firms were aged three years or less in 1994, this fraction fell to about 20% by 2015. This long-term continuous decline in the fraction of startups in the economy translated into a decrease in the share of employment accounted for by these young firms. This share fell from almost 15% in 1994 to around 5% in 2015.

We then turn to the investigation of the evolution of startups' quality in France. Leveraging the detailed information available in the SINE survey, we derive three proxies for entrepreneur ex-ante ability or "skills" based on their education and past employment. Panel A of Figure 3 plots the evolution of the share of entrepreneurs in each cohort from 1994 to 2014, that hold a college (at least two-year) degree. A rise in the share of educated entrepreneurs is evident throughout the whole sample period: from 25% in the 1994 cohort to 45% in the 2014 cohort. Panel B of Figure 3 presents the share of entrepreneurs previously employed as executive or CEO. It also highlights an upward trend: from 17% in the 1994 cohort to 27% in the 2014 cohort. Finally, Panel C of 3 shows the fraction of serial entrepreneurs in each cohort, i.e., entrepreneurs who had already created a business in the past. It is around 25% in 1994 and increases up to 45% in the 2002 cohort. It then remains around 35% in the last three cohorts. Overall, the pattern depicted by Figure 3 is that entrepreneurs have become more educated and skilled on average over time.

3.3.2 The rise of superstar firms

What are the reasons of such a decline in entrepreneurship in France? One explanation might be that very large firms (superstars) have expanded at the expense of younger and smaller firms. To investigate this possibility, Figure 4 (orange line) plots the evolution of the average productivity of the top 10 largest firms in each industry each year from 1994 to 2015. We construct productivity as the average of log sales per employee over all firms in each industry. Figure 4 shows that there is a clear upward trend in the productivity of top firms over this period: we estimate that the productivity of superstars has increased by about 50% across industries between 1994 and 2015.

Henrekson and Sanandaji (2014) find that the rate of self-made billionaires correlates negatively with self-employment rates. Schoar (2010) discusses the need to differentiate subsistence and transformational entrepreneurs. Levine and Rubinstein (2017) argue that incorporation is a better proxy of US entrepreneurship than self-employment.

At the same time, we do not observe such a trend in the evolution of the productivity of other firms in the economy. Indeed, Figure 4 (blue line) also shows how the average productivity of all firms has evolved over the same period. It stands out that the average productivity of all firms in 2015 is pretty much the same as in 1994, with a relatively flat evolution. A similar pattern emerges when we focus on young firms. The green line on Figure 4 presents the evolution of the average productivity of young firms (age between 1 and 3). As for all firms, we do not observe a clear change in the productivity of young firms between 1994 and 2015.

Table 2 demonstrates the considerable heterogeneity in the emergence of superstar firms across various sectors. Column 3 highlights the change in the share of sales attributed to the Top 20 firms between 1994 and 2015 at the sector level (using the naf1 1-digit classification). It is evident that the Top 20 share exhibits significant variations among sectors. For example, the Textile and Clothing sector has experienced a 26% increase in the Top 20 share, while the Hotels and Restaurants sector has seen an 8% decrease. We leverage these cross-sectional differences at the industry level in our empirical analysis presented in section 4.

3.4 Case study: The auto dealership industry

Automobiles in France are sold by two types of retailers: Independent dealers and retail subsidiaries owned by auto manufacturers. Historically, both groups were of roughly equal importance. An article in the trade publication *Automotive News* reported that 57% of sales were from independent dealers in 1995.[6] Of these, only 6.5% were owned by large distribution groups that owned a large number of dealers. Most other dealers were independently owned small businesses, many run by families with multiple generations of ownership. These small, family-run businesses were not efficient by international standards: According to *Automotive News*, retail outlets in France averaged only 93 sales per retailer in 1995, well below the American or British sales at the time.

Beginning in the early 1990s, the French automobile retail sector underwent substantial consolidation and reorganization. Consolidation took place in two ways. On the one hand, auto manufacturers reduced their number of retail outlets in order to increase branches per location. On the other hand, many independent dealers either shut down or were bought by centralized distribution networks. In many cases, this took the form of family firms not being continued by the next generation. At the same time, fewer individual entrepreneurs have started new auto dealerships.

The consolidation in the French auto retailing has resulted in greater concentration and a larger

[6] https://www.autonews.com/article/19960916/SUB/609160805/france-many-factory-dealerships

share of sales from the very largest retailers, i.e., manufacturer subsidiaries. According to French tax data, in 1994, the top 20 auto retailers accounted for about 21% of sales and the top 50 accounted for 25%. By 2016, these shares had risen to 30% and 35% respectively. Increasing concentration has been accompanied by a decrease in firm creations: Over 800 new auto dealers were registered with the French government in 1994. By 2016, the number had fallen to about 300, despite an increase in the number of automobiles sold.

It is instructive to compare consolidation among French auto retailers to what has happened in the United States. Unlike in France, car manufacturers in the U.S. sell essentially no automobiles directly. Antitrust law is also different in both countries, with some research arguing that U.S. enforcement is relatively lax (Philippon, 2019). Despite these institutional differences, the U.S. has experienced a similar trend as France. Large regional auto dealer networks, many focused on a particular region and a few car brands, make up an increasing share of U.S. auto sales.[7] Statistics from the U.S. economic census show that the sales share of the top 20 auto retailers rose from 9.5% in 2002 to above 12% in 2017. The *Automotive News Research and Data Center* reports that the top 150 dealer networks owned 13.9% of dealers in 2011, but 22.7% in 2021.

What explains the rise in consolidation in France — and in the United States? French trade publications we read are vague about this, with many ascribing the rise in concentration to the better service that dealer networks are able to provide. However, industry participants we spoke to emphasized two factors. First, the role of changing business practices. Large dealer networks in France are able to hire specialized professionals in human resources, finance, accounting, etc. Centralizing business practices saves money and improves operating efficiency. Second, the role of changing technology. The rise of the internet enables consumers to shop around for the best deals, favoring firms that provide services most efficiently. In addition, increasingly complex technology (e.g., from electric cars) harms auto dealers who do not have sufficiently advanced tools or educated mechanics. Both trends favor large, professional distribution networks — and in the case of France, manufacturer-owned dealers — over small, independent retailers. While there are some concerns about consolidation leading to lower competition, industry insiders notably did not attribute these trends to changing regulation or changes to antitrust law.

We draw two lessons from the experience of French auto dealers. First, that the secular trends driving consolidation favor efficient firms with professional management over traditional, family-run

[7] See https://www.kbb.com/car-news/consolidation-fewer-companies-operate-car-dealerships-every-year/ and https://jalopnik.com/you-re-not-wrong-car-dealerships-are-getting-bigger-an-1848745399

businesses. Second, that the qualitative industry trends are similar in both the U.S. and France. This would favor economic explanations linked to changes in the underlying business environment or technology, rather than matters of local policy. Changes in technology or business practices that work to the advantage of the most efficient firms seem to be an important cause, but we also cannot reject a role for changing demographics or educational opportunities, which might explain why family firms do not continue into the next generation.

In section 4, we provide systematic quantitative evidence on how the decline in entrepreneurship and the increase in entrepreneurs' quality are related to the expansion of superstar firms and their productivity growth.

4 Declining Business Formation and the Rise of Superstar Firms

4.1 Main result

Our central prediction concerns the relationship between superstar concentration and firm creations. The theory implies that the growth of superstars comes at the expense of potential new entrepreneurs who decide not to start new firms in industries where superstars dominance has increased. We provide evidence for this result using two tests. First, we show directly that in industries where superstar productivity increased the most, entry by new firms has fallen. Second, we show that the share of output, value-added and sales coming from young firms is lower in these industries.

Figure 6 provides graphical evidence for our first result, that greater superstar productivity is associated with fewer new firms. The X axis of this figure shows the change in top-20 sales share by industry. The Y axis of this figure is the change in log number of new firms. As elsewhere, differences are calculated from 1994-2015 and the size of each industry is proportional to the number of firms in the base year. The figure shows a strong negative relationship which is confirmed by the best fit kind through the points.

Table 3 shows regression estimates relating superstar concentration to new business formations. The results confirm the relationship apparent in Figure 6. We measure concentration using the top-20 share, the top-8 share, the top-4 share and the log of the HHI. We note in this table that the coefficient and regression R^2 is larger when we measure the productivity of the top-20 firms, suggesting that broader measures of superstar productivity perform somewhat better. At the same time, the log of HHI has a somewhat lower R^2, suggesting that concentration among *superstars* matters, rather than industry-wide dynamics.

The effects of concentration across industries Table 6 examines the consequences of increasing concentration on firm creation within four specific industries: manufacturing, commerce and repair, property and business services, and transportation and communication. Across these industries, we consistently observe that a rise in concentration tends to result in a decline in new firm creation. The findings are not always statistically significant, which may be attributed to the reduced number of observations.

4.2 Time dynamics

Autor et al. (2020) shows that the effect of concentration on worker income share is present, using differences measured over five or ten years. Is the effect on entrepreneurship also present at shorter intervals? Our results show that this is the case.

Table 4 replicates our primary specification, where this time we investigate the effects of concentration on new firm creation over shorter time horizons. Columns 1-4 of Table 4 estimate ten-year differences. The coefficients are qualitatively and quantitatively similar whether we use ten-year changes or 21-year changes. Interestingly, 5-year changes in concentration appear to have a lesser impact. An additional benefit of using this specification is that we use more of the time series data, rather than being restricted to just using data from 1994 and 2015.

A second benefit of studying ten-year differences is that the estimates allow us to do a second important check of the results. As discussed above, a potential concern is that there is a mechanical link between concentration and new firm creations. Any sort of exogenous shock which increases entry into an industry will increase the output of small firms, and thus, will decrease the output share of superstars. A measure of concentration not subject to this critique instead uses the change in concentration *among existing firms*. To create this measure, we calculate the change in superstar output share from year $t-10$ to t excluding any firms that are created from years $t-10$ to t. The estimates are shown in columns 5-8 of Table 4. Why do we show these results for ten-year changes and not for long differences? In principle, it should be possible to measure long differences changes in concentration excluding newly created firms. The problem with doing this is that over such a long time period, many firms in an industry will have disappeared, so concentration measured excluding new firms and concentration including them is a very weak one. For this reason, we prefer to show these estimates using ten-year changes.

We continue our analysis in Table 8, inquiring which time periods' changes in concentration are responsible for the primary effect of long-term differences in concentration on new firm creation.

The effects of changes in concentration from 1994 to 2001 and 2001 to 2008, shown in the first and second panels of Table 8, are substantial and significant. In contrast, we do not find any significant effect of changes in concentration from 2008 to 2015 on new firm creation. Although concentration continued to rise after 2008, Appendix Table A2 reveals that this increase was not significantly correlated with changes in the number of newly established firms.

The estimates shown here raise the question of how quickly rising superstar growth decreases entrepreneurship. To understand this relationship, we estimate local projections of the change in entrepreneurship on concentration. The local projections consist of regression estimates of the following form. For j from 1 to 10, we estimate:

$$Log(Creations)_{t+j} = \beta_{1j} Concentration_{tj} + \beta_{2j} Log(Creations)_{tj} + \gamma X_{t-1j} + \gamma_2 X_{t-2j}$$

The local projection (LP) consists of the β_{1j} plotted as a function of j. Importantly, we include lagged values of concentration and firm creations in X_{t-1} and X_{t-2}, and have also estimated similar results using other lags and other control variables.

Estimates are shown in Figure 7. The estimates shown here show that concentration takes several years to have an effect on firm creation rates, indeed, as long as ten years depending on the estimates. In addition to providing new evidence on the time dynamics of concentration, the findings here provide further evidence against a mechanical link between concentration and entrepreneurship (in which case we would see all the results in year 0).

Finally, we ask whether the effects are similar in all time periods, or are stronger in some years than others. The time series summary statistics show that the rise in concentration was greatest in the late 1990s and early 2000s, in line with estimates of previous research. In Table 8, we show that the effects of concentration on new creations are also greatest in the early years. This table splits the sample into early years and late years and shows that the coefficients are substantially stronger in early years.

Overall, our findings suggest that the consequences of changes in concentration take time to manifest, implying a lag between the initial change and its influence on new businesses.

4.3 Alternative explanations

Until now, we have limited the results to bivariate regressions limited to show the relationship between concentration and superstars. In this section, we discuss possible alternative explanations.

One alternative explanation is that our estimates reflect sector-wide changes rather than changes specific to a particular industry. The fifth panel of Table 3 repeats our main specification but includes sector-specific fixed effects. In effect, these estimates control for the *sector* level changes in concentration and entrepreneurship and use only within-sector variation. The findings are still robustly statistically significant, although they are somewhat smaller.

A second set of alternative explanations is that the estimates reflect differences across industries in overall growth rates. For example, if an industry is shrinking for exogenous reasons, it may lead to a discouragement of new entrepreneurs and an increase in firm exits — both of which would raise concentration and lead to a spurious relationship between concentration and new firm creations. Panels 3 and 4 of Table 3 control for industry growth rates and industry initial value-added. Again, the effects remain robust to these controls, despite the relatively small number of observations in the sample.

