Demographic Origins of the Startup Deficit*

Fatih Karahan
Federal Reserve Bank of New York
fatih.karahan@ny.frb.org

Benjamin Pugsley
Notre Dame University
bpugsley@nd.edu

Ayşegül Şahin
Federal Reserve Bank of New York
aysegul.sahin@ny.frb.org

March 19, 2018

Abstract

This paper studies the causes of the declining startup rate over the past three decades. The stability of firms' lifecycle dynamics throughout this period along with the widespread nature of the declining startup rate place strong restrictions on potential explanations. We show that declines in the growth rate of the labor force explain an important share of the startup rate decline while leaving incumbent dynamics unaffected in a Hopenhayn-style firm dynamics model. Moreover, using cross-sectional demographic variation we estimate a quantitatively and statistically significant labor supply growth elasticity of the startup rate, which is robust to alternative specifications. Our findings suggest that the decline in the growth rate of the working age population through its general equilibrium effects on firm dynamics are an important driver of the decline in firm entry.

*We thank Rob Dent, Sam Kapon, Harry Wheeler and Rene Chalom for their superb research assistance. We are also indebted to Felipe Schwartzman for his insightful comments on a previous paper and suggestion to consider the effect of labor supply shocks on firm entry. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau, Federal Reserve Bank of New York or the Federal Reserve System. All results have been reviewed to ensure that no confidential information is disclosed.

1 Introduction

Recent empirical work has identified an unmistakable shift in U.S. firm dynamics since the late 1970s. Principally, the startup rate, measured as the share of new employer firms out of all employer firms, declined by about 30 percent from around 13 percent in 1979 to approximately 10 percent in 2007, before further declines in the Great Recession. Remarkably, this steady decline occurred relatively uniformly within geographic areas and narrow industry groups and without disturbing the lifecycle dynamics of incumbent firms. What explains this apparent decline in the rate of business formation is important since startups are both a vital source of job creation and productivity gains.[1] Understanding the source (or sources) of the declining startup rate is crucial to understanding whether the decline is an efficient response to technological shifts or escalating misallocation that could be eased through policy reforms.[2]

In this paper we propose and evaluate a simple hypothesis for the declines in the entry rate: changes in the growth rate of the labor supply feed back into the pace of business formation. In the U.S., the declining startup rate coincides with an abrupt decline in the growth rate of the labor force in the late 1970s. Figure 1 shows that trend labor force growth peaked in the mid 1970s. Its subsequent decline coincides with a gradual reduction in the startup rate, which is plotted against the right axis.

There are various reasons why demographic shifts would affect business formation. Most directly, an older population might be associated with a lower rate of business formation in the economy if younger workers are more likely to engage in entrepreneurial activity. However, in the U.S., the age group with the highest propensity to form businesses, ages 35-54, grows over this period so this is unlikely to fully explain the decline in startups.[3] We highlight a different channel.

Changes in the growth rate of the working age population affect the expected expansion in labor supply and could have important effects on business formation through a general equilibrium channel. Positive shocks to the labor supply put downward pressure on wages and create incentives for incumbent firms to expand, but they also create opportunities for potential entrants. Any effect on the entrant share will depend on how the shocks to the labor supply are accommodated by expanding incumbents and new firms. A lesson from models of firm dynamics, starting with Hopenhayn (1992), is that in the long run, the free entry of new firms make labor demand infinitely elastic. Ultimately, shifts in labor supply are absorbed not at the intensive margin by incumbent firms, but at the extensive margin by adjusting the quantity of entrants. Crucially, labor supply

[1] Bartelsman and Doms (2000), for example, find about 25 percent of within-industry gains in TFP occurs from new entrants in the manufacturing sector. In a structural estimation of an endogenous growth model, Lentz and Mortensen (2008) find a similar share of aggregate TFP growth due to the entry margin.

[2] For example, fixed costs of starting or running a business may have risen from increased regulations. One difficulty with many potential explanations based on costs or technology is that ceteris paribus they impact the value of an incumbent firm and thus its lifecycle dynamics, in contrast with experience in the U.S. over this period. The widespread nature of the startup rate declines poses a similar challenge.

[3] Ouimet and Zarutskie (2014) show that new and young firms hire a disproportionate share of young workers and show the share of young workers in a state is predictive of the startup rate, especially in high-tech sectors. Liang, Wang, and Lazear (2014) show that countries with older workforces have lower rates of entrepreneurship. In a recent paper, Engbom (2017) links the decline in firm and worker dynamics to the aging U.S. population.

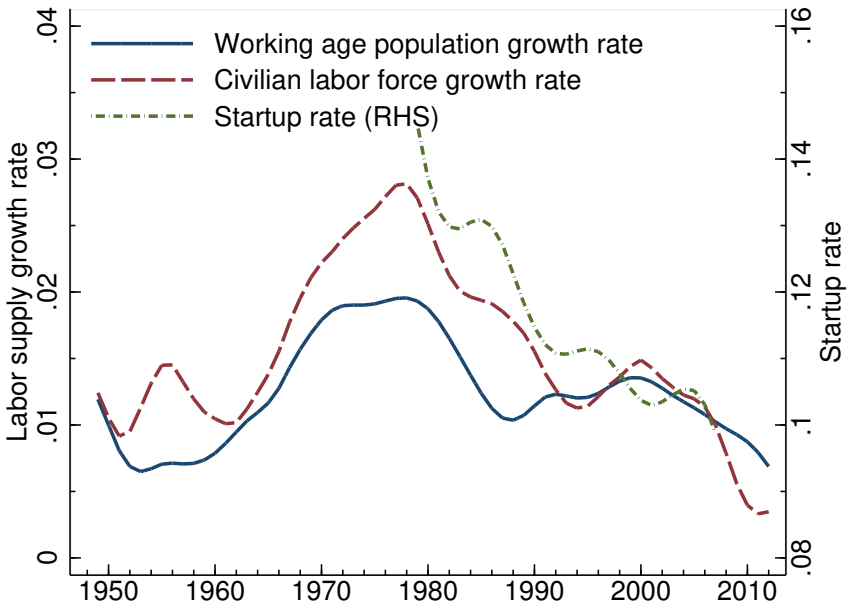


Figure 1: Trend declines of startup rate and labor supply growth coincide

Note: BLS, Annual Census Bureau population estimates, and Business Dynamics Statistics. Annual data, HP filtered with smoothing parameter 6.25. Working age population is defined here as population ages 20 to 85. Civilian labor force is measured by the BLS for the adult (16+) civilian non institutional population. Startup rate is number of age 0 employers as share of the total number of employers within a year.

changes have no direct effect on the value of an incumbent firm, except indirectly through transient effects on the real wage. A careful examination of different margins of firm dynamics since 1980 provides direct support for this channel. Despite the stark decline in firm entry we observe in the last three decades, we see relatively few changes in the size distribution of new firms as well as the dynamics of incumbents, measured by their exit rate and average employment growth by firm age and firm size.

We build on this insight and evaluate the quantitative significance of the *demographic channel* in accounting for the decline in firm entry. We do so by first setting up a Hopenhayn-style firm dynamics model with population growth. Our calibrated model solidifies the intuition that the entry margin is the most responsive margin to changes in growth rate of labor supply in the economy. When we compare the steady-state entry rates in two otherwise identical economies with different growth rates of labor supply, we find that these two economies only differ in their firm entry rates. Other margins of firm dynamics—such as the size of entrants and growth and survival rates of incumbents—remain identical. The demographic channel is also quantitatively important. A decline in the growth rate of labor supply from two percent to one percent—similar to the one we observe in the U.S. from 1980 to 2000—corresponds to a significant decline in the startup rate from 13 percent to 10 percent, explaining around two thirds of the decline.

We also use our framework to examine the plausibility of alternative hypotheses that focus on barriers to entry and stricter regulations as drivers of declining entry. While changes in the cost

structure of the economy arising from changes in entry and operating costs could have potentially caused the decline in business formation, they are not likely to be the main drivers since these changes would have affected the incumbent firm dynamics profoundly, while in the data firm life cycle dynamics have not changed. In addition, changes in the cost structure of the economy to account for the decline in firm entry would be very large in magnitude—a change for which there is little evidence in the data.

We then use cross sectional demographic variation in labor supply growth to measure its effects on firm dynamics. In particular using cross-state and industry data pooled over time, we estimate a linear model of the startup rate and several incumbent margins on two measures of labor force growth, controlling for state, time, and industry effects. To generate variation in labor force growth from fully anticipated demographic shifts we instrument with 20-year lags of each state's fertility rate. The goal is to see whether changes in the growth rate of the working-age population and lower frequency component of the civilian labor force growth, our two measures, predicted by purely demographic forces have an effect on firm dynamics.

This strategy relies on two identifying assumptions. The first and most important is the exclusion restriction that conditional on state and time effects, whatever determines the cross sectional variation in the fertility rate has no other long lasting effects that would still affect business dynamics 20 years later. We argue this is likely to be the case. Second, and less important, is that the mobility costs are large enough to prevent geographic mobility from completely equating the real wage across these segmented markets. To the extent this is not true, our estimates will understate the effect of demographics on startups and incumbent dynamics. The fertility instrument changes the age composition of the labor force in addition to capturing population growth. We further control for compositional differences across-states to ensure the effects we estimate follow from changes in the growth rate of the labor supply and not its composition.

We find that the demographic changes have a large effect on the startup rate. We estimate a startup rate semi-elasticity of labor supply growth of roughly 1 to 1.5. Given these estimates and the declines in the working age population growth and civilian labor force growth over this period in the aggregate time series, the demographic shifts explain an important share of the decline in the startup rate. Further, we find that the demographic shifts have little effect on incumbent dynamics. A demographically induced increase in the growth rate of labor supply causes a short run increase in incumbent firm size, which diminishes over longer horizons. This is consistent with the transient adjustment of incumbent firms as the entry margin fully adjusts to its long run levels.

Our framework also provides an additional time-series implication. Since the U.S. working age population grew in the 1960-80 period, the start-up rate must have increased to accommodate the new entrants into the labor force. While the main sources of firm dynamics statistics lack information on firm entry, we derive an imputation procedure disciplined by the overlapping time period after 1980 between the Business Dynamics Statistics (BDS) and County Business Patterns (CBP). The earlier data support the relationship we have established in the post-1980 data: the increase in the growth rate of the working-age population in the 1965-1980 period coincides with

increasing firm entry.

Overall, our results show that a large fraction of the declining startup rate is consistent with the demographic shifts over this same period. At first, this may be a surprising result since changes in the entry costs or other technological explanations would have a more direct effect on entry. However, the stability of the incumbent survival and growth margins combined with the widespread nature of the declines in entry point to the demographic channel. Generically, shifts in costs and technology affect incumbent dynamics through their effects on the value of a firm, whereas in standard formulations labor supply changes do not.[4] In the long run with free entry the labor supply shifts are completely absorbed by the entry margin. In the U.S., the aggregate time series and cross-state evidence strongly supports this explanation of the declining startup rate.

Related Literature Our paper is closely related to the emerging literature on the declining dynamism in the U.S. economy. Reedy and Strom (2012) first called attention to a decline in the aggregate entry rate of new employers. Using more disaggregated data, recent papers by Pugsley and Şahin (2014), Decker, Haltiwanger, Jarmin, and Miranda (2014), Hathaway and Litan (2014a), Gourio, Messer, and Siemer (2014) and Davis and Haltiwanger (2014) all document that declines in the entry rate are pervasive within geographic areas and relatively narrow industry aggregations. In addition to our paper, Hathaway and Litan (2014b) note a correlation between declining startup rates and population growth as well as business consolidation. All of these papers have also drawn attention to the relevance of the declining startup rate for the ongoing health of the labor market.

Haltiwanger, Jarmin, and Miranda (2013) have shown that new and young businesses are a key input into net employment creation. Taking the trend decline or "startup deficit" as given, Pugsley and Şahin (2014) show that both its direct effect on startup job creation and its indirect cumulative effect through shifts in the employer age distribution partly explain the emergence of slower employment recoveries with each business cycle. Startups are also a source of aggregate productivity gains from the advances in technology embodied in the new firms. A slowdown in business entry may portend an overall slowing of employment and labor productivity growth. Changes in laws and regulations, market concentration, education and licensing requirements, and shifts in economies of scale might discourage firm entry by creating higher barriers to enter and/or a higher fixed cost of operating. Noting this possibility, Davis and Haltiwanger (2014) among others highlight the introduction of several forms of labor market and occupation regulations in the U.S. that are a potential source of these slowdowns in the entry margin.

Without offsetting effects, these individual changes to the economy are difficult to reconcile with the empirical evidence. We show generically that changes to costs and technology will change the value of an operating firm and by extension its lifecycle dynamics. However, as we show that both in the time series and the cross section, conditional on age, firm dynamics have been unaffected by the net forces causing the decline in the startup rate. This incumbent stationarity, along with the widespread nature of the entry declines, leads us to consider changes to the supply of labor

[4]With a frictional labor market, decreasing returns in the aggregate matching function would be one example where changes in the labor supply would effect the value of the firm through their effects on the match surplus.

input as a potential general equilibrium source of the decline.[5] Specifically we study the role of slow moving demographic shifts in driving the decline in startups.

2 Data description

To support our hypothesis, we use data on firms and labor market demographics from several sources. We summarize these here and provide additional details in our online data appendix A.

2.1 Measuring firm dynamics

We use firm-level data on employer businesses from the U.S. Census Bureau Longitudinal Business Database (LBD) and its public-use tabulations, the Business Dynamics Statistics (BDS). The LBD microdata cover all nonfarm private-sector employer establishments in the U.S. starting in 1976, and the database is based on a longitudinally-linked version of theCensus Bureau's Business Register, whicn includes all private-sector establishments with paid employees. [6] An establishment is a physical location of business activity, and when multiple establishments belong to the same firm, they can be grouped by a shared firm identifier.[7] This is an important detail, since we are interested in the decline of new firms than new locations (new establishments) of an existing firm. We construct firm-level measures by aggregating across establishment microdata.

Startups and incumbent age The first key measure is firm age, which we use both to identify startups and to distinguish young and old incumbent firms. To be consistent with the BDS, we assign firm age as the age of each firm's oldest establishment, where an establishment enters in the year it first reports employment and ages naturally thereafter (regardless of any ownership changes). New firms or *startups* are age 0 firms. This measure of firm entry is robust to mergers and other reorganizations: any age 0 firm is a bona fide new entrant since the firm is composed entirely of one or more age 0 establishments. This is also the measure of firm age and entry popularized in Haltiwanger, Jarmin, and Miranda (2013). Incumbents are all other firms, age 1 and higher. One drawback of the measure of firm age is that birth year is left censored for any firms founded prior to 1977, when new establishments may first be observed. These firms are part of the database but their age can only be bounded below. For example, we can infer starting in 1981 that the group of left censored establishments is at least 5 years old. We split incumbents into two age groups: young and mature. The young age group contains firms ages 1 to 10, and the mature age group contains all other firms, ages 11 and higher. Because our definition depends on identifying age 10 firms, our measures of incumbent firms by age group begin in 1987.

[5] We thank Felipe Schwartzman for his insightful suggestion to consider this margin.

[6] Jarmin and Miranda (2002) provide a detailed description of the linking procedure and the construction of the LBD.

[7] The Census Bureau identifies the boundaries of firms through the annual Company Organization Survey and quinquennial Economic Census as the highest level of operational control of an establishment.

Employment, growth and exit The LBD contains establishment-level employment reported for the week containing March 12 of each calendar year. To measure firm-level employment, we aggregate across the one or more establishments within each firm. Firm-level employment growth is measured as the employment weighted average of establishment-level employment growth, grouped by the current year firm identifier.[8] To be consistent with the BDS measure of firm death, we measure firm exit when all of its establishments have 0 employment and are reported closed. This is a more conservative measure of exit, which by construction, counts among survivors firms that temporarily have 0 total employment.

Aggregating by industry and geography We tabulate the firm-level measures by both state and 4 digit NAICS industries, individually and jointly.[9] The *startup rate* is measured as the number of age 0 firms within the cell as a fraction of the total number of firms. The survival rate is measured as 1 minus the number of firm exits as a fraction of the total number of firms within cell, and the employment growth rate is the employment-weighted average of the firm-level growth rate. public-use tabulations to construct these measures are available in the BDS by state, but not by industry or by state $\times$ industry jointly. To encourage replicability, we use the publicly available BDS-based measures wherever possible and supplement our analysis with the LBD-based measures when necessary.

2.2 Measuring labor market and population demographics

Our analysis links firm demographics to population demographics, and to do this we construct measures of worker and population demographics over time both by state and the U.S. overall. Our central focus is the growth rate of the labor supply, which we estimate using two alternative proxies.

Working age population Our first measure of labor supply growth is the growth rate of the working age population, which we define to be adults ages 20 to 64, which corresponds to age range with the highest labor force participation. This definition is slightly more expansive than the customary definition of prime age workers of ages 25 to 54, where labor force participation is both high and relatively constant. Historically, participation among ages 20-24 and 55-64 is slightly lower than prime-age, and falls off steeply outside of these age ranges.

We construct the growth rate of the working-age population using annual Census Bureau population estimates by age group. These annual data are based on the decennial population census and intercensal estimates formed using data on births, deaths, and migration.[10] We sum the an-

[8] This is equivalent to measuring the growth rate of firm employment, but it is robust to changes in firm identifiers across years for technical reasons. We provide additional discussion of our construction of these longitudinal measures in the data appendix section A.1.1.

[9] We describe the procedure to assign firm-level measures of state and industry in data appendix sections A.1.2 and A.1.3, respectively.

[10] The Census Bureau annual population estimates and a description of the the estimation methodology are available from https://www.census.gov/programs-surveys/popest.html.

nual estimates by age group to estimate the population ages 20-64 and then take the one year growth rates. This is a benchmark measure of the growth rate of the working age population. We also compute the growth rate over alternative definitions of the working age population, which we examine in our robustness appendix section B.2.4.

Civilian labor force Our second measure of labor supply growth is based on the growth rate of the adult civilian labor force. While the working age population includes all individuals ages 20 to 64, the adult civilian labor force (CLF) includes all adults, 16 or older, who are either currently employed or actively searching for a job. At the national level the CLF is estimated monthly by the BLS using the Current Population Survey (CPS), and for states the labor force is estimated as part of the Local Area Unemployment Statistics program, which combines the CPS with information from state-level unemployment insurance programs, the BLS establishment survey, as well as local population estimates from the Census Bureau. We average the monthly estimates of the CLF by year, and then take one year growth rates. Since, even at an annual frequency the CLF is procyclical, see for example Elsby, Hobijn, and Şahin (2015), we also use a version of the CLF growth rates purged of business cycle fluctuations using an HP filter with a smoothing parameter of 6.25 as recommended by Ravn and Uhlig (2002).

2.3 Labor supply instruments

To generate plausibly exogenous variation in our measures of labor supply growth we describe in section 5.1 an instrumental variables strategy that relies on cross-state differences in lagged fertility and migration patterns. We describe the data used to construct these instruments.