Third, a potential concern is that the characteristics of startups are changing in industries with more concentration. If this is the case, perhaps new firms are responsible for a similar share of output, even though there are fewer of them. Thus, we measure the new firm share of industry sales in the last panel of Table 3. The results show that the effects are robust to using the new firm share of industry sales.

4.4 Creation rates

The literature on entrepreneurship has largely favored measuring firm creation *rates* rather than output. The literature disagrees somewhat on the correct measure of rates. The reason for the disagreement is that industries with a falling creation rate will ultimately have fewer firms. Since the number of firms is the denominator of the creation rate, an exogenous decrease in the benefits to entrepreneurship will show a smaller change in the rate than in the number of new firms that are being created.

Our model implies that there is a second reason to be concerned about measuring creation rates: If the same shock causes both a decrease in the number of entrepreneurs and an increase in firm destruction, then the number of firms in an industry will fall the most in industries where entrepreneurship is discouraged *due to the destructions channel*. These two critiques both mean that estimates using creation rates may be biased towards zero.

Our primary estimates are robust to these concerns because we measure firm creations using the log change in the number of firms in an industry. Table 7 shows that these concerns are warranted. In columns 1 and 2 of this table, we show the relationship between concentration and the change in creation rates from 1994-2015, measuring creation rates as $(Creations_t)/(Firms_t)$, i.e., measuring the change in annual rate of creation over the entire time period. The estimates show there is no statistically or economically significant relationship between the variables, and in fact is slightly positive. Columns 3-4 repeat the same specifications as 1-2, but use $Firms_{1994}$ as the denominator. Since this holds the number of firms in the industry fixed, it is not subject to the same critique.

5 Other Implications of the Rise of Superstar Firms

5.1 New firms' quality

The relation between firm's ex-ante quality and the rise of superstar Given the finding in Section 4 that increased superstar productivity explains the reduction in the *quantity* of businesses created, we now analyse whether changes in superstar productivity also had an impact on the *quality* of new businesses. We present in Table 8 the results of similar industry-level regression specifications estimating the relationship between the change in entrepreneur quality and the change in productivity of the largest (superstar) firms.

We use three measures of CEO quality. First, we measure education. The rise in concentration increases the likelihood that an entrepreneur has a high school degree (note, the effect on college is not statistically significant, but both are positive). Second, we show that they affect the background of entrepreneurs. Industries with more concentration have more entrepreneurs with a CEO/executive background and fewer with a blue collar background. Finally, there is a large and highly significant effect on the likelihood that the firm will favor innovation. We do not estimate a significant impact on the likelihood that creators are a serial entrepreneur.

We also show the effect on entrepreneurs' age. We estimate no significant impact. This suggests that the estimates are not driven by a mechanical relationship between quality and executive demographics, as we would expect for example if it were related to an aging workforce.

Note that the number of observations in the regressions presented in Table 8 is slightly lower compared to Table 3. The reason is that some industries are not present in both the initial and final cohorts of the SINE survey, either because some industries appeared or disappeared from the sample or because no entrepreneur was surveyed in either of these years.

Most specifications detect a strong relationship between changes in ex-ante quality of entrepreneurs and changes in the productivity of the largest firms. The correlations are all positive and are statistically significant in all specifications except the ones in columns 5 and 9. The latter correspond to specifications regressing respectively the change in the share of entrepreneurs previously employed as executive or CEO on the change of productivity of the top 3 largest firms, and the change in the share of serial entrepreneurs on the change of productivity of the top 3 largest firms.

In terms of magnitude, we observe that a rise by one standard deviation in the change in top 5 firms productivity (0.73) is associated with an increase by 4.7 p.p. in the share of entrepreneurs with a college degree in the industry, by 3.1 p.p. in the share of entrepreneurs previously employed as executive or CEO, and by 2.9 p.p. in the share of serial entrepreneurs. Graphical evidence is provided in Figure 10. Overall, these results are consistent with the view that growing productivity of superstar firms has led to a rise in the quality of new entrepreneurs in the cross section of industries. Low-ability entrepreneurs are likely to be discouraged from starting a business in the specific industries where superstar firms have become more productive, resulting in a higher average ability of new entrepreneurs.

5.2 Exit of incumbent firms

In Table 10, we ask whether greater increases in superstar productivity lead to more firm exits. Again, we measure long-term changes in superstar productivity from 1994 to 2015 as the increase in the productivity of the top 3, 5, or 10 firms in each industry. The dependent variable is the cumulative fraction of incumbent firms surviving over the same period. Our firm survival measure is defined at the industry level as:

$$\frac{\text{Count of existing in both 1994 and 2015}}{\text{Count of firms existing in 1994}}$$

We also create an analogous survival rate variable limited to firms with above- and below-median productivity in 1994.[8]

Columns (1)–(2) show that increases in either measure of superstar concentration lead to a lower cumulative fraction of firm survival, i.e., a greater rate of exit. The estimates are substantial

[8] Results are very similar when using a variable that weights firms by their 1994 total employment.

and statistically significant one standard deviation industry-level increase in an industry's top-20 concentration growth, which is about 0.7 (from Table 1) is associated with a survival rate that is 2.5 percentage points lower. The results are similar for the top-20 and top-8 concentration. The results in these two columns are not statistically significant. However, columns (3) and (4) weight firms by their employment and show results that are both larger and statistically significant. In other words, greater concentration causes firms to disappear *with more employment*. Potentially, smaller firms such as sole proprietorships are not affected. Alternatively, we think that low-employment firms are poorly measured in the data.

6 Comparing our results with evidence from the United States

6.1 Main result using US data

In this section we test whether the link between concentration and firm entry holds in the United States as it does in France. An implication of our theory is that similar cross-industry patterns should hold in any setting where superstar firms are increasingly dominant. Therefore, an important test of our results is whether they hold outside the French context. Here we show that similar results hold in the cross-section of U.S. industries. Specifically, we show that in U.S. industries where concentration has increased by more, entry by new firms has decreased as has the share of employment in new firms.

The United States is a natural comparison country for two reasons. First, both the rise of superstars and the decrease in business formation have been exhaustively documented for the U.S. Therefore, we would like to show that our findings extend to a setting that is already well-understood. Second, the reasons for the rise of concentration in the United States are disputed, with some papers blaming lax antitrust enforcement (Philippon, 2019) and others emphasizing the superior productivity of superstars (Autor et al., 2020). That similar results hold in the U.S. as in France suggests that the findings are not due to any peculiarities of the French labor market or international differences in antitrust enforcement.

Our analysis follows, as closely as possible, the same empirical specification as our analysis of French industries. The main limitation to studying the United States is the lack of data. There are no systematic measures of superstar productivity for the United States, and detailed productivity measurement has mostly focused on the manufacturing sector. Instead of measuring productivity directly, we estimate the relationship between industry-level changes in concentration and changes in the number of new firms. Table 3 estimates the concentration-entry relationship for France. For the the U.S. estimates, we try to replicate that table as closely as possible. This makes our results similar to other recent papers that have measured concentration directly rather than trying to estimate superstar productivity.

There is only one public data source measuring concentration at the industry level for the United States: The Economic Census, which occurs every five years. Concentration data is only available from 2002-2017. Our concentration measures are the share of sales in each industry from the top 4, 8, 20 and 50 firms. We focus on these measures since they are the only ones available for all years. Data on firm creation is more easily available, through the Business Dynamics Statistics program of the U.S. census bureau. Here, we focus on two measures. First, the log number of new firms created in each year. Second, the share of employment in each industry coming from new firms. We choose these measures because they are most similar to the variables we study for France.

Concentration and firm creations are both measured at the 4-digit NAICS industry level. We keep only those industries where all four sales concentration measures are available in both 2002 and 2017. We also remove industries which ever undergo extreme changes in concentration — defined as a 30% or more change in the top-4 share in any 5-year period — since such changes may be due to firm reclassification rather than actual changes in concentration.

Appendix D shows statistics of the variables we use for U.S. industries. Like our main results, the summary statistics are weighted by the ex ante (i.e., 2002) number of firms in each industry. The summary statistics show that the share of sales coming from top firms increased from 2002-2017 for the average U.S. industry. However, the increases are small, ranging from 2.55 percentage points (for the top-4 share) to 3.5 percentage points (for the top-50 share). There is also substantial variation across industries: The 25th percentile increase is negative for two of the four concentration measures, indicating that superstar firms *fell* in importance in many industries. The standard deviation of concentration changes are between 5 and 6 for all four measures. As a result, there is substantial variation across industries that we can use to estimate the results.

We estimate bivariate long-differences specifications relating startup entry to changes in concentration. Our main dependent variable is a measure of new firms and our independent variable is a measure of the top-X share of sales. Both are in 15-year differences from 2002 to 2017.

$$\Delta NewFirms = \beta \Delta TopXSalesShare$$

As for our results from France, we begin with graphical evidence, also shown in Appendix D. The figure shows the relationship between the log change in number of new firms and the change in Top-8 share of sales by industry. The relationship is negative and immediate in the figure.

Table D.1 shows estimates of the bivariate regressions. Column 1-4 show regression results where the dependent variable is the log number of new firms. Column 1 shows results where the log change in new firms is the dependent variable and the change in the top-4 share of sales by industry is the independent variable. In columns 2-4, the independent variable is the top-8 share, top-20 share and top-50 share respectively. The coefficient is stable across all specifications and rises slightly, from -0.019 to -.025, when we use broader concentration measures. The concentration variables all have similar standard deviations in the cross section of industries, between 5 and 6. These estimates mean that a one standard deviation increase in the top share of industry sales will result in at least 0.1 log points less new creations. This is an economically large amount, given that the average change in firm creations is 0.18 log points. Notably, our estimates for the cross section of French industries (in Table 3) showed that the R^2 was larger as we used broader measures of concentration. The results here are similar: Both the magnitude of the coefficient, and the R^2 of the regression, are larger when we measure concentration using a larger group of superstars.

Columns 5-8 show results where the share of new firms' employment is the dependent variable. The average share of new firms' employment is about 1.41% in the sample. We multiply the dependent variable by 100 so the estimates can be interpreted as the effect on percentage points. Relative to the effects on log creations, the effect on employment shares are smaller but still economically significant. The coefficients themselves vary from -.017 to -.053, as in columns 1-4, are larger for broader concentration measures. One standard deviation increase in an industry's top-4 sales share therefore reduces the new-firm employment share in that industry by between between about 0.1 to 0.25 percentage points, or between 7 and 17% of the average effect. Overall, these effects are smaller than when we measure the log number of creations, but they are still economically large.

We view the estimates from the United States as confirming the negative relationship between concentration and new entry that we observe in France. An implication of our theory is that industries where superstars have become more productive also have less new firm entry as a result. We leave this for future work.

6.2 Reconciling seemingly contradictory U.S. evidence

The findings in this paper imply that the least productive entrepreneurs are increasingly discouraged from starting new businesses because the productivity of superstars has increased. At first glance, our results would seem to contradict an important stylized fact from the United States: that the fraction of new firms with extremely high growth rates has fallen (Decker et al., 2016b). The U.S. findings would seem to imply that the most productive entrepreneurs are being discouraged, rather than the least productive (as we find). However, we do not think that there is a contradiction: our measure of productivity is *ex ante*, i.e., we measure the *potential* productivity of entrepreneurs using their characteristics *at the time of entry.* As such, our finding is closely related to the result of Guzman and Stern (2020), who measure the ex ante quality of entrepreneurs in the United States. In contrast, Decker et al. (2016b) shows that the number of *ex post* highly productive entrepreneurs has fallen.

What mechanism can reconcile rising *ex ante* quality but declining *ex post* success? We think that superstars could be one such mechanism. If entrepreneurs must increasingly compete with the top firms in the market, then even high-quality entrepreneurs may have a lower likelihood of success, even as the least-productive entrepreneurs are discouraged altogether. Rather than leave this as a theoretical possibility, we would like to show directly that the skewness of firm growth has fallen in France, as Decker et al. (2016b) show for the United States.

Following the results of Decker et al. (2016b), we study the evolution of business dynamism by examining the skewness of the firm growth rate distribution. Specifically, following Decker et al. (2016b), we compute percentile differentials of the distribution of young firm growth rates. Using the DADS and the tax files datasets, we compute from 1994 to 2015 each firm's employment and sales growth rate from the previous year. Then in each year, we compute the 10th, 50th and 90th percentiles of the growth rate distributions among young firms, i.e., with age below 5-year. Finally, we examine the evolution of the differential between the 90th and the 50th percentiles as well as the difference between the 50th and the 10th percentiles, of the firm growth rate distribution.

Figure 11 shows the evolution over time of these differentials from the employment-weighted distribution of firm-level employment and sales growth rates. To facilitate focusing on the trends, we smooth the resulting time series using Lowess smoothing. In the upper panel of Figure 11, it is apparent that there is a secular decline in the 90-50 differential of employment growth among young firms. This echoes the results of Decker et al. (2016b), showing a similar decline in the US. This result, together with those presented in section 5, highlight that an increase in entrepreneur quality (also document by Guzman and Stern, 2020, in the US) can be associated with a declining skewness of the young firm employment growth rate distribution.