Fertility State-level annual birthrates are tabulated from the Natality Data from the National Vital Statistics System of the National Center for Health Statistics. For each state and year we measure the number of births per 1000 adults, which is known as a "crude birth rate." We compute these annual state-level birth rates for the years 1960 to 1987, which we will use with a 20 year lag to predict labor supply growth for the years 1980 to 2007. Even conditional on permanent differences across-states there is considerable variation birth rates. In data appendix section A.3 we include a set of figures that plots the residual variation in these birth rates.

Migration Our migration instrument relies on measuring the distribution of birthplaces of individuals within a state. The long form of the Decennial Census (until replaced by the annual American Community Survey in 2001) asks respondents for the state of birth of each person in the household. It also asks for the address of each person 5 years ago, if different from the current residence. We use the IPUMS microdata, see Ruggles, Genadek, Goeken, Grover, and Sobeck (2017), for the long form responses to the 1970, 1980, 1990 and 2000 Decennial Censuses.[11] For each Decennial census and state, we aggregate over native-born persons in that state to estimate

[11] We use the 5% microdata samples for all years except 1970, where we use the 1% Form 1 metro sample.

the distribution of birth states. We also do the same using the measure of location 5 years ago on each census long form.

2.4 County Business Patterns

We supplement the LBD and the BDS, which begin in 1977, with historical data from the Census Bureau County Business Patterns. The County Business Patterns (CBP) provide annual tabulations of the stock of establishments and employment by state and industry extending back to 1964.[12] We provide additional details on the historical CBP data in data appendix section A.4. There are a number of limitations that prevent us from using these data for our main analysis. The two most important for our purposes, are that CBP are static cross sections of establishments, with no longitudinal links to infer establishment age, and that the CBP tabulates establishments rather than firms. Nevertheless, as we describe in section 6 we can use the CBP data to impute a longer time-series on the startup rate.

3 Linking trends in firm dynamics with worker demographics

Our analysis focuses on the period 1980–2007, prior to the Great Recession and its slow recovery. This period features a steady decline in the firm startup rate that coincides with a parallel slowdown in the growth rate of the labor *supply*. We begin this section with a simple example of why these two trends may be related. With this example in mind, we then document several aggregate facts that will discipline a standard equilibrium model of firm dynamics extended to incorporate demographic change.

3.1 A motivating example

To see why changes to the growth rate of the labor supply might affect the firm entry rate, it is helpful to start with the following simple example. Consider an economy with measure μ_t of identically sized firms in year t, where, over the year, measure M_t firms enter. We would define the "startup rate" for this economy as $SR_t \equiv M_t/\mu_t$. Exit is exogenous at rate x_t. The economy is growing with the labor supply N_t, which is expanding exogenously at rate η_t. At full employment, the measure of firms per worker $\bar{\mu}_t \equiv \mu_t/N_t$ evolves as

$$\bar{\mu}_{t+1} = \frac{(1-x_t)}{1+\eta_t}\bar{\mu}_t + \bar{M}_t.$$

The number of firms per worker declines with increases in firm exit, x, and faster growth in the number of workers, η. Existence of the balanced growth path with a constant measure of firms per worker $\bar{\mu}$ requires that increases in labor demand equal the increases in labor supply, which is

[12] The CBP program provides data as early as 1946 at roughly triannual frequencies. In these years, establishments may also be combined within county and detailed industry. See https://www.census.gov/programs-surveys/cbp/technical-documentation/methodology.html for additional details.

	Labor Supply GR(%)				Predicted SR (%)	
	WAP	CLF	Exit Rate (%)	SR (%)	η =WAP	η=CLF
1980 to 1984	1.67	2.07	9.43	12.51	10.91	11.26
2003 to 2007	1.11	1.26	8.47	10.37	9.48	9.60
Decline	0.56	0.81	0.96	2.14	1.43	1.66

Table 1: Actual and flow balance startup rates in 1980-1984 and 2003-2007 periods.

implied here by the constant average firm size. From the law of motion, we can determine that along such a balanced growth path, the startup rate must equal

$$SR = \frac{\eta + x}{1 + \eta}. \tag{1}$$

We refer to this as a "flow balance" startup rate, because it equates the inflow of new firms per worker with the outflows per worker, where outflows are determined both by the exit rate and by the growth in the number of workers. With a constant exit rate, the faster the labor supply expands (increases in η) the higher the required startup rate. In this example, as labor supply expands, the additional labor demand to clear markets must come from *new* firms. Said differently, the entry margin makes labor demand infinitely elastic. This feature follows immediately in this simple example because of the constant firm size as well as the constant exit rate, but as we will show in section 4 it extends to a more realistic equilibrium setting, where we relax these rigid assumptions on size and exit.

In spite of its simplicity, the simple example does capture quantitatively some of the decline in entry.

In table 1, we use equation (1) to compute the flow balance startup rate for the early 1980s and the mid 2000s. We first report the average labor force growth, columns 1 and 2, as well as the average exit rate, column 3, for those periods that we use in the calculation. Next we show the actual average startup rate and then the flow balance startup rate computed using either the WAP or CLF measure of labor force growth. The actual startup rate declined from its 1980-84 average of 12.5 percent to an average of 10.4 percent for 2003-07, a decline of 2.1 percentage points. For the same period, the flow balance startup rate from equation (1) using WAP (CLF) declines from 10.9 (11.3) percent to 9.5 (9.6) percent, a decline of 1.5 (1.7) percentage points and, averaging across these measures, explains roughly 3/4 of the decline.[13]

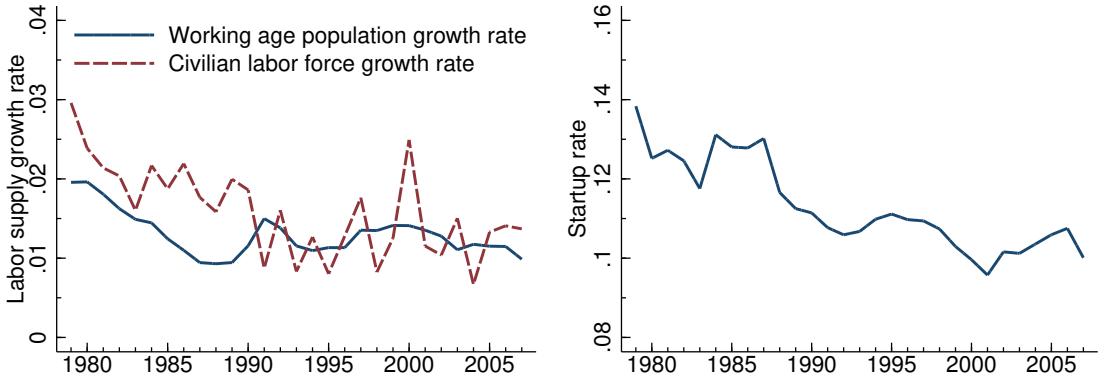
There are a number of limitations with this simple example: (i) the exit rate is constant, (ii) firm size is fixed and there is no incumbent employment growth, and (iii) the flow balance startup rate for each period is only valid in the long run. In practice, there is substantial heterogeneity

[13]Note that the decline in the exit rate is consistent with the stability of exit rates *conditional* on age and size. As shown in Pugsley and Şahin (2014), the startup deficit caused a shift in the firm age distribution towards older firms. Due to this compositional shift, the aggregate exit rate trended down.

in firm dynamics, exit and employment growth may change, and the transition dynamics may be slow. In section 4, we relax these strong assumptions to extend a standard model of firm dynamics to incorporate the effects of demographic change. Through the model we can examine the short- and long-run effects of a demographic shock of a similar magnitude. Importantly, we can assess not only the effects of a demographic shock on entrants, but also its effects on incumbents of varying age and size. In order to judge the model's plausibility, we first document the time series patterns for each of these margins as well as the demographic shock.

3.2 Trends in labor supply growth and firm dynamics

We document the trends in labor supply growth and firm dynamics since 1979. These serve both to characterize the shape and magnitude of the demographic changes and to provide empirical restrictions on the calibrated model. We show a relative steep decline in two measure of labor supply growth, which coincide with a decline in the startup rate over the same period. Remarkably, we see relatively few changes in the size distribution of new firms as well as the dynamics of various types incumbents, measured by their exit rate and average employment growth by firm age and firm size.



(a) Labor supply shifts

(b) Startup rates

Figure 2: Declines of labor supply growth and the startup rate from 1979 to 2012

Note: Census population estimates, Current Population Survey and Census BDS. Startup rate is number of age 0 employers as share of the total number of employers within a year. Working age population is defined as the noninstitutional population ages 20-64. Civilian labor force is the non institutional population ages 20-64 who are presently employed or unemployed.