Furthermore, in the lower panel of Figure 11, we present the evolution of the same percentile differentials from the distribution of sales growth rates. A very different picture emerges from this analysis. Indeed, we observe an increase in the 90-50 differential of sales growth among young firms. This highlights that even if the skewness of young firms' employment growth rate has been declining, a similar trend is not apparent when considering the distribution of young firms' sales growth rates. These results are directly related to Barkai and Panageas (2021), showing that while young firms had an underwhelming performance in terms of creating jobs in the last couple of decades, their performance in terms of generating sales was not similarly weak. Overall, our results question whether the well-documented decline in young firm employment growth rate skewness should be interpreted as a decline in business dynamism.

7 Conclusion

Why has the rate of new business formation declined? Using administrative data from France, we document the decline in the number of firm creations from 1994 to 2005 similar to that documented in the United States (e.g., Decker et al., 2016a; Barkai and Panageas, 2021). We show that this decline was accompanied by a decrease in the share of aggregate employment, and to a lesser extent of aggregate value added, accounted for by young firms. Meanwhile, we show that the *quality* of new firms has increased, implying that the decline in the quantity of new businesses is due to less business formation by low-ability potential new entrants.

We offer a novel explanation for this decline in new business formation which is based on the rise of the largest and most productive "superstar firms" (Autor et al., 2020). We propose a simple model of superstar firms and firm entry and show support for our model's predictions. First, we find that new firm creation has decreased in industries experiencing a larger rise of superstar firms – i.e., industries with larger increases in product market concentration. Second, the rise of superstar firms has discouraged low-ability entrepreneurs from starting a business, but not high-ability ones, resulting in a higher average quality of new businesses. Third, superstar firms have displaced low-productivity incumbent firms from the market.

Figures

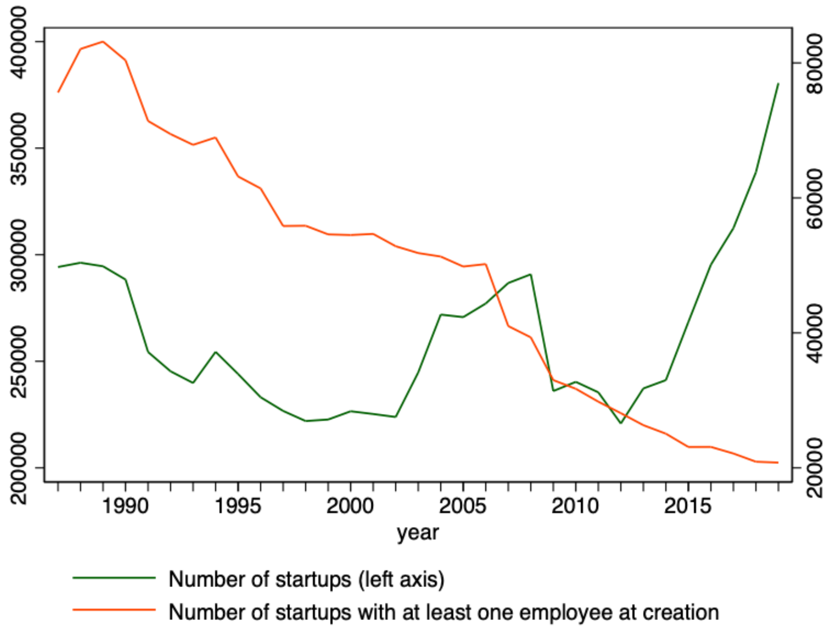


Figure 1: **Declining business formation for startups with at least one employee.** This figure plots the evolution over time of the number of startups in France (green line, left axis) and of the number of startups with at least one employee at creation (orange line, right axis). *Source:* Firm creations registry.

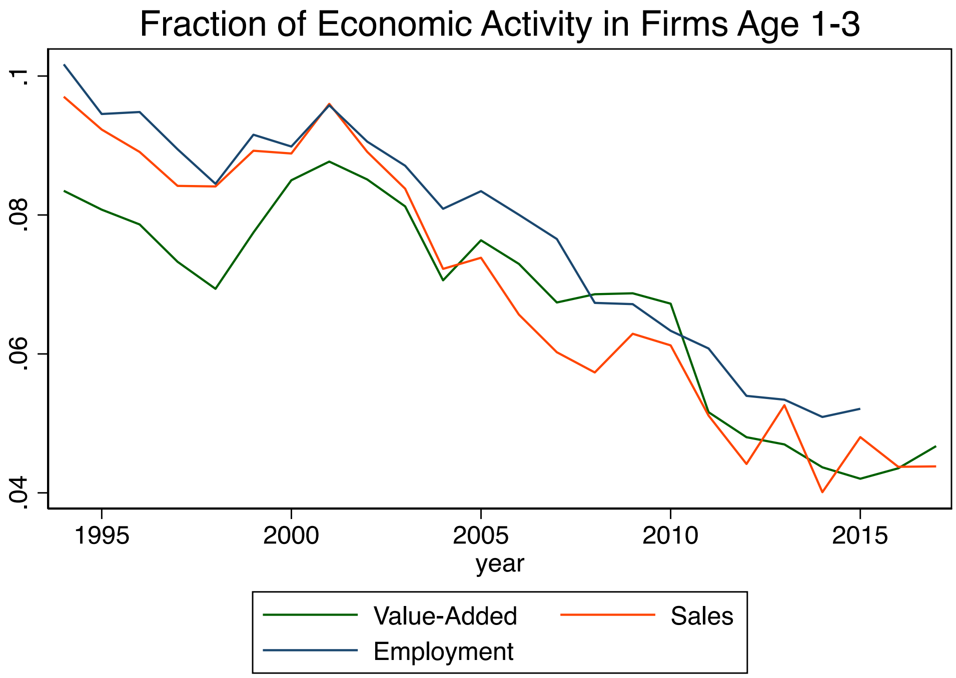
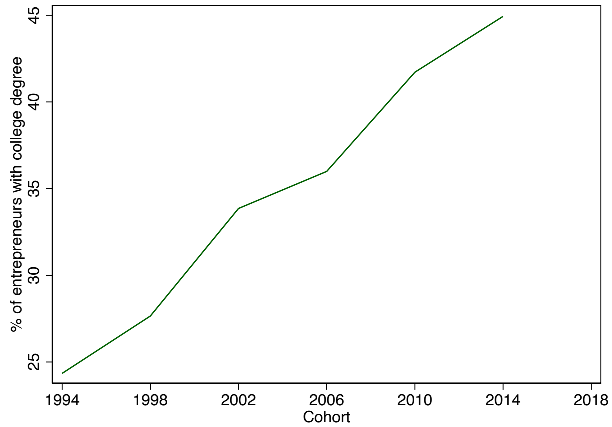
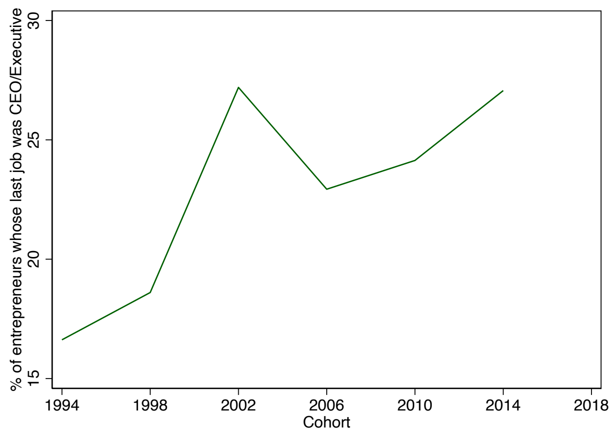


Figure 2: **Young firms' share of value added, sales, and employment.** This Figure shows the share of total value-added, sales and employment attributable to incorporated firms ages 1-3 in France by year. The sample excludes public sector, education, social services, agriculture, and energy sectors. *Source*: Value-added and sales are calculated from tax data, available until 2017. Employment is calculated using full-time employees from DADS (employer payroll) data, available until 2015.

Panel A: Entrepreneurs with college degree



Panel B: Entrepreneurs previously employed as executive or CEO



Panel C: Serial entrepreneurs

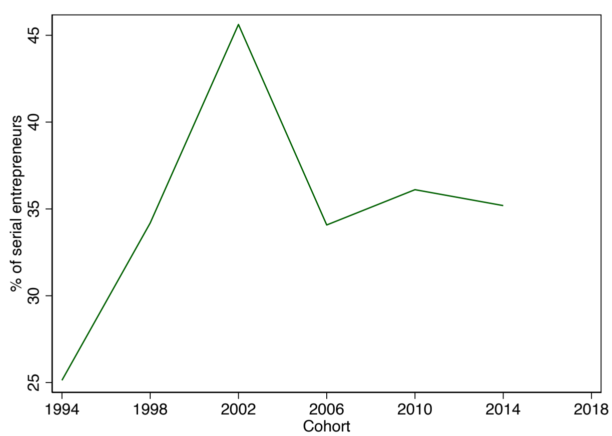


Figure 3: **The aggregate increase in entrepreneur ability.** This figure plots the evolution of proxies for entrepreneur quality over time. Panel A shows the fraction of entrepreneurs with a college degree (at least two-year college degree) in each cohort of the SINE survey from 1994 to 2014. Panel B shows the fraction of entrepreneurs previously employed as executive or CEO in each cohort of the SINE survey from 1994 to 2014. Panel C shows the fraction of serial entrepreneurs (who had previously created a business) in each cohort of the SINE survey from 1994 to 2014. *Source:* SINE survey.

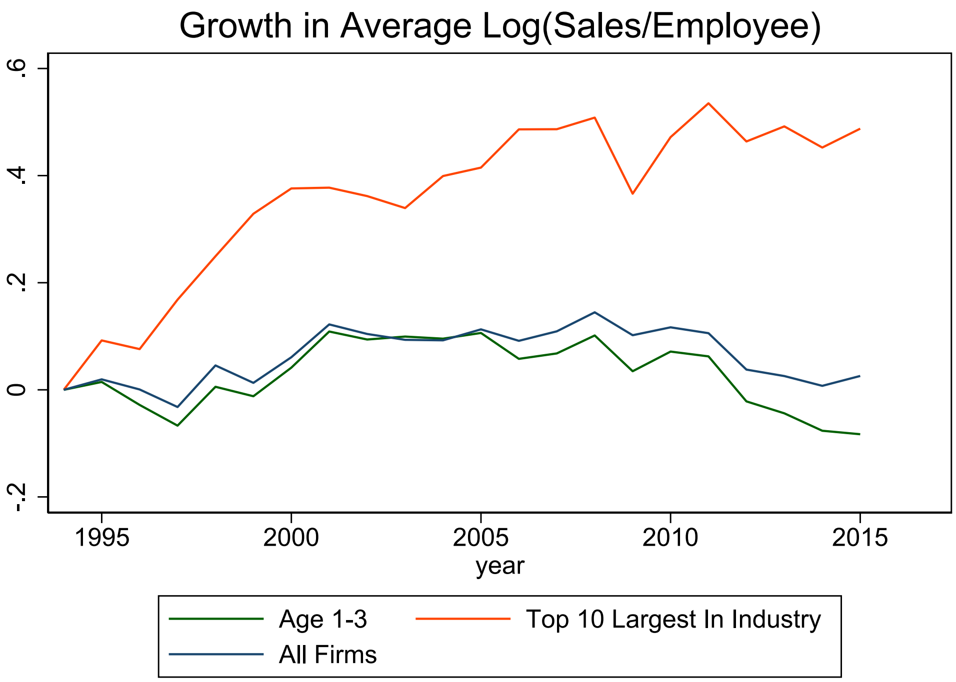
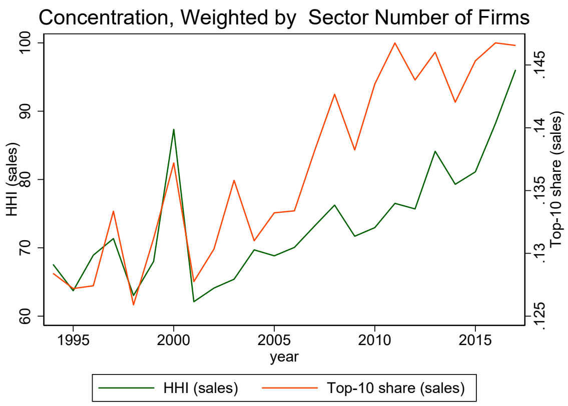
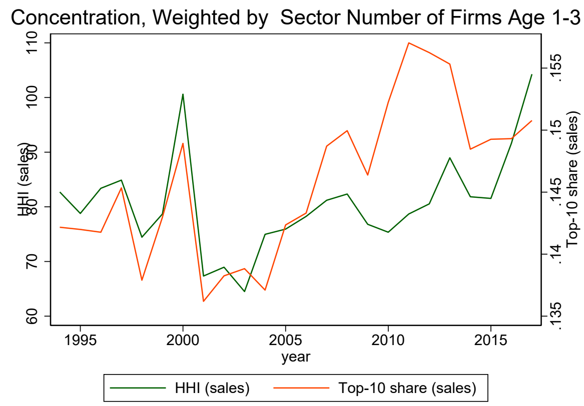


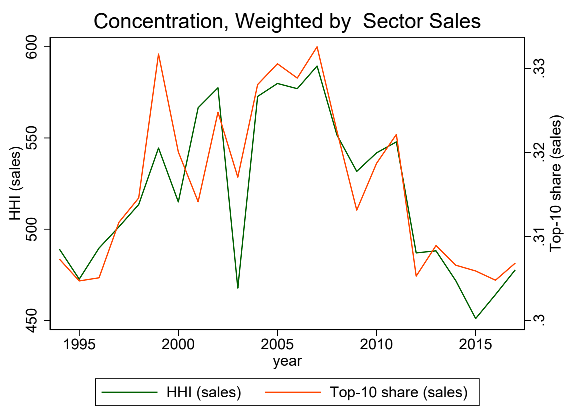
Figure 4: **The growing productivity gap between superstar firms and other incorporated firms.** Productivity is measured as the unweighted average of log sales per employee over all firms in each group; figure shows changes in average productivity since 1994. "Top-10 Largest in Industry" includes firms in the top-10 of sales in each industry each year. "All Firms" and "Age 1-3" include all firms and firms age 1-3 respectively. Results are similar, but noisier, using value-added per employee. We add one employee to each firm to include the owner and guarantee that every firm has at least one employee. Source: Sales are taken from tax data and employment is taken from DADS (employer payroll) data.



(a) Weighted by industry number of firms



(b) Weighted by industry number of firms age 1–3



(c) Weighted by industry sales

Figure 5: **The increase in market concentration and the rise of superstar firms.** HHI and top-10 share of sales are calculated within each 3-digit industry by year. Industry HHI and top-10 share are then aggregated across industries to create a national average each year, weighting industries by number of firms (top-left panel), number of firms ages 1-3 (top-right panel) and total industry sales (bottom panel). The sample excludes public sector, education, social services, agriculture and energy sectors. Source: Sales are calculated from tax data, available until 2017.

Figure 6: **Changes in log creations vs change in top-20 share of firms.** The x-axis shows the change in the top 20 share of firm sales from 1994-2020. The y-axis shows the change in the log creations between the years 1994 and 2015. Circle sizes are proportional to the number of firms in each industry in 1994. The fit line is from a WLS regression. Source: Business formation data comes from exhaustive firm registry data, sales are taken from tax data, and employment is taken from DADS (employer payroll) data.

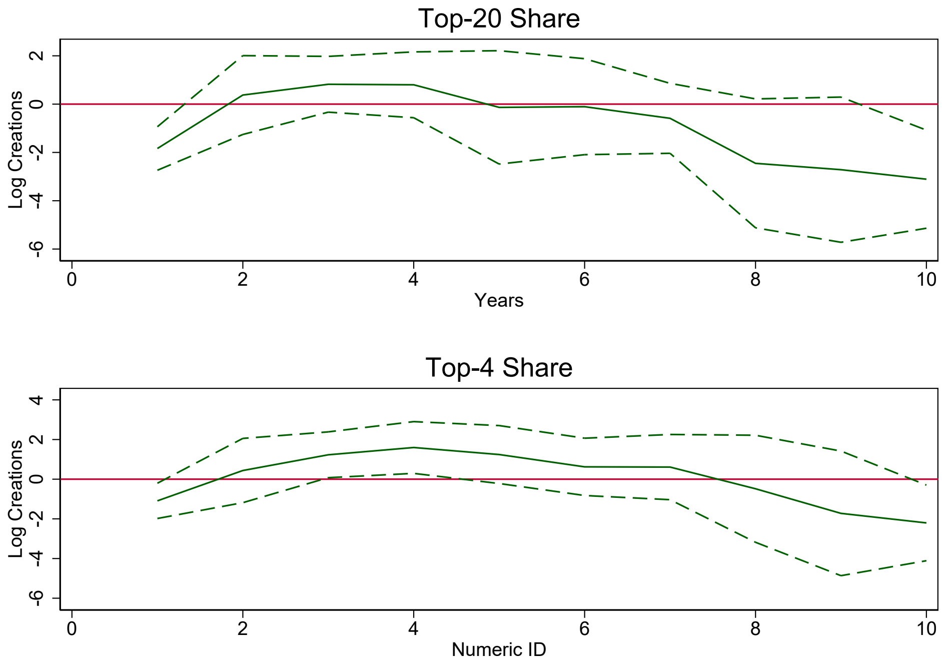


Figure 7: **Local projection of log creations on concentration.** Estimates from a local projection of log creations on measures of concentration. Estimates are from a regression of future values of creations on current concentration, controlling for two lags of concentration and creations. Upper panel uses the top-20 share as a measure of concentration, and lower panel uses the top-4 share of sales.

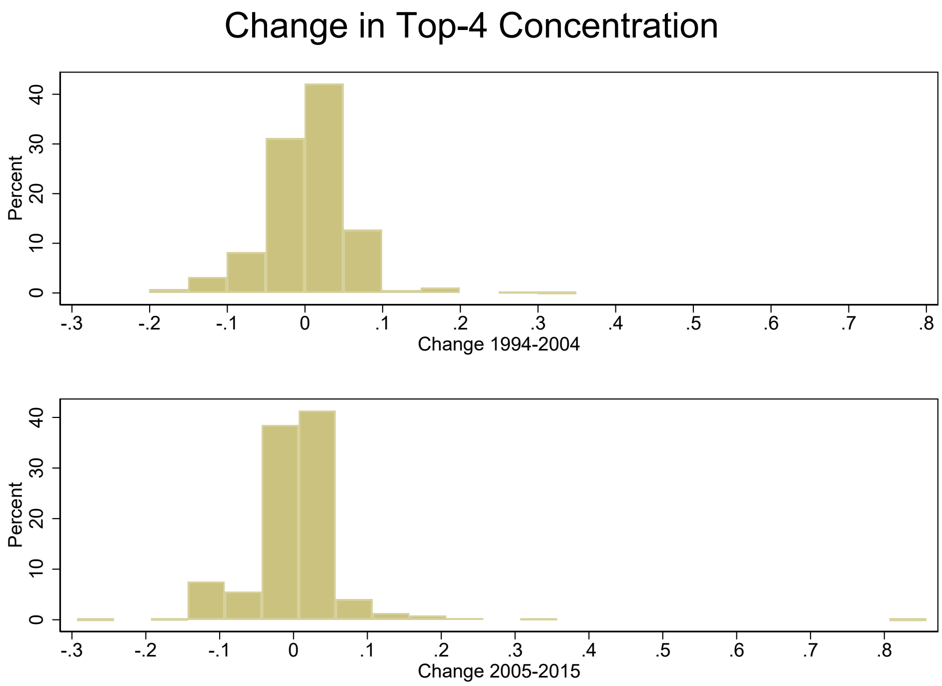


Figure 8: **Histogram of industry-level concentration changes, measured using top-20 share of sales. Upper figure shows concentration changes from 1994-2004 and lower figure shows concentration changes from 2005-2015.** The x-axis shows the change in concentration at the industry level over the specified time period. Y-axis shows the percent of firms in industries at that level of concentration change. Industries are weighted by number of firms in 1994.

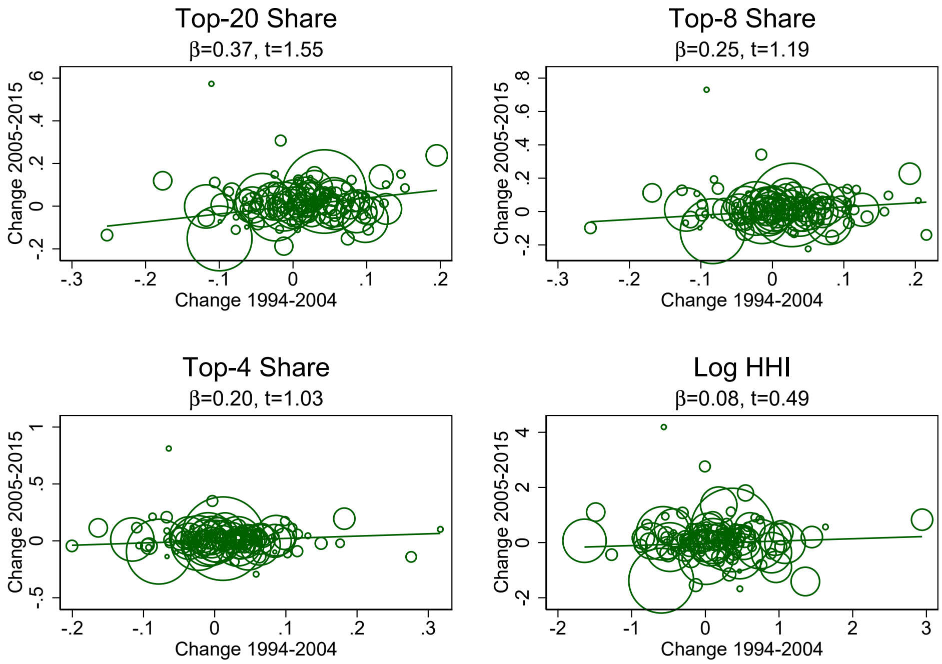


Figure 9: **Concentration change by industry from 1994-2004 vs concentration change by industry from 2005-2015.** The x-axis in each figure shows the change in superstar concentration from 1994-2004. The y-axis in each figure shows the change in superstar concentration from 2004-2015. Circle sizes are proportional to the number of firms in each industry in 1994. The fit line is from a WLS regression. Sales are taken from tax data.

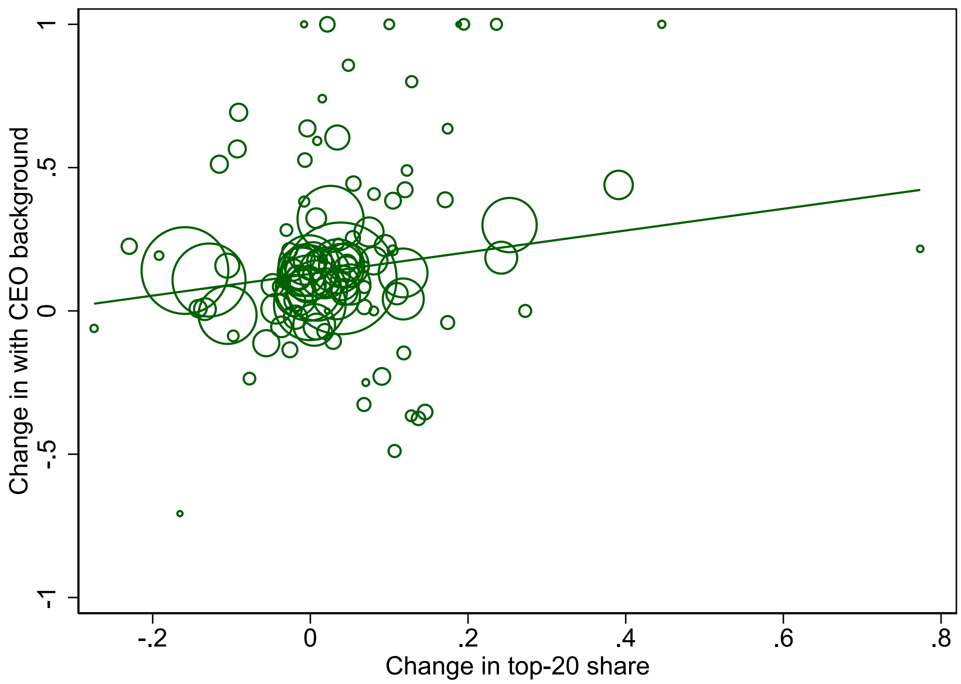
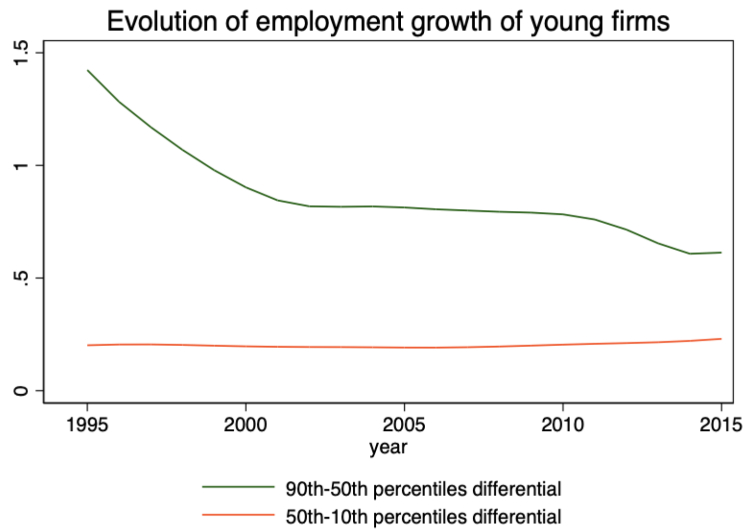


Figure 10: **The increase in entrepreneur background and the increase in the concentration of superstar firms.** The x-axis shows the change in superstar concentration measured as the change in top-20 share of sales, from 1994-2014. The y-axis shows the change in the fraction of entrepreneurs with college degree starting a new business between the initial (1994) and final (2014) cohorts of the SINE survey. Circle sizes are proportional to the number of firms in each industry in 1994. The fit line is from a WLS regression. Source: Education data comes from the SINE, sales are taken from tax data, and employment is taken from DADS (employer payroll) data.