Labor supply growth and the startup rate For the period 1980-2007 Figure 2a shows that this long-run decline has coincided with a shift in two supply side factors. The solid line shows that over the same period the growth rate of the labor force has declined from around two percent to about 0.8 percent. Similarly, the growth rate of the working age population fell from over two percent to a low of just over one percent. Figure 2b shows the startup rate, which is the number of newborn firms as a fraction of the overall stock of firms, for the period 1977 to 2007. The startup

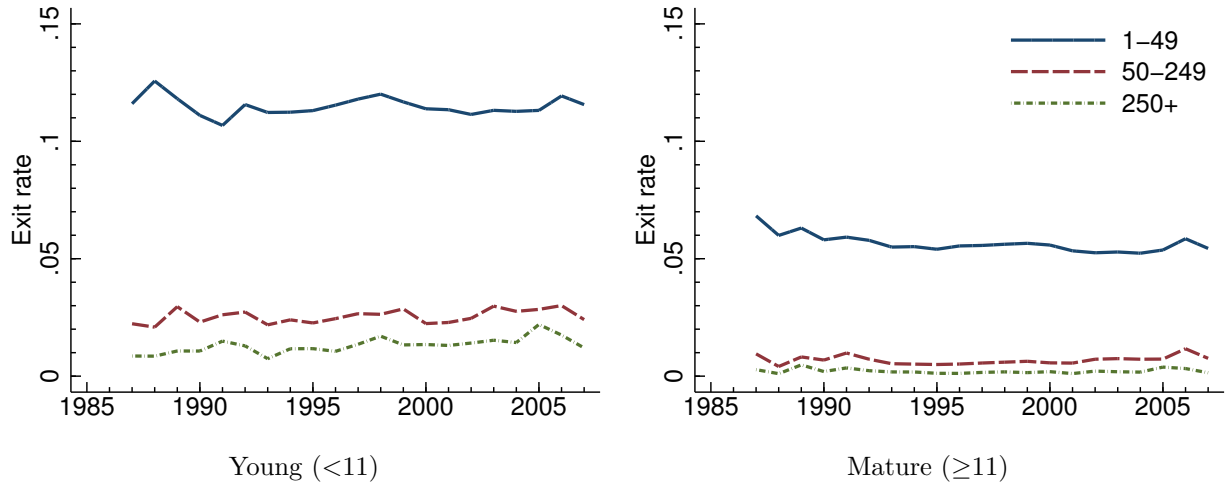


Figure 3: Incumbent exit rates by firm size for young and mature firms

Note: BDS. Exit rate is the number of incumbent firm exits as a fraction of the the number of firms in the same size and age cohort for the previous year.

rate has declined steadily from an average of roughly 13 percent in the early 1980s to around 10 percent in 2007, before the onset of Great Recession. As we show in the appendix, it continues its decline and eventually to 8 percent by 2012.

Changes in incumbent firm dynamics Firm dynamics are determined not only by entry, but also the survival and growth of incumbent firms. The determinants of the declines in the entry rate could affect incumbents directly or indirectly by changes in the selection of new entrants (that become incuments). One way to assess any effects on incumbents is to condition on both size and age and examine any changes over time.

We consider two distinct metric of performance to shed light on this question. In particular, we report firm exit rates and conditional employment growth rates by firm size and age. The conditional growth rate is defined as the change in average firm size. Defined in this way, the product of 1 minus the exit rate with the conditional growth rate is the net employment growth rate of that age and size group cohort.[14] Figure 3 plots the exit rates of firms by three size categories: 1–49, 50–249 and 250+ for firms younger than 11 years old (*young*) and for those older than 11 years old (*mature*) firms. While there are business cycle fluctuations in survival rates, these rates have remained surprisingly stable in a period of declining entry.

While exit appears little changed, the dynamics of firms could still change through their patterns of employment growth. One immediate question is whether the decline in firm entry affected the quality of entrants. Given that the entry rate declined over time, it is reasonable to expect that entrants have become better over time in terms of their post-entry performances. Figure 4 plots the employment growth of firms for the same size categories. Aside from the large swings associated with the business cycles, there is no evidence to support the idea that entrants in later years

[14]See the discussion in appendix A.2 for additional details.

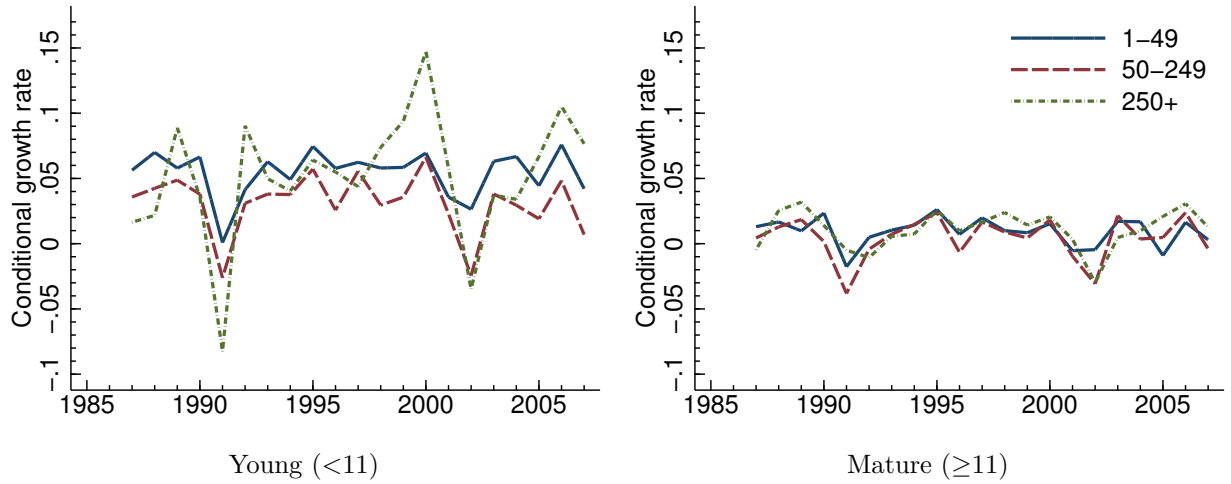


Figure 4: Incumbent employment growth rates by firm size from 1980 to 2012 for young and mature firms

Note: Business Dynamics Statistics. Incumbent conditional growth rate is the growth rate of average employment size within each size group (excludes all startups) relative to the average employment of the size group cohort in the previous year.

performed better. As we explain later on, these facts help us discipline our structural model by providing additional margins of variation. More importantly, they provide valuable information in telling apart the relative contributions of labor supply shifts from shifts in operating and entry costs to the decline in the startup rate.

4 Model of firm dynamics with labor supply growth

In this section, we construct an equilibrium model of firm dynamics with endogenous entry and exit with labor supply growth building on the workhorse Hopenhayn and Rogerson (1993) model. We calibrate the model to match several salient features of the data to evaluate the quantitative importance of different sets of explanations. Specifically, we entertain three factors as possible drivers of declining entry: changes in the cost of firm entry, changes in operating and adjustment costs and changes in labor supply growth rate. The first two factors capture the various frictions that are associated with starting and running a business. First, new firms face barriers to entry that we model as *entry costs*. Second, running a firm is subject to a fixed cost and a variable cost in case the firm needs to adjust its size. Lastly, we consider changes in the growth rate of labor supply and show that a lower labor supply growth results in lower firm entry. While each factor can generate a decline in entry by itself, each has a different implication for the behavior of incumbent firms. This restriction allows us to test and quantify the role of each factor.

4.1 Model environment

Our model economy consists of a continuum of firms that operate in a closed economy, owned by a growing population of households. Time is discrete. There is no aggregate uncertainty.

Households The economy consists of a representative household who is growing in size. It has one unit of time per member, which it supplies inelastically. Its preferences over per-capita consumption, c_t, are given by

$$\sum_{t=0}^{\infty} H_t \beta^t \log c_t \,, \tag{2}$$

where β is the discount factor and c_t is the per-capita consumption. The household has access to a one period bond that has a return of r_{t+1}. Bond holdings in period t are denoted by b_t. The population or household size is given by H_t, and it grows at a constant rate η.

The price of labor is given by w_t. We use consumption as the numeraire, so that the price of the consumption good is normalized to 1. Firms are owned by the households and all profits are distributed immediately as an aggregate dividend π_t to each household. The population of measure H_t chooses consumption c_t and savings b_t to maximize (2) subject to the following budget constraint

$$c_t + b_{t+1} = (1 + r_t) b_t + w_t + \pi_t \,. \tag{3}$$

Incumbent firms The economy is populated by a continuum of firms that use labor as the only input to produce the consumption good. Each firm has access to a decreasing returns to scale technology

$$f(\alpha, s_t, n_t) = e^{\alpha} s_t n_t^{\theta}. \tag{4}$$

Here, α denotes the fixed component of a firm's productivity and s_t denotes the idiosyncratic component, which evolves exogenously according to an AR(1) process:

$$\log s_t = \rho s_{t-1} + \varepsilon_{t+1} \tag{5}$$

where ε is normally distributed with mean zero and standard deviation σ_s.

Firms are required to pay a time-invariant fixed cost c_f for each period that they operate. In addition, changing its employment size from n_{t-1} to a different level n_t requires the firm to pay an adjustment cost of $\Phi(n_{t-1}, n_t)$, denominated in units of output.

Firms exit the economy in two ways. They may exit exogenously with probability δ, or by choice, if the value of remaining operational becomes negative. This exit decision takes place at the end of a period, after production and paying the fixed cost in period t, and in advance of learning next period's productivity s_{t+1}.

We now describe the problem of an incumbent firm. The value V_t of operating a firm with

productivity s_t and employment n_{t-1} is given by the following Bellman equation

$$V_t(\alpha, s_t, n_{t-1}) = \max_{n_t} f(\alpha, s_t, n_t) - c_f - \Phi(n_{t-1}, n_t) - w_t n_t + \frac{1-\delta}{1+r_t} \max\{0, E_t V_{t+1}(\alpha, s_{t+1}, n_t)\} \ . \quad (6)$$

The inner maximization reflects the end of period exit decision. Let $X_t(\alpha, s_t, n_{t-1})$ be the optimal exit decision of a firm that starts the period with productivity s_t and size n_{t-1}, which takes the value 1 if the firm decides to exit and 0 otherwise. Similarly, let $g_t(\alpha, s_t, n_{t-1})$ be the decision for firm's optimal size, n_t.