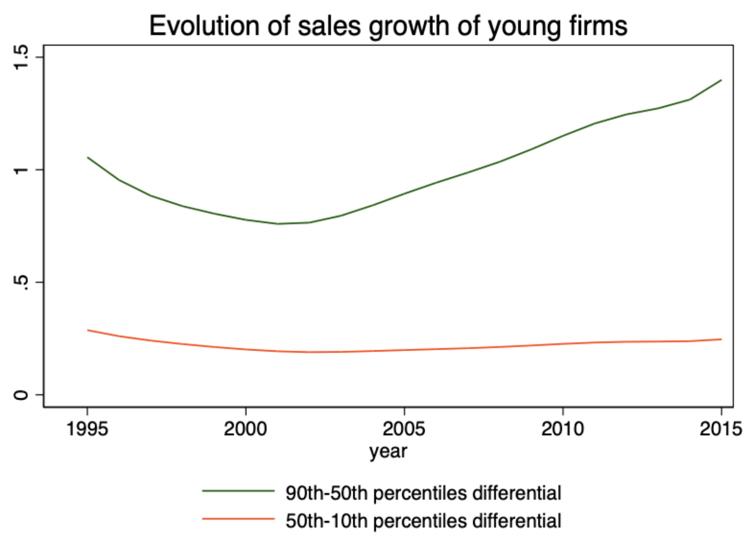


Figure 11: **The decline (increase) in the skewness of employment (sales) growth rate of young firms.** The figures shows the 90th-50th and 50th-10th percentiles differential in young firm growth rates in terms of employment (upper panel) and sales (lower panel). The 90th-50th and 50th-10th percentiles differentials are respectively the difference between the 90th and the 50th percentiles and the difference between the 50th and the 10th percentiles, of the employment-weighted distribution of young firm growth rates. The upper panel uses the employment growth rate from the year $t-1$ to t. The lower panel uses the sales growth rate from the year $t-1$ to t. These growth rates are calculated for firms with age below 5 in year t. Times series are smoothed using Lowess smoothing. Sources: Firm creation registry, DADS and tax files.

Tables

Table 1: **Summary Statistics.** *Source*: 1994, 1998, 2002, 2006, 2010, and 2014 SINE surveys, Tax files and DADS. This table contains summary statistics for the main entrepreneur characteristics we use in our analysis. Most variables are dummies so that the reported means stand for percentage in the category. The only exceptions are Nb. Employees x-Y and Sales x-Y. Appendix A1 provides a description of each variable.

	Mean	Sd	25%	50%	75%	95%	99%	N
Panel A: Regulatory Filings								
Levels								
# New Firms (Log)	3.93	1.97	2.30	3.74	5.45	7.38	8.39	3,285
Sales Share of New Firms	0.05	0.05	0.01	0.03	0.06	0.15	0.24	3,308
Employment Share of New Firms	0.05	0.05	0.01	0.03	0.06	0.16	0.25	3,032
Sales Share of Top 20 Firms	0.48	0.24	0.29	0.45	0.67	0.90	0.98	3,308
Sales Share of Top 8 Firms	0.36	0.22	0.20	0.31	0.51	0.79	0.93	3,308
Sales Share of Top 4 Firms	0.27	0.19	0.14	0.22	0.37	0.69	0.86	3,308
HHI (Log)	5.28	1.47	4.46	5.35	6.27	7.54	8.65	3,308
Sales Share of Young Firms	0.08	0.07	0.03	0.06	0.11	0.21	0.31	3,308
Industry # of Firms in 1994	4165	8023	493	1135	3480	18844	38399	3,308
Long-term changes (1994-2015)								
Change in (log) # of New Firms	-0.73	0.88	-1.39	-0.87	-0.01	0.77	1.04	135
Change in # of New Firms	-0.28	0.78	-0.75	-0.58	-0.01	1.17	1.83	135
Change in Sales Share of New Firms	-0.06	0.06	-0.08	-0.05	-0.03	-0.01	0.02	135
Change in Employment Share of New Firms	-0.07	0.05	-0.08	-0.06	-0.03	-0.01	0.00	135
Change in Sales Share of Top 20 Firms	0.03	0.12	-0.03	0.02	0.10	0.21	0.46	135
Change in Sales Share of Top 8 Firms	0.03	0.12	-0.02	0.02	0.10	0.24	0.44	135
Change in Sales Share of Top 4 Firms	0.03	0.12	-0.02	0.02	0.09	0.23	0.46	135
Chg in (log) HHI	0.23	0.91	-0.24	0.17	0.62	2.08	3.59	135
Change in Sales Share of Young Firms	-0.09	0.08	-0.12	-0.08	-0.05	-0.00	0.11	135
Panel B: Entrepreneur Survey Statistics (SINE)								
Levels								
College Education	0.4	0.5	0.0	0.0	1.0	1.0	1.0	37,260
Previously CEO-Executive	0.2	0.4	0.0	0.0	0.0	1.0	1.0	37,260
Serial Entrepreneur	0.4	0.5	0.0	0.0	1.0	1.0	1.0	37,260
$\mathbb{1}$(Employment 3-Y > 0)	0.6	0.5	0.0	1.0	1.0	1.0	1.0	29,127
$\mathbb{1}$(Employment 5-Y > 0)	0.5	0.5	0.0	0.0	1.0	1.0	1.0	29,127
$\mathbb{1}$(Employment 10-Y > 0)	0.3	0.5	0.0	0.0	1.0	1.0	1.0	13,299
Nb. Employees 3-Y	4.4	13.4	1.0	2.0	4.0	14.0	42.0	22,650
Nb. Employees 5-Y	5.4	23.0	1.0	2.0	5.0	17.0	54.0	18,606
Nb. Employees 10-Y	7.8	29.4	1.0	3.0	7.0	24.0	83.0	6,189
Sales 3-Y	695.9	4458.2	105.8	230.1	525.2	2047.7	7754.0	29,516
Sales 5-Y	922.6	5704.2	123.7	282.0	649.0	2732.0	10357.6	18,606
Sales 10-Y	1349.1	6152.3	140.5	366.0	949.5	4380.6	15831.4	8,714

Sector	Av. # Firms 1994 (1)	Av. # Firms 2015 (2)	Δ Top-20 (3)	Δ Top-8 (4)	Δ Top-4 (5)	Δ Log HHi (6)	Δ New Firms (7)
Textile and Clothing	3789	2448	0.26	0.20	0.15	1.64	-0.82
Leather and Footwear	631	531	0.21	0.20	0.20	0.98	-0.74
Other Manufacturing	2382	3104	0.11	0.12	0.11	0.88	-0.72
Chemicals	693	700	0.10	0.09	0.08	0.44	-0.70
Woodworking	1153	1446	0.09	0.07	0.06	0.65	-0.52
Education	3940	20000	0.09	0.11	0.10	1.45	0.26
Automotive Trade and Repair	28079	50598	0.08	0.06	0.04	0.69	-0.36
Machinery and Equipment	2119	3306	0.07	0.06	0.05	0.34	-0.54
Electronic Equipment	1121	1651	0.04	0.03	0.02	0.10	-0.59
Publishing and Printing	6762	8347	0.03	0.02	0.01	0.23	-0.78
Transport Equipment	748	942	0.02	0.03	0.04	0.12	-0.64
Non-energy Extraction	906	778	0.02	0.02	0.03	0.29	-0.64
Transport and Communications	13715	30544	0.02	0.02	0.01	0.17	0.33
Metallurgy	4781	6955	0.02	0.02	0.01	0.40	-0.55
Rubber and Plastics	3050	3002	0.02	0.03	0.02	0.17	-0.81
Mineral Products	904	1075	-0.00	-0.02	-0.04	-0.16	-0.69
Health Care	3881	17688	-0.03	-0.02	-0.01	-0.49	0.54
Construction	21132	69168	-0.05	-0.04	-0.03	-0.75	0.26
Commercial Real Estate	18860	88980	-0.06	-0.04	-0.03	-0.40	0.21
Hotels and Restaurants	19478	65816	-0.08	-0.07	-0.07	-1.47	0.27

Table 2: **Summary statistics on long-term changes at the sector level (1994–2015).** *Source*: Tax files and DADS. This table contains summary statistics at the sector level for 20 sectors. *Top-X* are the long-term differences in the share of the largest X firms in a sector from 1994 to 2015, defined as the weighted average of the long-term differences in the share of these firms in each industry within the sector, using tge number of firms in each industry in 1994 as weights. Log HHI is the HHI of firm sales. Δ New Firms is the long-term difference in the number of new firms. Appendix A1 provides a description of each variable.

	Concentration Measure			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
Baseline Effect on New Firms	-2.7***	-2.6***	-2.6***	-.25***
	(.48)	(.61)	(.74)	(.05)
R^2	.2	.15	.12	.17
IV Purging New Firms	-2.1**	-1.9*	-1.9*	-.15*
	(.82)	(1)	(1.1)	(.088)
R^2	.19	.13	.11	.14
Industry Growth Ctrl	-1.7***	-1.8***	-2***	-.17***
	(.53)	(.55)	(.53)	(.045)
Industry Growth Coef	.85***	.89***	.92***	.87***
	(.11)	(.1)	(.1)	(.1)
R^2	.47	.46	.46	.47
Initial VA Ctrl	-2.6***	-2.5***	-2.4***	-.24***
	(.49)	(.62)	(.76)	(.049)
Initial VA Coeff	.09	.086	.075	.086
	(.071)	(.071)	(.074)	(.07)
R^2	.21	.15	.13	.18
Within-Sector	-1.3***	-1.1**	-1.2**	-.13**
	(.48)	(.52)	(.49)	(.06)
R^2	.48	.47	.47	.48
Startup Output Share	-.095**	-.11**	-.11*	-.0082
	(.045)	(.052)	(.057)	(.0058)
R^2	.053	.055	.042	.038

Table 3: **New firm creations have decreased in industries where the market share of superstar firms has increased.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. Estimates are weighted by the number of firms in each industry in 1994. The dependent variable in each regression is the long-term change in market concentration at the industry level from 1994-2015. Each column reports regressions using a different measure for the long-term change in market concentration. *Top-X* is the change in the share of the largest X firms in an industry from 1994 to 2015. Log HHI is the HHI of firm sales. Row labeled "Baseline Effect on New Firms" shows estimates from a bivariate regression of the log change in number of new firms on the change in concentration. "IV Purging New Firms" uses instruments for the concentration change using a measure of concentration that omits the effects of any new firms on concentration. Rows labeled "Industry Growth Ctrl" and "Initial VA Ctrl" add controls for industry-level growth and initial industry-level value added, respectively. Rows labeled "Industry Growth Coef" and "Initial VA Coef" report the regression coefficients for these controls. Row labeled "Within-Sector" adds sector fixed effects. Row labeled "Startup Output Share" uses as the dependent variable the total share of output coming from firms ages 0-3. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Change in Firm Creations over 5- and 10-Years			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
10-Year Changes	-2.5***	-2.3***	-1.9***	-.13**
	(.57)	(.64)	(.67)	(.057)
R^2	.051	.035	.021	.017
10-Year Changes, No New Firms	-4.6***	-4.3***	-4.3***	-.33*
	(1.5)	(1.5)	(1.5)	(.17)
R^2	.089	.072	.063	.04
5-Year Changes	-1*	-.81	-.46	-.037
	(.54)	(.53)	(.49)	(.048)
R^2	.0053	.003	.00063	.00069

Table 4: **Effect on Firm Creations, 5 and 10 Year Changes.** Differences regressions at the industry (3-digit) level for 135 industries. The dependent variable is the change in log firm creations, measured over ten-year differences (rows 1-2) or five-year differences (row 3). Rows 1 and 3 use overall concentration measures as the independent variables and row 2 uses a measure of concentration that purges the effects of new firm creations. All estimates are weighted by the number of firms in 1994. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Change in Firm Creations by Concentration Time Period			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
1994-2001	-4.3***	-3.6***	-3.9***	-.39***
	(1.2)	(1.3)	(1.3)	(.13)
R^2	.1	.072	.069	.096
2001-2008	-4.2***	-3.5***	-3.8**	-.3***
	(1.2)	(1.2)	(1.5)	(.076)
R^2	.11	.067	.06	.084
2008-2015	-3.5*	-2.4	-1.8	-.19
	(1.9)	(1.9)	(1.7)	(.17)
R^2	.07	.028	.012	.02

Table 5: **The effects of concentration on creations are greatest from 1994-2001.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable is the change in log creations at the industry level over the specified time period. Estimates are weighted by the number of firms in each industry in 1994. *Change in top X share* is the % change in the share of output of the largest X firms in an industry over the specified time period. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Change in Firm Creations by Sector			
	Top-20 (1)	Top-8 (2)	Top-4 (3)	Log HHI (4)
Manufacturing	-1.2*	-.92	-.53	-.042
	(.68)	(.78)	(.81)	(.13)
R^2	.028	.0074	-.0077	-.011
Observations	71	71	71	71
Commerce & Repair	-1	-.59	-1.1	-.074
	(1.2)	(1.5)	(1.4)	(.15)
R^2	-.031	-.051	-.038	-.048
Observations	19	19	19	19
Property & Business Services	-.97	-.32	-.077	.073
	(1.3)	(1)	(.84)	(.26)
R^2	-.11	-.12	-.12	-.12
Observations	10	10	10	10
Transportation & Communication	-1.1*	-1.1*	-1.1**	-.11**
	(.54)	(.56)	(.51)	(.051)
R^2	.046	.031	.035	.029
Observations	18	18	18	18

Table 6: **Effects on Creations by Top-Level Sector.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable is the change in log creations at the industry level over the specified time period. Estimates are weighted by the number of firms in each industry in 1994. *Change in top X share* is the % change in the share of output of the largest X firms in an industry over the specified time period. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Change in Firm Creation Rates			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
Creation Rate, Current Firms	.026	.029	.025	.0035
	(.028)	(.031)	(.032)	(.0029)
R^2	.018	.018	.0075	.038
Total Rate Since Baseline	-4.8***	-4.9***	-4.7***	-.51***
	(1.2)	(1.4)	(1.5)	(.11)
R^2	.24	.19	.15	.27

Table 7: **Effect on Firm Creation Rates.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable in each regression is the change in log creations at the industry level from 1994-2015. Estimates are weighted by the number of firms in each industry in 1994. *Top-X* is the change in the share of the largest X firms in an industry from 1994 to 2015. Log HHI is the HHI of firm sales. Row labeled "Creation Rate, Current Firms" shows results from a regression where the dependent variable is $\frac{NewFirms_{2015}}{TotalFirms_{2015}} - \frac{NewFirms_{1994}}{TotalFirms_{1994}}$ and the independent variable is the change in concentration from 1994-2015. Row labeled "Creation Rate, Current Firms" shows results from a regression where the dependent variable is $\frac{NewFirms_{2015}}{TotalFirms_{2004}} - \frac{NewFirms_{1994}}{TotalFirms_{1994}}$.

	Change in Founder Characteristics			
	Top-20 (1)	Top-8 (2)	Top-4 (3)	Log HHI (4)
HS Degree	.24*	.26*	.24	.02
	(.13)	(.15)	(.15)	(.015)
R^2	.018	.016	.0085	.0079
College Degree	.34	.39	.38	.032
	(.25)	(.27)	(.28)	(.022)
R^2	.049	.049	.038	.04
Serial Entrepreneur	.15	.21	.22	.018
	(.15)	(.19)	(.22)	(.015)
R^2	.0046	.013	.011	.01
Founder Age	.069	.38	2.3	.035
	(2.3)	(2.6)	(2.8)	(.25)
R^2	-.0093	-.0092	-.0061	-.0092
CEO/Executive	.26*	.33**	.38**	.027**
	(.13)	(.14)	(.15)	(.012)
R^2	.033	.044	.051	.036
Blue Collar	-.11*	-.13**	-.12**	-.011*
	(.056)	(.06)	(.058)	(.0059)
R^2	.04	.042	.029	.04
Innovative	.37**	.4**	.37*	.028
	(.16)	(.19)	(.21)	(.022)
R^2	.068	.061	.042	.033

Table 8: **Greater concentration results in founders with more experience and education.** Cross-sectional regressions at the industry (3-digit) level. The dependent variable is the change in average founder characteristics from 1994-2014. Estimates are weighted by the number of firms in each industry in 1994. *Top X* is the % change in the share of output of the largest X firms in an industry from 1994 to 2014. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Concentration Measure			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
Effect of Concentration	-2.2***	-2**	-1.9**	-.24***
	(.82)	(.92)	(.96)	(.055)
College Completion	1.5**	1.5**	1.6**	1.4**
	(.62)	(.64)	(.65)	(.61)
HS Completion	-.4	-.37	-.34	-.13
	(.6)	(.61)	(.61)	(.58)
Share CEO Founder	.17	.1	.1	-.2
	(.79)	(.8)	(.81)	(.77)
Avg Founder Age	-.02	-.018	-.016	-.009
	(.027)	(.027)	(.027)	(.025)
Share Blue-Collar	4.5*	4.6*	4.5*	4.2*
	(2.4)	(2.5)	(2.5)	(2.1)
R^2	.18	.15	.14	.22

Table 9: **New firm creations have decreased in industries where superstar concentration has increased, After Controlling For Demographic Differences by Industry.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable in each regression is the change in log creations at the industry level from 1994-2015. Estimates are weighted by the number of firms in each industry in 1994. *Top-X* is the change in the share of the largest X firms in an industry from 1994 to 2015. Log HHI is the HHI of firm sales. Demographic controls measure average founder characteristics from the SINE survey. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	The Effects of Concentration on Firm Survival			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
Overall Survival	-.13	-.12	-.11	-.0064
	(.093)	(.1)	(.11)	(.013)
R^2	.044	.028	.016	.0048
Employment-Weighted	-.25**	-.28**	-.28**	-.02
	(.1)	(.12)	(.13)	(.013)
R^2	.099	.09	.08	.058
High Productivity	-.042	-.036	-.036	.00018
	(.069)	(.076)	(.08)	(.0088)
R^2	-.0019	-.0042	-.0047	-.0075
Low Productivity	-.16*	-.15*	-.14	-.011
	(.082)	(.088)	(.091)	(.011)
R^2	.071	.047	.033	.029

Table 10: **Concentration causes incumbent firms to exit.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable is the share of 1994 firms that still exist in the industry as of 2015. *Top X* is the change in the share of output of the largest X firms in an industry from 1994 to 2015. "Overall Survival" shows the effects on all firms. "Employment-Weighted" weights each firm by its share of employment within industry. "High Productivity" considers only firms with value added per worker above the industry median and "Low Productivity" considers only firms with value added per worker below the industry median. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

Bibliography

Aghion, Philippe, Ufuk Akcigit, Matthieu Lequien, and Stefanie Stantcheva. 2017. "Tax simplicity and heterogeneous learning." Tech. rep., National Bureau of Economic Research.

Akcigit, Ufuk and Sina T Ates. 2021. "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory." *American Economic Journal: Macroeconomics* 13 (1):257–98.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The fall of the labor share and the rise of superstar firms." *The Quarterly Journal of Economics* 135 (2):645–709.

Baker, Jonathan B. 2019. *The antitrust paradigm.* Harvard University Press.

Barkai, Simcha and Stavros Panageas. 2021. "Value without Employment." Tech. rep., National Bureau of Economic Research.

Bento, Pedro and Diego Restuccia. 2019. "The role of nonemployers in business dynamism and aggregate productivity." *NBER working paper* (w25998).

Bessen, James. 2020. "Industry concentration and information technology." *The Journal of Law and Economics* 63 (3):531–555.

Bijnens, Gert and Jozef Konings. 2020. "Declining business dynamism in Belgium." *Small Business Economics* 54 (4):1201–1239.

Bornstein, Gideon. 2018. "Entry and profits in an aging economy: The role of consumer inertia."

Criscuolo, Chiara, Peter N Gal, and Carlo Menon. 2014. "The dynamics of employment growth: New evidence from 18 countries." .

De Loecker, Jan, Jan Eeckhout, and Simon Mongey. 2021. "Quantifying market power and business dynamism." *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-54).

De Ridder, Maarten, Basile Grassi, Giovanni Morzenti et al. 2021. "The hitchhikerâs guide to markup estimation." Tech. rep.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The role of entrepreneurship in US job creation and economic dynamism." *Journal of Economic Perspectives* 28 (3):3–24.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. 2016a. "Declining business dynamism: What we know and the way forward." *American Economic Review* 106 (5):203–07.

———. 2016b. "Where has all the skewness gone? The decline in high-growth (young) firms in the US." *European Economic Review* 86:4–23.

Decker, Ryan A, John C Haltiwanger, Ron S Jarmin, and Javier Miranda. 2018. "Changing business dynamism and productivity: Shocks vs. responsiveness." Tech. rep., National Bureau of Economic Research.

Del Gatto, Massimo, Gianmarco IP Ottaviano, and Giordano Mion. 2006. "Trade integration, firm selection and the costs of non-Europe." .

Engbom, Niklas. 2019. "Firm and worker dynamics in an aging labor market." Tech. rep.

Fazio, Catherine, Jorge Guzman, Fiona Murray, and Scott Stern. 2016. "A new view of the skew: Quantitative assessment of the quality of American entrepreneurship." *Kauffman Foundation New Entrepreneurial Growth, Kansas City, MO: February* .

Furman, Jason and Peter R Orszag. 2018. "Slower productivity and higher inequality: Are they related?" *Peterson Institute for International Economics Working Paper* (2018-4).

Gourio, François, Todd Messer, Michael Siemer et al. 2014. "What Is the Economic Impact of the Slowdown in New Business Formation?" *Chicago Fed Letter* (Sep).

Guzman, Jorge and Scott Stern. 2016. "The State of American Entrepreneurship: New Estimates of the Quality and Quantity of Entrepreneurship for 32 US States, 1988-2014." Tech. rep., National Bureau of Economic Research.

———. 2020. "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy* 12 (4):212–43.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda. 2013. "Who creates jobs? Small versus large versus young." *Review of Economics and Statistics* 95 (2):347–361.

Hathaway, Ian and Robert Litan. 2014a. "Declining business dynamism: It's for real." *Economic Studies at Brookings, Brookings Institution. https://www. brookings. edu/research/declining-business-dynamism-its-for-real* .

Hathaway, Ian and Robert E Litan. 2014b. "Whatâs driving the decline in the firm formation rate? A partial explanation." *The Brookings Institution* 4.

Henrekson, Magnus and Tino Sanandaji. 2014. "Small business activity does not measure entrepreneurship." *Proceedings of the National Academy of Sciences* 111 (5):1760–1765.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar. 2020. "Can unemployment insurance spur entrepreneurial activity? Evidence from France." *The Journal of Finance* 75 (3):1247–1285.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania. 2018. "From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share." Tech. rep., National Bureau of Economic Research.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg. 2019. "The industrial revolution in services." Tech. rep., National Bureau of Economic Research.

———. 2021. "The industrial revolution in services." Tech. rep.

Jiang, Helu and Faisal Sohail. 2022. "Skill-biased entrepreneurial decline." *Review of Economic Dynamics* .

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2019. "Demographic origins of the startup deficit." Tech. rep., National Bureau of Economic Research.

Kozeniauskas, Nicholas. 2018. “Whatâs driving the decline in entrepreneurship.” .

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann. 2021. “100 Years of Rising Corporate Concentration.” *Available at SSRN 3936799* .

Landier, Augustin and David Thesmar. 2008. “Financial contracting with optimistic entrepreneurs.” *The Review of Financial Studies* 22 (1):117–150.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard. 2018. “Information technology and returns to scale.” *Available at SSRN 3458604* .

Levine, Ross and Yona Rubinstein. 2017. “Smart and illicit: who becomes an entrepreneur and do they earn more?” *The Quarterly Journal of Economics* 132 (2):963–1018.

Lyonnet, Victor and Léa H Stern. 2022. “Venture Capital (Mis) allocation in the Age of AI.” *Fisher College of Business Working Paper* (2022-03):002.

Melitz, Marc J. 2003. “The impact of trade on intra-industry reallocations and aggregate industry productivity.” *econometrica* 71 (6):1695–1725.

Melitz, Marc J and Gianmarco IP Ottaviano. 2008. “Market size, trade, and productivity.” *The review of economic studies* 75 (1):295–316.

Philippon, Thomas. 2019. *The great reversal.* Harvard University Press.

Pugsley, Benjamin W, Petr Sedlacek, and Vincent Sterk. 2019. “The nature of firm growth.” *Available at SSRN 3086640* .

Pugsley, Benjamin Wild and Eric Hurst. 2011. “What do small businesses do?” *Brookings Papers on Economic Activity* 43 (2):73–142.

Salgado, Sergio. 2020. “Technical change and entrepreneurship.” *Available at SSRN 3616568* .

Schoar, Antoinette. 2010. “The divide between subsistence and transformational entrepreneurship.” *Innovation policy and the economy* 10 (1):57–81.

Wooldridge, Jeffrey M. 2009. “On estimating firm-level production functions using proxy variables to control for unobservables.” *Economics letters* 104 (3):112–114.

Appendix A Additional Tables

We describe below the list of firm-level variables that we construct for firms which participate in the SINE survey. Variables related to employment and sales are constructed using the DADS (matched employer-employee) and tax files datasets, matched with the SINE survey using the unique firm identifiers.

Table A1: Description of Variables

Variables	**Description**
College Education	Dummy variable that takes value one if the firm is created by an entrepreneur with a college degree (at least two-year). *Source:* SINE.
Previously CEO-Executive	Dummy variable that takes value one if the firm is created by an entrepreneur who was previously employed as executive or CEO (before the firm creation). *Source:* SINE.
Serial Entrepreneur	Dummy variable that takes value one if the firm is created by an entrepreneur with entrepreneurial experience, i.e., who had created at least one firm. *Source:* SINE.
$\mathbb{1}$(Employment x-Y > 0)	Dummy variable that takes value one if the firm has at least one employee x years after its creation. *Source:* DADS.
Nb. Employees x-Y	Number of employees at the firm x years after its creation. *Source:* DADS.
Sales x-Y	The amount of sales (in thousands euro) of the firm x years after its creation. *Source:* Tax files.
Top 5% Nb. Employees x-Y	Dummy variable that takes value one if the firm is in the top 5% of its SINE cohort in terms of number of employees x years after creation. *Sources:* SINE, DADS.
Top 5% Sales x-Y	Dummy variable that takes value one if the firm is in the top 5% of its SINE cohort in terms of sales x years after creation. *Sources:* SINE, Tax files.