Entry New firms can enter the economy by paying an entry cost of c_e units of output. Upon entering, firms draw the fixed component of productivity α from the distribution F and the idiosyncratic component from the distribution G and start operating in the same period. We assume that in every period there is a supply of potential new entrants. This implies that the mass of entrants M_t is determined in equilibrium by the free entry condition:

$$c_e = \int V(\alpha, s, 0) F(d\alpha) G(ds) \ . \quad (7)$$

Entrant and incumbent dynamics Let $\mu_t(\alpha, S, N)$ be the measure of firms producing in period t with size $n \in N$, idiosyncratic productivity $s \in S$ and fixed productivity α. This measure of firms includes the new entrants M_t that start out with productivity α and $s \in S$, and choose to operate with size $n \in N$, as well as incumbent firms that decided to remain in business at the end of the previous period. The measure of firms evolves according to

$$\mu_t(\alpha, S', N') = \underbrace{(1-\delta) \int_{g(\alpha,s,n)\in N'} \mathbf{1}\{X_{t-1}(\alpha, s, n) = 0\} G(S'|s) \, d\mu_{t-1}(\alpha, s, n)}_{\text{incumbents}} + \underbrace{M_t F(\alpha) G(S') \mathbf{1}\{g(\alpha, s', 0) \in N'\}}_{\text{new entrants}} . \quad (8)$$

where G is the conditional distribution for productivity implied by the stochastic process for idiosyncratic productivity s_t in equation (5).

Aggregation Given the distribution of firms, the aggregate supply of goods net of entry, operating and adjustment costs is given by

$$Y_{st} = \int \Big(f(\alpha, s_t, n_{t-1}) - c_f - \Phi(n_{t-1}, g_t(\alpha, s_t, nt-1)) - c_e M_t \Big) d\mu_t(\alpha, s_t, n_{t-1}) \quad (9)$$

Similarly, total profits that accrue to households as dividend payments are given by

$$\Pi_t = \int \Big(f(\alpha, s_t, n_{t-1}) - n_{t-1} - c_f - \Phi(n_{t-1}, g_t(\alpha, s_t, n_{t-1})) - c_e M_t \Big) d\mu_t(\alpha, s_t, n_{t-1}) \quad (10)$$

Aggregate labor demand is given by $L_t = \int g_t(\alpha, s_t, n_{t-1}) d\mu_t(\alpha, s_t, n_{t-1})$. The per-capita counterparts, y_{st}, π_{st}, and l_t, are simply given by the ratio of the aggregates to population H_t.

Balanced growth path

Definition 1. An equilibrium is a sequence of per-capita consumption c_t, firm size and exit decisions g_t and X_t, per-capita profits π_t, wages w_t, interest rates r_t, a measure of new entrants M_t, and a measure μ_t of firms over productivity and size, such that

1. Given $\{w_t, r_t, \pi_t\}_{t=0}^{\infty}$, household consumption and bond holding decisions $\{c_t, b_{t+1}\}$ maximize (2) subject to (3).
2. Given $\{w_t, r_t\}_{t=0}^{\infty}$, firms' size and exit decisions $\{g_t, X_t\}$ solve (6).
3. The sequence of wages w_t satisfies the free entry condition in (7) in all periods.
4. The goods market clears ($y_{st} = c_t$), the labor market clears ($l_t = 1$), and the bond market clears ($b_{t+1} = 0$).
5. The distribution of firms, μ_t evolved according to the law of motion dictated by (8)

A balanced growth path equilibrium is an equilibrium in which prices w_t, r_t, per-capita consumption and profits c_t, π_t, firm decisions g_t^h and X_t^h, per-capita entry rate m_t, and the resulting distribution μ_t are time invariant.

4.2 Computation and calibration

We calibrate our model to match various key statistics of firm dynamics assuming that the economy is in its steady state in early 2000s. The model period is assumed to be a year.

Functional forms and distributions Startups draw their fixed effect α from a normal distribution with mean zero and standard deviation σ_α. The idiosyncratic component of log productivity, $\log s$, is drawn from a normal distribution with mean μ_0 and standard deviation σ_0; i.e. $G = \ln \mathcal{N}(\mu_0, \sigma_0)$. We discretize the process for s using the Tauchen procedure on 39 grid points and obtain the transition matrix G.

Cost for adjusting labor takes the following form:

$$\Phi(n_{t-1}, n_t) = \gamma \left(\frac{n_t}{n_{t-1} + 1} - 1 \right)^2. \tag{11}$$

Parameters Some parameters are set following the literature without solving the model. We set the time discount rate to 0.96. This corresponds to a steady state interest rate of around $\beta^{-1} - 1 = 4.17\%$. We set the curvature parameter of the production function θ to match the labor's share of total revenue. This requires setting θ to 0.64. Population growth rate is set to 1%, its level in early 2000s.

The remaining nine parameters are calibrated within the model. These nine parameters are: the entry cost c_e, the fixed operating cost c_f, the adjustment cost parameter γ, the exogenous exit rate δ, the productivity parameters σ_α, ρ and σ_ε and the parameters governing the distribution of startups μ_0 and σ_0. The targeted moments are reported in Table 2.

Moment	Data	Model
Startup rate	10.13	9.59
Average startup size	6.00	5.89
Average incumbent size	20.68	20.71
Exit rate by size		
0-49	8.66	9.08
50-249	1.71	1.35
250+	0.86	1.35
Conditional growth by size		
0-49	1.90	1.91
50-249	0.40	0.40
250+	-0.90	-0.87

Table 2: Targeted and model implied moments.

We estimate the values of these parameters by minimizing the sum of squared distance between a set of targets in the data and their model counterparts. We target nine moments regarding firm dynamics in the early 2000s. Specifically, we target the startup rate, the average size of startups and incumbent firms, and the exit and conditional growth rates of firms in three size categories: 0–49, 50–249, 250+. Table 2 summarizes the moments we target along with the model implied statistics and Table 3 shows the estimated values of parameters.

Parameter	Value
Entry cost, c_e	1.7
Operating cost, c_f	1.1451
Adjustment cost, γ	0.0005
Exogenous exit rate, δ	1.35%
Std. deviation of fixed productivity, σ_α	0.0189
Persistence of idiosyncratic shocks, ρ	0.9941
Std. deviation of idiosyncratic shocks, σ_ε	0.0006
Mean productivity of startups, μ_0	-0.5838
Std. deviation of startup productivity, σ_0	0.1776

Table 3: Internally calibrated parameters

4.3 Implications of changing cost structure and demographics in the long run

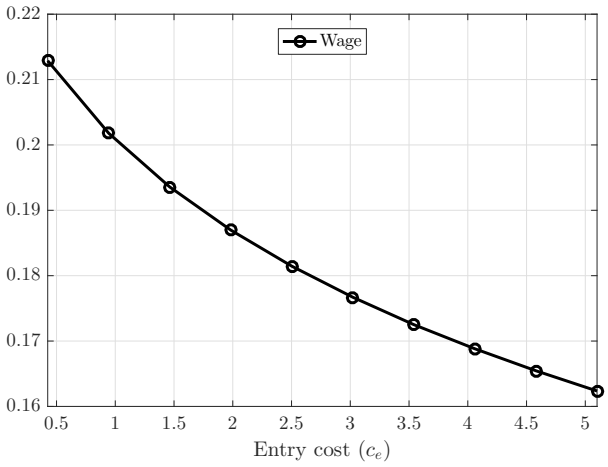
Firm entry can decline because of changes in entry, operating and adjustment costs, as well as due to changes in population growth. A key question is how to disentangle the different sources of declining entry. Using the calibrated model, we first show which of these factors can generate declines in entry and then study the implications of those factors for other margins of firm dynamics such as the average size of startups, and the behavior of incumbent firms.

For each factor that can generate declining entry, we compute the balanced growth path equilibrium for a range of values around the calibrated parameter governing that factor. Figures 5–8 plot the long-run response of the economy to changes in these parameters. In each figure, the two top panels (a and b), show the equilibrium wage and the startup rate, respectively. The bottom left panel (panel c) shows the average startup size and the bottom right panel (panel d) shows the exit and and conditional growth rates of small firms (fewer than 50 employees) that are age 2.[15]

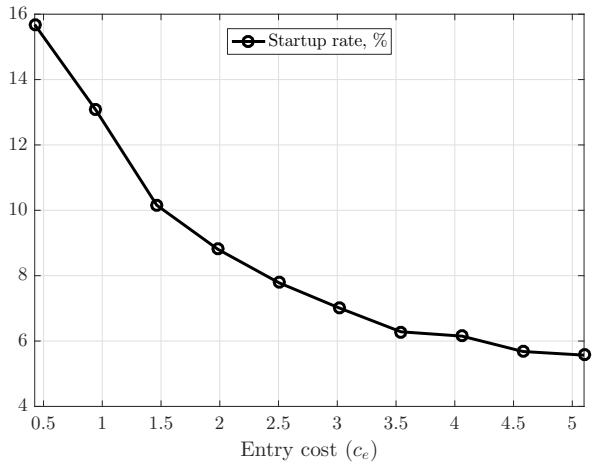
Changes in the entry cost We start by considering change in the cost structure as a potential explanation for declining entry. A common intuition is that stricter regulations make it harder for new businesses to enter. Our model captures this type of friction as an increase in *entry costs*. Figure 5 shows the changes in key firm dynamics in our economy in the long run when entry costs increase. Free entry pushes down the equilibrium wage. With lower cost of labor, new firms find it optimal start larger. Lower wages also help incumbent firms to survive negative productivity shocks better thereby lowering the exit threshold for productivity. Weaker selection results in a lower exit rate and a lower growth rate for surviving firms. The decline in the exit rate of incumbent firms results in a lower entry rate. Note that one can attribute a low entry rate to higher entry costs. However, the change in entry costs lead to an increase in the startup size and a decrease in the exit and growth rates of incumbent firms.