	Change in Firm Creations by Concentration Time Period			
	Top-20	Top-8	Top-4	Log HHI
	(1)	(2)	(3)	(4)
1994-2001	-1.4***	-1.1**	-1.2**	-.13**
	(.55)	(.53)	(.56)	(.055)
R^2	.061	.035	.033	.06
2001-2008	-1.2	-1	-1.2	-.038
	(.97)	(.94)	(1)	(.065)
R^2	.021	.01	.013	-.0035
2008-2015	-1.4**	-.93	-.65	-.088
	(.6)	(.56)	(.47)	(.056)
R^2	.042	.014	.0029	.016

Table A2: **The effects of concentration on creations are greatest from 1994-2001.** Cross-sectional regressions at the industry (3-digit) level for 135 industries. The dependent variable is the change in log creations at the industry level over the specified time period. Estimates are weighted by the number of firms in each industry in 1994. *Change in top X share* is the % change in the share of output of the largest X firms in an industry over the specified time period. Robust standard errors are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

Appendix B Framework

Our framework illustrates the effects of an increase in superstar firms' productivity on new business formation in an otherwise standard Melitz (2003) model.

Demand. We assume a standard CES utility function in equation (5).

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

where σ is the elasticity of substitution between varieties and Ω is the mass of available goods. This utility function yields the usual demand and expenditure for each variety ω,

$$q(\omega) = \left(\frac{p(\omega)}{P} \right)^{-\sigma} \frac{R}{P} \tag{6}$$

$$r(\omega) = \left(\frac{p(\omega)}{P} \right)^{1-\sigma} R, \tag{7}$$

where R is the nominal income and P is the aggregate price index given by

$$P = \left[\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}}. \tag{8}$$

Production. Firms face a constant exogenous death probability δ. To produce goods, they must pay a fixed cost f and a marginal cost $\frac{1}{\varphi(\omega)}$, where productivity φ is the realization of a random variable. We assume that the productivity distribution follows a Pareto distribution with a cumulative distribution function given by equation (3) and a probability density function

$$g(\varphi) \equiv \theta \underline{\varphi}^{\theta} \varphi^{-\theta-1} \text{ for } \varphi \geq \underline{\varphi}, \tag{9}$$

with positive support on $(0, \infty)$. We assume that $\theta > \sigma + 1$.[9] The Pareto assumption is standard in extensions of Melitz (2003) (e.g., Melitz and Ottaviano, 2008). It has two advantages: combined with a CES utility function it delivers closed form solutions, and it is a good approximation of the upper tail of firms' productivity in the data (e.g., Del Gatto, Ottaviano, and Mion, 2006), which is the focus of this paper.

Firms use labor as the sole factor of production. The total labor endowment is L and we normalize wages at one so that the firm's profit is

$$\pi(\omega) = r(\omega) - \left(f + \frac{q(\omega)}{\varphi(\omega)} \right). \tag{10}$$

[9] The assumption that $\theta > \sigma - 1$ is sufficient to ensure that integrals converge, but $\theta > \sigma + 1$ is needed for the mass of entering firms to be positive.

Under monopolistic competition, firms consider aggregate prices as given, which implies

$$p(\varphi) = \frac{\sigma}{(\sigma - 1)\varphi}. \tag{11}$$

Plugging (11) into (7), we get

$$r(\varphi) = \left(\frac{\sigma - 1}{\sigma} P\varphi\right)^{\sigma - 1} R. \tag{12}$$

Plugging (12) and (6) into (10), we rewrite profits as a function of productivity:

$$\pi(\varphi) = \frac{r(\varphi)}{\sigma} - f. \tag{13}$$

Equilibrium. In a stationary equilibrium, a firm either exits immediately or produces and earns the same profits $\pi(\varphi)$ given by (13) in each period. The expected value of a firm with productivity φ is

$$v(\varphi) = \max\left\{0, \sum_{t=0}^{+\infty}(1-\delta)^t \pi(\varphi)\right\} = \max\left\{0, \frac{\pi(\varphi)}{\delta}\right\}. \tag{14}$$

Firms remain on the market if and only if $\pi(\varphi) \geq 0$. Since profits are increasing in productivity, there exists a unique productivity level $\varphi^* \equiv \inf\left\{\varphi \geq 0; \frac{\pi(\varphi)}{\delta} > 0\right\}$ such that firms only stay in the market if $\varphi > \varphi^*$. The productivity cutoff φ^* implies the "Zero Cutoff Profit" (ZCP) condition

$$\pi(\varphi^*) = 0. \tag{15}$$

We denote $\bar{\pi}$ as the average profits per period for surviving firms. From (13), we know that

$$\bar{\pi} = f\left[\frac{r\left(\tilde{\varphi}(\varphi^*)\right)}{\sigma f} - 1\right]. \tag{16}$$

Plugging the definition of φ^* given in (15) into (13), we get that

$$r(\varphi^*) = \sigma f. \tag{17}$$

Plugging equation (17) into (16) and using the definition of $r(\varphi)$ given by (12) allows us to rewrite the ZCP condition as

$$\bar{\pi} = f\left[\left(\frac{\tilde{\varphi}(\varphi^*)}{\varphi^*}\right)^{\sigma - 1} - 1\right]. \tag{18}$$

Under our assumption that the productivity distribution follows a Pareto distribution (equations (3) and (9)), the ZCP is flat and average profits are independent of φ^*:

$$\bar{\pi} = \frac{f(\sigma - 1)}{\theta - \sigma + 1}. \tag{19}$$

Free entry requires the total expected value of profits to be equal to the fixed cost of entry, yielding the "Free Entry" (FE) condition

$$\bar{\pi} = \frac{\delta f}{1 - G(\varphi^*)} \tag{20}$$

$$= \delta f \left(\frac{\varphi^*}{\underline{\varphi}}\right)^{\theta} \text{ for } \varphi \geq \underline{\varphi}, \tag{21}$$

where the second row follows from our assumption that the productivity distribution follows a Pareto distribution. Assuming that $\delta < \frac{1}{\frac{\theta}{\sigma-1}-1}$,[10] the ZCP equation (19) and the FE equation (21) determine a unique $(\bar{\pi}, \varphi^*)$:

$$\bar{\pi} = \frac{f(\sigma - 1)}{\theta - \sigma + 1} \tag{22}$$

$$\varphi^* = \underline{\varphi}\left(\frac{\sigma - 1}{\delta(\theta - \sigma + 1)}\right)^{\frac{1}{\theta}}. \tag{23}$$

Figure B1 illustrates the determination of $(\bar{\pi}, \varphi^*)$ in equilibrium.

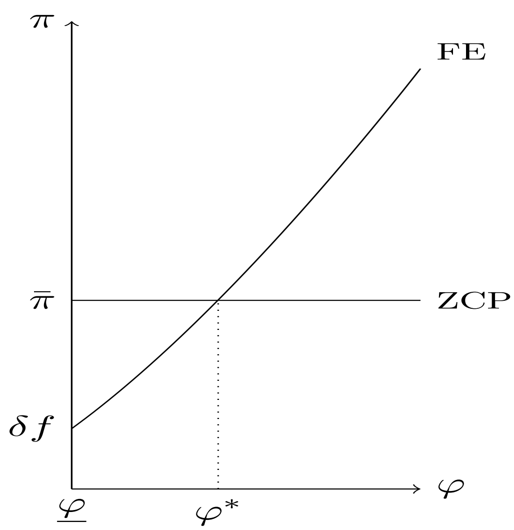


Figure B1: Determination of φ^* in equilibrium

Given $(\bar{\pi}, \varphi^*)$, there exist three types of firms in equilibrium:

1. Firms that pay the fixed cost f, draw a productivity $\varphi > \varphi^*$, and operate. We denote M the mass of surviving firms.
2. Firms that pay the fixed cost f, draw a productivity $\varphi < \varphi^*$, and exit. We denote M_E the mass of firms that survive or exit.
3. Firms that do not pay the fixed cost f.

[10] This condition is always met if $\sigma > 3$ given that $\delta < 1$ because δ is the death probability.

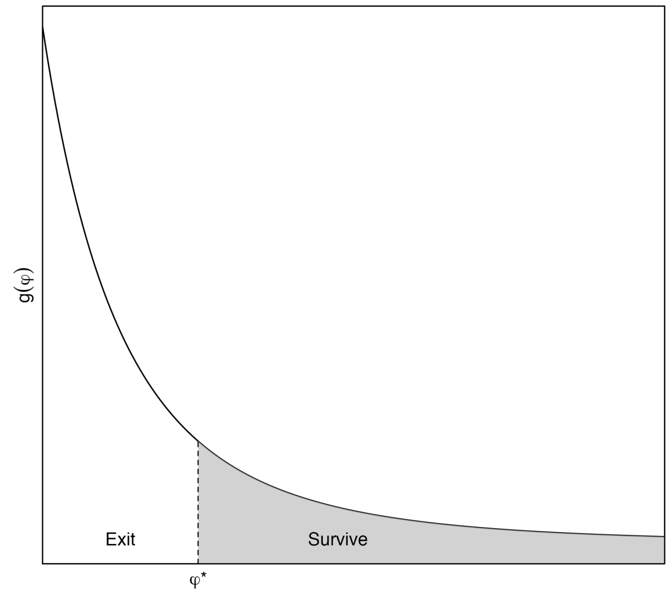


Figure B2: Productivity distribution and selection into production

Figure B2 plots the distribution of firm productivity when it is distributed Pareto (equations (3) and (9)). The gray area under the curve and above φ^* is equal to $1 - G(\varphi^*)$. Hence, the cumulative distribution of the M_E firms that pay the fixed cost f and draw a productivity $\varphi > \varphi^*$ is $\frac{G(\varphi)}{1-G(\varphi^*)}$. We denote $\mu(\varphi)$ as the conditional probability density function of surviving firms' productivity levels in equilibrium:

$$\mu(\varphi) = \begin{cases} \frac{g(\varphi)}{1-G(\varphi^*)} & \text{if } \varphi \geq \varphi^* \\ 0 & \text{if } \varphi < \varphi^* \end{cases} \tag{24}$$

Proofs. We model the increase in superstar firms' productivity as a decrease in the Pareto parameter θ, that is, a shift in the productivity distribution that benefits the most productive firms relatively more. This choice allows us to keep the model tractable and to obtain closed-form solutions.

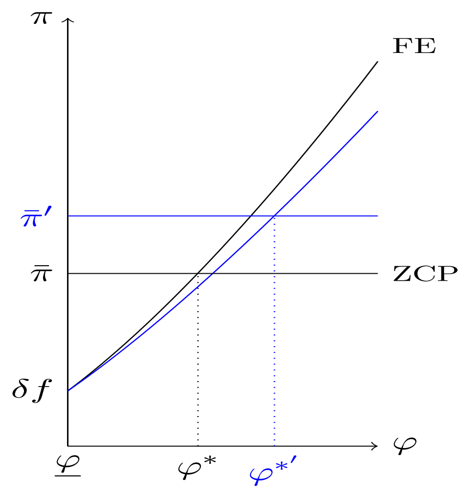
Proof of Prediction 1. The mass M_E of entrants that pay the fixed cost f is determined in equilibrium such that $M_E = \frac{R}{\bar{r}}$, where from (13) we have $\bar{r} = \sigma(\bar{\pi} + f)$. Using the equilibrium value of $\bar{\pi}$ given in (22), we obtain

$$M_E = \frac{R\left(1 - \frac{\sigma+1}{\theta}\right)}{\sigma f}. \tag{25}$$

The mass M of surviving firms represent the portion of the mass M_E of entrants that survive,

$$M = [1 - G(\varphi^*]M_E, \tag{26}$$

where both $[1 - G(\varphi^*)]$ and M_E are increasing in θ. Therefore, M is increasing in θ, such that an increase in the productivity of superstar firms (i.e., a decrease in θ) leads to a decrease in M. ■



[Increase in φ^*] [Decrease in $(1 - G(\varphi^*))$]

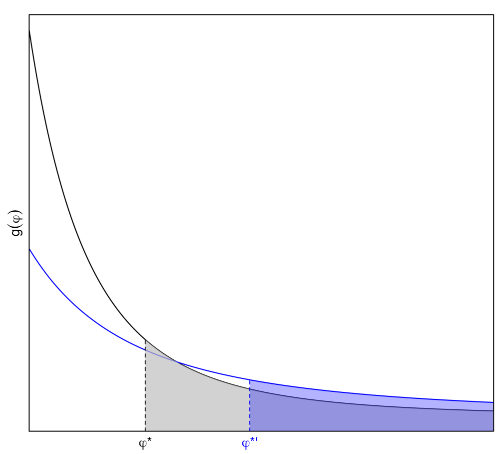


Figure B3: Effects of a decrease in θ

Figure B3 shows how a decrease in θ increases the equilibrium value of φ^* and the resulting distribution of surviving firms' productivity. The area in blue is smaller than that in gray, illustrating the result in Prediction 1 that following a decrease in θ, the fraction $1 - G(\varphi^*)$ of firms that pay the fixed cost f and operate decreases.