Changes in the operating costs Figure 5 conducts a similar exercise for operating costs, c_f. At a fixed wage, an increase in operating costs pushes up the productivity threshold for remaining in business. This change increases exit rates of existing firms and reduces the value ofoperating a firm. Free entry requires a lower equilibrium wage, resulting in larger startups on average. Lower wages counteract the direct effect of higher operating costs on exit, but are not enough to overturn it quantitatively. Stronger selection of firms implies that surviving firms grow faster. The startup rate increases in response to the higher exit rate. Our results suggest that for operating costs to be the main factor behind declining entry, they must have gone *down* over time. This stands in contrast to the conventional wisdom that emphasizes increasing regulatory costs of running a business in the U.S. The equilibrium effect is key for the result: The entry margin responds to, among other things, changes in the exit rate of firms.

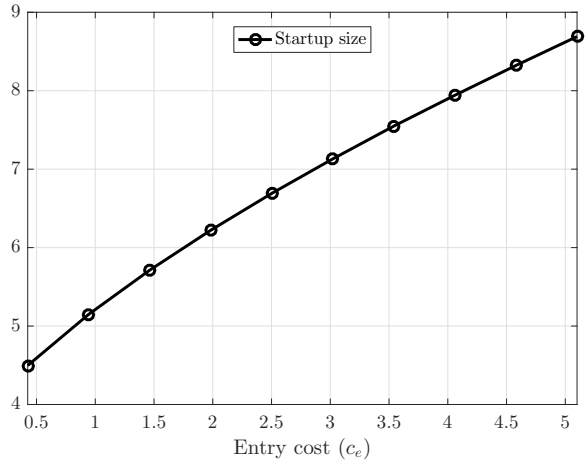
[15] We only plot the behavior of small age 2 firms, but the qualitative predictions are similar for other age and size categories.



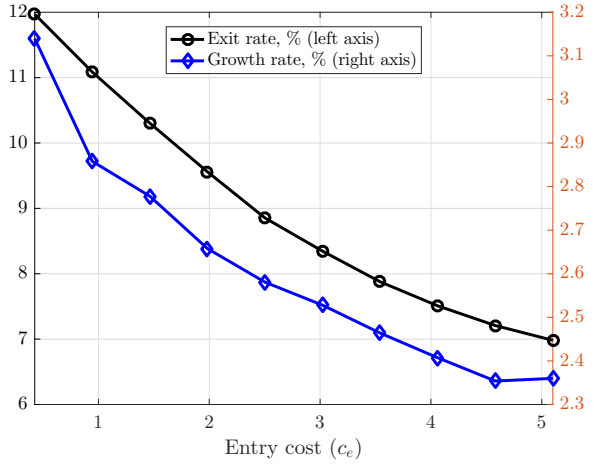
(a) Startup rate



(b) Startup size



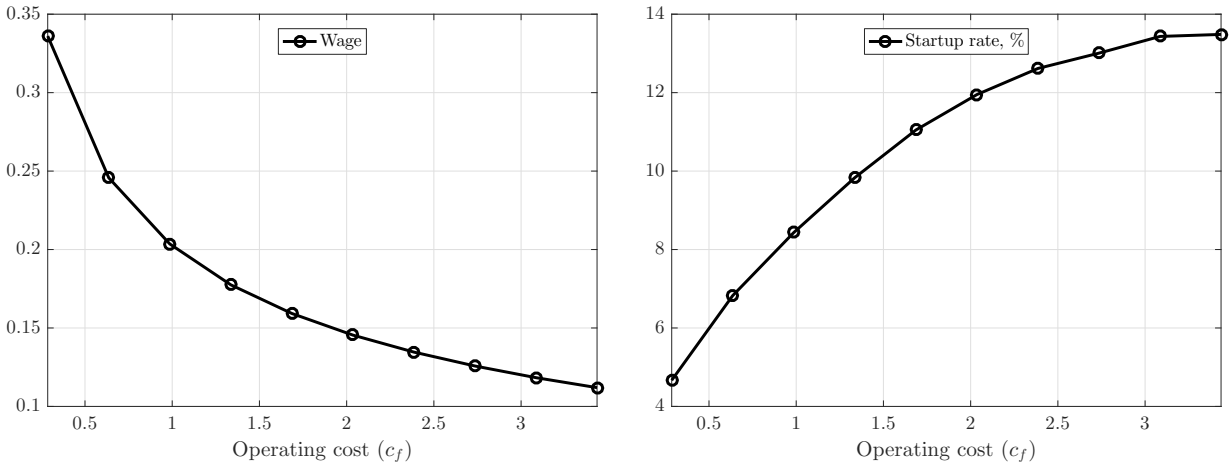
(c) Exit rate of young, small firms



(d) Conditional growth of young, small firms

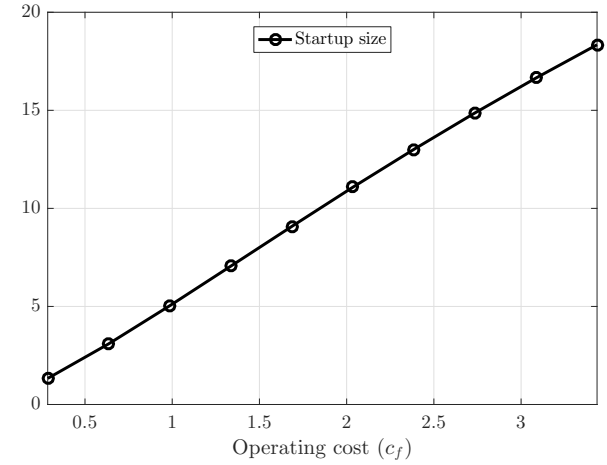
Figure 5: Entry costs and firm dynamics

Note: Four panels show how various dimensions of firm dynamics change with entry costs. Young and small firms shown on panel (d) refer to age 2 firms that have less than 50 employees.

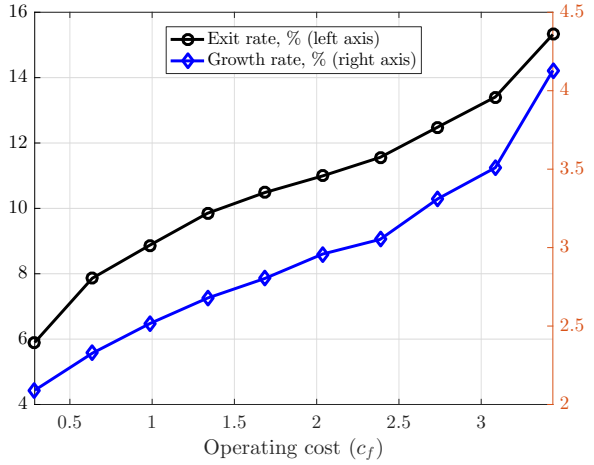


(a) Wage

(b) Startup rate



(c) Startup size



(d) Incumbent firm dynamics

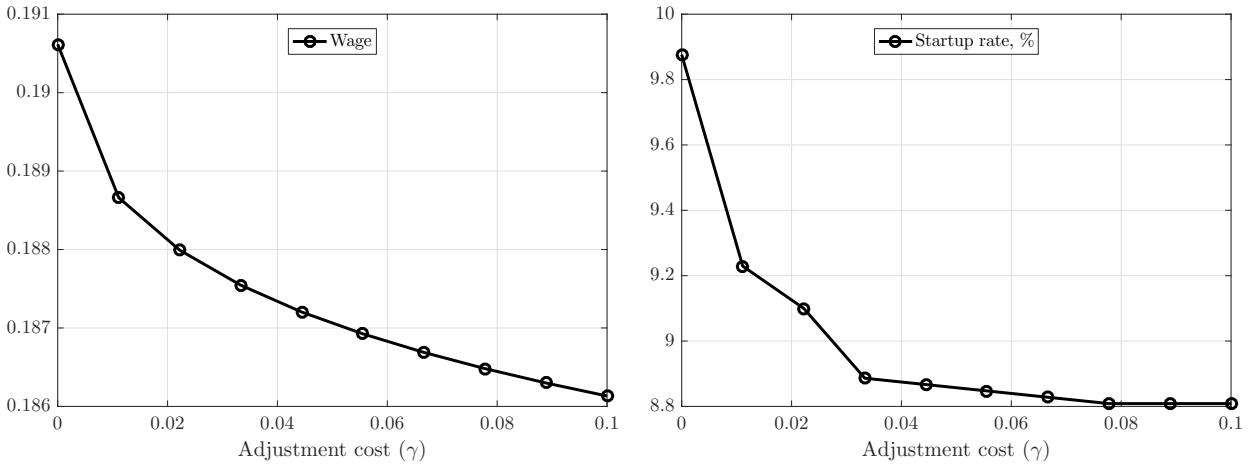
Figure 6: Operating costs and firm dynamics

Note: Four panels show how various dimensions of firm dynamics change with entry costs. Young and small firms shown on panel (d) refer to age 2 firms that have less than 50 employees.

Changes in adjustment costs Decker, Haltiwanger, Jarmin, and Miranda (2018) argue that firms are less responsive to productivity shocks, which they interpret as evidence for increasing frictions in changing firm size. In our framework, this friction corresponds to higher adjustment costs. In Figure 7, we study the effect of adjustment costs. Similar to operating costs, a higher adjustment cost reduces the value of a firm, lowering the equilibrium wage. There are two opposing effects on the size of a startup. While a lower wage pushes down the size, a higher adjustment cost implies that it is more costly for a new firm to hire more people. Quantitatively, the latter dominates, resulting in lower startup sizes when adjustment costs are higher. The wage effect dominates along the exit and growth margins as well: The direct effect of higher adjustment costs pushes up the exit rate and down the growth rate of surviving firms, whereas lower wages push in the opposite direction. We find that the indirect effect is stronger quantitatively, so that firm exit declines. The latter results in a lower startup rate.