Proof of Prediction 2. The equilibrium aggregate productivity level writes

$$\tilde{\varphi} = \left[\int_{0}^{+\infty} \varphi^{\sigma-1} \mu(\varphi) d\varphi \right]^{\frac{1}{1-\sigma}}, \tag{27}$$

Using the definition of $\mu(\varphi)$ in Equation (24) and the fact that only the firms with a productivity level φ such that $\varphi \geq \varphi^*$ operate, we rewrite the aggregate productivity level as

$$\tilde{\varphi}(\varphi^*) = \left[\frac{1}{1 - G(\varphi^*)} \int_{\varphi^*}^{+\infty} \varphi^{\sigma-1} g(\varphi) d\varphi \right]^{\frac{1}{\sigma-1}}, \tag{28}$$

where $G(\varphi^*)$ is defined in (3). Using the value of φ^* given in Equation (23), we rewrite the aggregate productivity level as

$$\tilde{\varphi} = \varphi^* \left(\frac{1}{1 - \frac{\sigma-1}{\theta}} \right)^{\frac{1}{\sigma-1}}, \tag{29}$$

where both φ^* and the term in parenthesis are decreasing in θ. Therefore, $\tilde{\varphi}$ is decreasing in θ, such that an increase in the productivity of superstar firms (i.e., a decrease in θ) leads to an increase in $\tilde{\varphi}$.

■

Proof of Prediction 3. Only the firms with a productivity level φ such that $\varphi \geq \varphi^*$ operate. Since φ^* is decreasing in θ, an increase in the productivity of superstar firms (i.e., a decrease in θ) implies a new productivity threshold $\varphi^{*'} > \varphi^*$ such that the least productive incumbent firms (for which $\varphi^* < \varphi < \varphi^{*'}$) exit the market. ■

Proof of Prediction 4. An increase in the productivity of superstar firms (i.e., a decrease in θ) implies a new productivity threshold $\varphi^{*'} > \varphi^*$ under which firms do not enter and incumbents exit. Every surviving firm now has a productivity level φ that is higher than previously surviving firms that no longer survive (i.e., firms for which $\varphi^* < \varphi < \varphi^{*'}$). Since expenditures per variety (hence per firm) given by Equation (12) are increasing in φ, a decrease in θ increases the sales of all firms. This increase in sales, coupled with the fact that M decreases following a decrease in θ (Prediction 1), implies that the market share of each firm increases following a decrease in θ.

This paper uses several measures of market concentration: The Herfindhal index (HHI), which is the sum of squared market shares of all firms, and the share of the largest (superstar) firms' market share. Each of these concentration measures implies that market concentration increases following a decrease in θ.

■

Appendix C TFP Measurement Using the Ackerberg, Caves Frazer (2015) / Wooldridge (2009) Approach

We assume that production is given by

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it}$$

Where y is output, k is capital, m is materials, l is labor, and ω is the firm's idiosyncratic productivity. OLS estimates of omega are biased because input levels are chosen as a function of omega, creating a correlation between inputs and the error term. Our approach uses the assumptions and methodology in Wooldridge (2009). There are two essential assumptions for this approach. First, that productivity can be written as an invertible function of variable inputs (such as materials) and state variables (such as the capital stock). Second, that innovations to productivity are uncorrelated with the capital stock or with any past values of labor and materials. Under these assumptions, we can recover estimates of the ceofficients in (1) by estimating an instrumental variables specification. The reduced form specification is

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + f(m_{i,t-1}, k_{i,t-1}) + u_{it}$$

Valid instruments for l, k and m are contemporaneous and lagged values of k, and lagged values of m and l, as well as functions of these variables. In practice, we approximate $f()$ using a third-order polynomial in lagged materials and capital. Our excluded instruments are $l_(t-1)$, $m_(t-2)$ and their product. We estimation this specification using GMM. This approach follows closely the application of Wooldridge (2009) made available by Levinsohn Petrin (see Levinsohn's web site for code). The main difference is that we use intermediate inputs as the proxy variable rather than electricity. We only observe output for select industries, e.g., manufacturing. Therefore, we follow the literature and use the value of the firm's production for y, instead of measuring units of output. De Ridder et al. (2021) show, using a subset of our data, that using revenues in place of quantities leads to accurate results when studying changes in a firm's markups. Importantly, De Ridder et al. (2021) also show that assuming Cobb-Douglass production rather than a more flexible translog production will result in noise in markup estimation. In our case, this will increase the standard errors of our cross-industry estimates and make us less likely to estimate a statistically significant relationship between firm entry and industry productivity. We estimate the production function using variables Winsorized at the 1% level and limit the sample to firms with at least 50 employees in the DADS data. We allow the parameters of the production function to vary over five-year periods (1994-1999, 2000-2004, 2005-2009, 2010-1015) and by one-digit industry code. Note that production function estimation at a higher level than the main analysis is common in the literature, in order to maximize sample sizes.

Appendix D Evidence from the United States

D.1 Exhibits

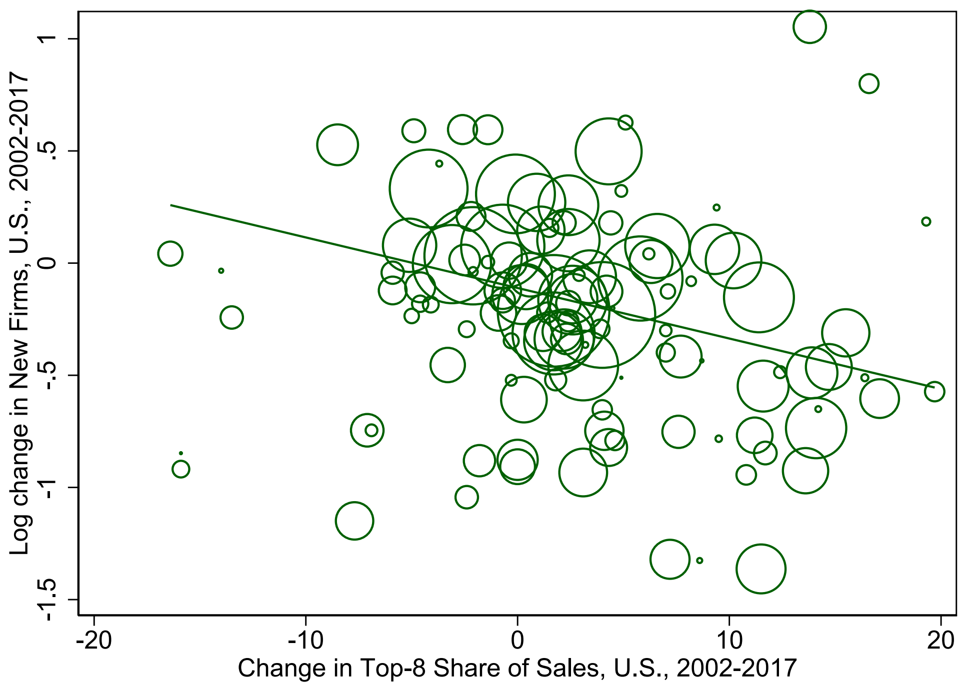


Figure D.1: **Changes in firm creation and changes in industry concentration.** The x-axis shows the change in the top-8 share of industry sales from 2002-2017. For readability, plot is limited to industries where the change in concentration is between -20 and 20. The y-axis shows the change in log new firms created from 2002-2017. Circle sizes correspond to the number of firms in each industry in 2002. The fit line is from a WLS regression using 2002 firms as weights. Source: Economic Census (concentration data); Business Dynamic Statistics (new firm data and firm count data).

Table D.1: Concentration and New Firms, US Industries

	Mean	SD	p25	p50	p75	p90	p99	N
15-Yr Change in Top-4 Share	2.55	5.57	-0.30	1.40	4.50	11.40	17.90	120
15-Yr Change in Top-8 Share	2.93	5.63	-0.30	2.20	5.80	11.50	17.10	120
15-Yr Change in Top-20 Share	3.34	5.65	0.40	2.70	5.90	11.80	18.70	120
15-Yr Change in Top-50 Share	3.95	5.38	1.20	4.50	6.10	12.20	16.20	120
15-Yr Change in New Firms	-0.18	0.39	-0.34	-0.15	0.06	0.31	0.60	120
15-Yr Change in New Firm Emp Share	-1.41	0.92	-1.85	-1.34	-0.90	-0.51	0.31	120

Appendix E The relationship between ex-ante entrepreneur quality and ex-post performance

	𝟙(Employee After 3Y)				Log(Employees After 3Y)				Log(Sales After 3Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(23)
College Education	.04***			.032***	.13***			.081***	.15***			.091***
	(.006)			(.0059)	(.022)			(.02)	(.022)			(.02)
Previously CEO-Executive		.06***		.055***		.33***		.3***		.39***		.37***
		(.0055)		(.0061)		(.019)		(.017)		(.021)		(.019)
Serial Entrepreneur			.0081	-.0038			.13***	.073***			.11***	.04**
			(.005)	(.0054)			(.022)	(.021)			(.02)	(.02)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.029	.03	.028	.031	.076	.093	.077	.096	.09	.1	.089	.1
Observations	35,050	35,050	35,050	35,050	20,158	20,158	20,158	20,158	33,273	33,273	33,273	33,273

Table E.2: **Entrepreneur Ex-ante Quality and Ex-post Performance After 3 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Employee After 5Y)				Log(Employees After 5Y)				Log(Sales After 5Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(23)
College Education	.051***			.043***	.17***			.11***	.2***			.14***
	(.0073)			(.0074)	(.025)			(.024)	(.027)			(.026)
Previously CEO-Executive		.056***		.051***		.35***		.31***		.43***		.39***
		(.0071)		(.007)		(.02)		(.017)		(.022)		(.021)
Serial Entrepreneur			.0018	-.0095			.11***	.048**			.11***	.027
			(.0085)	(.0086)			(.024)	(.023)			(.025)	(.025)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.036	.036	.034	.037	.079	.094	.076	.097	.093	.11	.091	.11
Observations	35,050	35,050	35,050	35,050	16,193	16,193	16,193	16,193	21,356	21,356	21,356	21,356

Table E.3: **Startup Ex-ante Quality and Ex-post Performance After 5 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	𝟙(Employee After 10Y)				Log(Employees After 10Y)				Log(Sales After 10Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(23)
College Education	.051***			.044***	.22***			.16***	.22***			.15***
	(.011)			(.012)	(.043)			(.044)	(.048)			(.044)
Previously CEO-Executive		.051***		.049***		.37***		.33***		.52***		.48***
		(.014)		(.016)		(.043)		(.043)		(.053)		(.044)
Serial Entrepreneur			-.008	-.02*			.13***	.047			.14***	.036
			(.01)	(.01)			(.035)	(.032)			(.046)	(.039)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.03	.03	.028	.032	.071	.083	.067	.087	.096	.11	.094	.11
Observations	17,092	17,092	17,092	17,092	5,679	5,679	5,679	5,679	10,263	10,263	10,263	10,263

Table E.4: **Startup Ex-ante Quality and Ex-post Performance After 10 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Top 5% Employees After 3Y)				$\mathbb{1}$(Top 5% Sales After 3Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College Education	.018***			.011**	.016***			.0073**
	(.0048)			(.0045)	(.0038)			(.0034)
Previously CEO-Executive		.05***		.046***		.054***		.052***
		(.006)		(.0058)		(.0053)		(.0052)
Serial Entrepreneur			.018***	.0087***			.013***	.0035
			(.0029)	(.0028)			(.0028)	(.0028)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.046	.053	.047	.054	.07	.079	.07	.079
Observations	26,704	26,704	26,704	26,704	33,858	33,858	33,858	33,858

Table E.5: **Startup Ex-ante Quality and Rank Within Cohort After 3 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	𝟙(Top 5% Employees After 5Y)				𝟙(Top 5% Sales After 5Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College Education	.022***			.013***	.02***			.012***
	(.0047)			(.0044)	(.0044)			(.0041)
Previously CEO-Executive		.055***		.051***		.054***		.052***
		(.0059)		(.0055)		(.006)		(.0061)
Serial Entrepreneur			.016***	.0059*			.0093***	-.0013
			(.0037)	(.0034)			(.0034)	(.0036)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.05	.058	.049	.058	.068	.076	.067	.077
Observations	21,835	21,835	21,835	21,835	21,835	21,835	21,835	21,835

Table E.6: **Startup Ex-ante Quality and Rank Within Cohort After 5 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	$\mathbb{1}$(Top 5% Employees After 10Y)				$\mathbb{1}$(Top 5% Sales After 10Y)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College Education	.025***			.017**	.019***			.012*
	(.0073)			(.0071)	(.0071)			(.0069)
Previously CEO-Executive		.049***		.045***		.043***		.041***
		(.0076)		(.0071)		(.0075)		(.007)
Serial Entrepreneur			.012**	.0013			.011***	.0021
			(.0057)	(.0053)			(.0037)	(.0028)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	.047	.052	.045	.053	.077	.081	.076	.082
Observations	7,603	7,603	7,603	7,603	10,536	10,536	10,536	10,536

Table E.7: **Startup Ex-ante Quality and Rank Within Cohort After 10 Years.** Cross-sectional regressions at the firm level for firms parts of the SINE cohorts from 1994 to 2014. All variable descriptions are provided in Appendix A1. All regressions include cohort, industry and zone (employment zone) fixed effects. Standard errors two-way clustered by industry and zone are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.