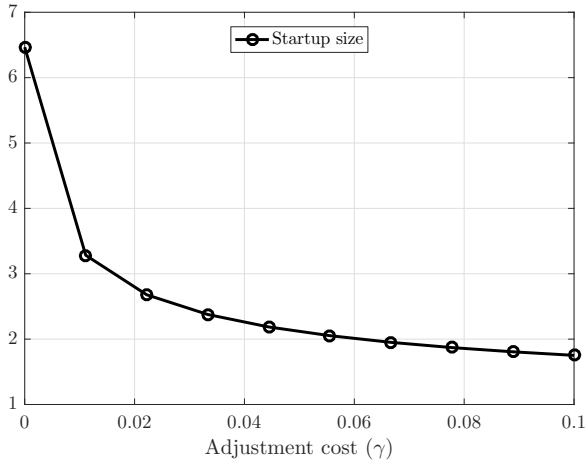
Changes in demographics Finally, we consider the role of changing labor force growth. As we have shown, over the past three decades, the growth rate of labor supply declined substantially. We highlighted via a simple example in section 3.1 that declining labor force growth can result in a lower entry rate. We now use our model to study the implications of changing demographics on firm dynamics in a more general setting. Figure 8 shows that changes in labor force growth are born entirely by the entry margin.

A change in population growth has no direct effect on the value of a firm, and thus leaves the equilibrium wage unchanged. Firm employment and exit decisions are unchanged. The extra labor supply as a result of high population growth is accommodated by an influx of more new firms, as opposed to an expansion of existing firms. In this regard, the decline in labor force growth distinguishes itself from other explanations based on changes in costs. While they all affect the behavior of incumbent firms, demographic changes leave those margins intact, and provide a coherent explanation for broader measures of firm dynamics in the U.S. over the past three decades.

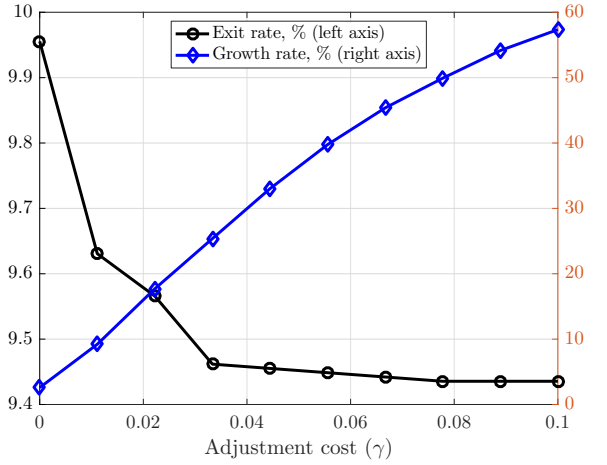


(a) Wage

(b) Startup rate



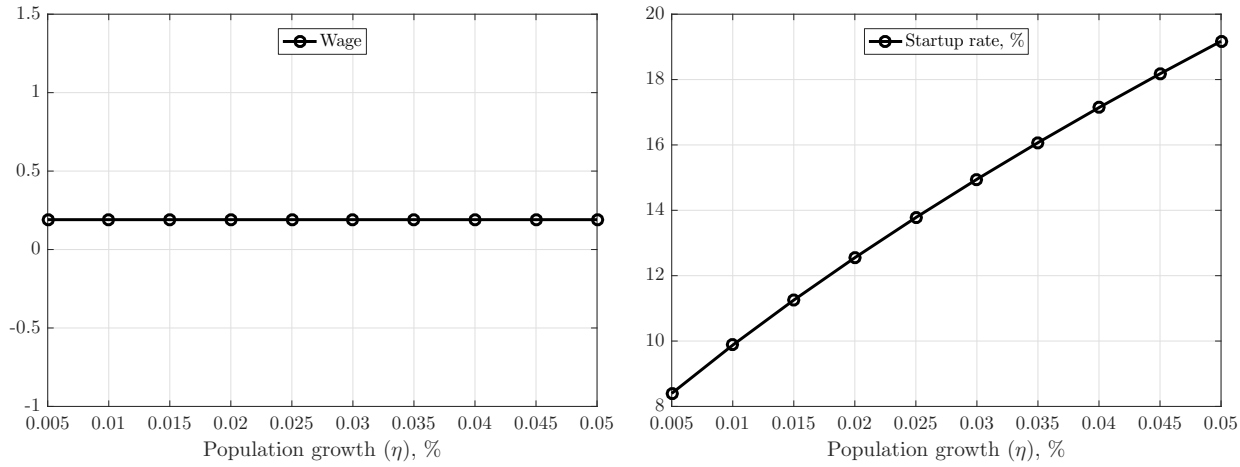
(c) Startup size



(d) Incumbent firm dynamics

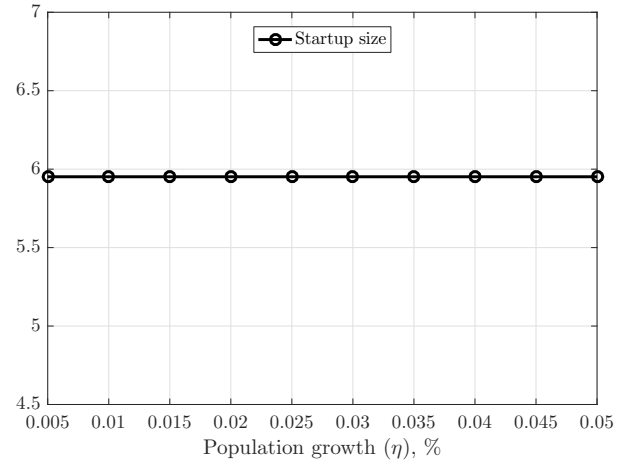
Figure 7: Adjustment costs and firm dynamics

Note: Four panels show how various dimensions of firm dynamics change with adjustment costs. Young and small firms shown on panels (c) and (d) refer to age 2 firms that have less than 50 employees.

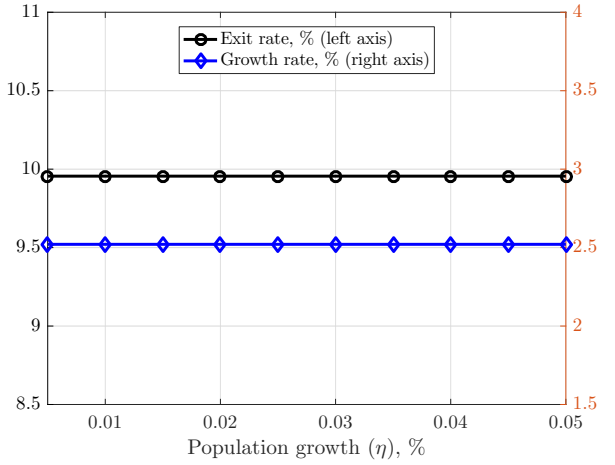


(a) Wage

(b) Startup rate



(c) Startup size



(d) Incumbent firm dynamics

Figure 8: Population growth and firm dynamics

Note: Four panels show how various dimensions of firm dynamics change with entry costs. Young and small firms shown on panel (d) refer to age 2 firms that have less than 50 employees.

4.4 Taking stock: Quantitative implications of potential explanations

Our quantitative analysis has shown that changes in the cost structure in the economy along the entry and operating cost margins and demographics can generate sizable changes in the startup rate in the long run. However, changes in the adjustment cost alone cannot account for declining entry since even with zero adjustment costs, the startup rate remains below 13% in our quantitative experiments.

Are cost driven hypotheses quantitatively relevant? So far, we have shown that cost driven explanations of declining entry also change the behavior of incumbent firms. This is in contrast to the data, as these margins are relatively stable in the data. We now investigate if the implied changes in these margins are quantitatively negligible. More specifically, we answer the following questions: What is the change in each cost required by the model to explain all of the decline in entry over the past three years? If the costs changed that much, what are the effects of those changes on the average size of startups and the exit and growth rates of incumbent firms? How do these changes compare to the data?

In Table 4, we consider entry and operating one by one and compute the required change in each cost parameter to explain the entire decline in the startup rate, We do this by keeping all parameters fixed at their calibrated values, except for the relevant cost parameter. We choose the cost parameter to match the startup rate of 13% in the model, which is the startup rate in early 1980s in the data. For entry costs to explain the decline in entry, they should have increased by almost 60% over the past decades. If a rise in entry costs of this magnitude had been the primary driver of the declining entry, then we should have seen a 12% increase in the size of startups and declines in the exit and growth rates of young and small incumbent firms of 1.16 ppts and 0.2 ppts, respectively. For operating costs to explain lower entry, they should have *declined* by 45%. Corresponding to that decline, we would have observed a large decline in the average size of startups of 47%, a decline in the exit rates of young and small incumbent firms by 1.39 ppts and a decline in their growth rate of 0.39 ppts.

	Entry cost	Operating cost	Population growth
Required change in parameter	59.47%	-45.28%	-1.2 ppts
Change in average startup size, %	11.77	-47.16	0
Change in exit rate (ppts)	-1.16	-1.39	0
Change in conditional growth (ppts)	-0.20	-0.39	0

Table 4: Implications of cost driven and demographic factors for firm dynamics

Our analysis thus shows that cost-based explanations have counter factual predictions for various margins of firm dynamics. Demographic changes, however, keep those margins intact and therefore provide a credible alternative for explaining declining startup rate. In the last column of Table 4, we repeat the same exercise for population growth rate and compute the necessary change in labor supply growth rate in the model required to generate the decline in the firm entry rate in the data.

The implied change is a decline of 1.2 percentage points in the labor supply growth rate. As we have shown in Table 1, the civilian labor force growth rate declined by 0.81 basis points from early 1980s to 2000s. This calculation suggests that changing labor supply growth rate could explain two thirds of the decline in the entry rate.

Our analysis using our general equilibrium framework has shown that a non-trivial amount of the decline in entry can be attributable to the demographic shift that the U.S. economy has been going through. Our framework also provides strong cross-sectional and time-series implications. First, our model implies that locations with more severe demographic shifts must have experiences a more notable change in their firm formation. Second, since the U.S. working age population has been growing 1960-80 period, start-up rate must have been increasing to accommodate the new entrants into the labor force.

5 Evaluating the mechanism in the cross section

As our model clearly illustrates, declines in the growth rate of labor supply may reduce the entry rate of new firms in general equilibrium without affecting their behavior after entry in terms of exit and employment growth. The goal of this section is to assess the empirical relevance of this hypothesis.

5.1 Cross sectional identification strategy

Our empirical analysis relies on cross-state variation in the timing and magnitude of labor supply growth rates. We exploit detailed data on firm dynamics and demographics at the state-level, which we describe in section 2.

With state-level firm dynamics and demographic data in hand, we estimate how the startup rate and several incumbent margins in state s, industry j and year t depend on the growth rate of labor supply, g_{st}. We estimate the following baseline specification for various outcomes related to firm dynamics:

$$y_{st} = \alpha_s + \gamma_t + \delta g_{st} + \epsilon_{st}. \tag{12}$$

The terms α_s and γ_t capture state s and year t fixed effects, respectively, and ϵ_{st} captures other sources of variation in y_{st}.

The key identification problem is that there are possibly unobservable factors that determine firm dynamics and also affect labor supply. Workers might relocate to states that are more profitable for firms and consequently attract more firm entry, causing a spurious correlation in the two measures. A successful empirical strategy of measuring the causal effects of labor supply growth should deal with this endogeneity issue and identify exogenous shifts in labor supply growth. To generate variation in labor force growth that is exogenous to demand-related factors affecting firm dynamics at the state level, we use two identification strategies.

Fertility instrument Our first strategy exploits variation in labor supply growth rates determined by fertility rates in the past. Specifically, similar to Shimer (2001) and Karahan and Rhee (2014), we obtain exogenous variation in population growth rates across-states by instrumenting with birthrates lagged by 20 years. These are measured in births per thousand residents and are available in the various Statistical Abstracts of the United States.[16]

This strategy relies on two identifying assumptions. The first and most important is the exclusion restriction that conditional on state and time effects, lagged economic conditions that cause variation in fertility rates in the past are not affecting or correlated with current business conditions, except indirectly through their effect on labor supply growth. This condition requires fertility decisions not to be driven by long-term expectations. The assumption would be violated if people in a given state had a higher fertility rate 20 years ago in anticipation of a persistently stronger labor market conditions relative to other states. The restriction also requires that higher fertility or the factors thereof have no other long lasting direct effects that would affect business dynamics 20 years later, except through their effects via labor supply. Second, and less importantly, to justify using states as the unit analysis, our empirical strategy requires the mobility costs to be large enough to prevent geographic mobility from completely equating differences across these segmented markets. To the extent this is not true, our estimates will likely understate the effect of demographics on startups and incumbent dynamics. Drawing on the literature on mobility costs, we argue that this is likely to be the case.[17]

One potential caveat of lagged birthrates is that they shift the age composition of the workforce together with its growth rate, as the primary effect of high fertility is an increase in the inflow of young workers. This shortcoming makes disentangling the size effect from the composition difficult. To examine if our results are contaminated by compositional shifts, we adopt a second strategy, which uses a Card (2001) style "migration" instrument.

Migration instrument We exploit variation in labor supply growth driven by stable migration patterns across-states. The basic premise is that changes in the population growth of a state k predict changes in the population growth of state s, if historically migrants out of state k tend to move to state s. We operationalize this idea as follows

$$\hat{m}_{st} = \sum_{k \notin C(s)} \omega^{k}_{st^*} g_{kt},$$

Here, $\omega^{k}_{st^*}$ is the share of residents of state s that were born in state k and g_{kt} is the growth rate of the working age population in k. In computing $\hat{m}_{st}$, we exclude states in the same Census division $C(s)$. To isolate the historical component of migration patterns, we use the birthplace shares from

[16] We are grateful to Rob Shimer for providing us with his data constructed from the Statistical Abstracts for the period 1940–91. Data are unavailable for Hawaii and Alaska prior to 1960. We drop these states entirely from the analysis.

[17] In an influential paper, Kennan and Walker (2011) estimate an average moving cost of \$312,000 (in 2010 dollars). This cost encompasses psychic as well as monetary costs and suggests that labor market differences across-states are unlikely to be offset by geographical mobility.

2 Censuses ago.

First-stage regressions To better understand the cross-state variation in the data, the left panel of Figure 9 plots the variation in the working age population growth rate against the fertility instrument, where both variables have been purged of state and time fixed effects. The x-axis plots the state's lagged fertility relative to its state average and the cross-state average that year. Similarly the y-axis plots the state's working age population growth rate relative to both its state average and the cross-state average for that year. A positive correlation between these two measures confirms the validity of the instrument. As Table A.1 in the Appendix shows, a 10 percentage point increase in lagged birthrate is associated with a 1.5 percentage point increase in the growth rate of working age population.

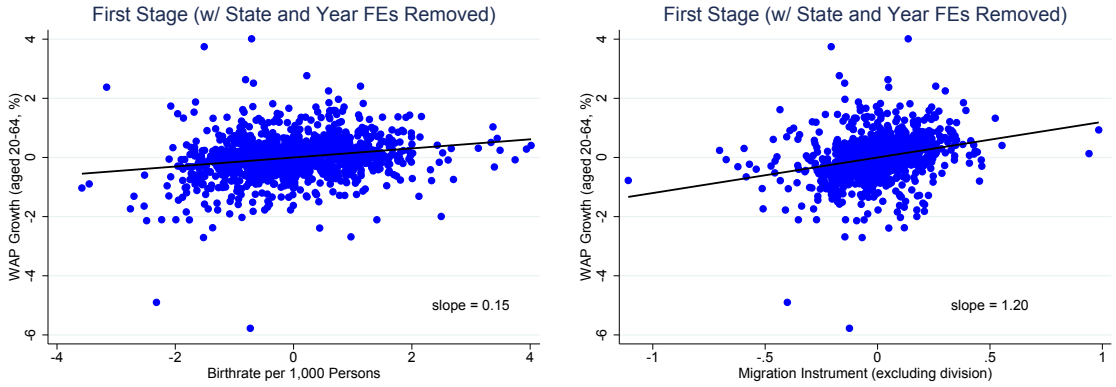


Figure 9: First-stage regressions using working-age population growth rate and civilian labor force growth rate.

Turning to the second instrument, we plot in the right panel of Figure 9 the variation in the working age population growth rate against the migration instrument, residualized in a similar fashion by regressing on state and year dummies. This figure points to a positive relationship between our instrument and population growth. Table 1 shows that this positive relationship is statistically significant. Moreover, the migration instrument contains independent information for the growth rate of working age population. The third row of Table A.1 shows that both of the instruments have power predicting WAP growth.

5.2 Labor supply growth and start-ups

We now turn to our main empirical analysis and evaluate the role of labor supply shifts in declining startup rates. To this end, we estimate the specification in (12) on pooled state-level data for the period 1980–2007. Our benchmark estimates use the growth rate of the working age population as a proxy for labor supply growth. We later investigate the robustness of our results to this choice. Unless otherwise stated, our specifications include state and year fixed effects. Later, we turn to the micro data from the LBD and include industry controls. Our standard errors are clustered

Table 5: Benchmark result: Start-up rate

	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.589***	0.776***	1.164***	1.030***
	(0.055)	(0.247)	(0.178)	(0.177)
Observations	1316	1316	1316	1316
R^2	0.90	0.89	0.86	0.87
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	24.86	14.98	13.38
J-stat	-	-	-	3.53

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state and year fixed effects, and cover the time period 1980-2007.

around state.

We start by presenting the OLS estimates. Column (1) of Table 5 shows that a 1 percentage point increase in labor supply growth is associated with roughly a 0.6 percentage point decline in the startup rate. This is a nonnegligible effect: Over our sample period, the growth rate of the working age population declined by a full percentage point and the startup rate declined by three percentage points. The OLS estimate implies that this decline explains about 20 percent of the decline.

As we discussed above, a concern with the OLS estimates is that the realized labor supply growth rates may be correlated with state-level demand shocks through worker flows, in which case the OLS estimate reflects the effects of demand and supply shifts. To identify the causal effect of declining labor supply shifts, we turn to our instrumental variables estimates.

Column (2) of Table 5 presents the results for our lagged birthrate instrument. According to this estimate, a one percentage point slow down in the working age population growth rate leads to around a 0.8 percentage point decline in the entry rate of firms. Since the working age population growth rate slowed down roughly by a full percentage point in our sample period, the results suggest that this labor supply shift can explain about a third of declining startup rates in this period.[18]

Clearly, changes in fertility rates shift labor supply growth and the age composition simultaneously. If this is the case, then our estimates could reflect the effect of the aging population on business formation and may either under- or overstate the effects of labor supply shifts. To avoid contaminating labor supply effects with that of population aging, we utilize the migration instrument discussed in Section 5.1. We estimate equation (12) by instrumenting labor supply growth rates with only the migration instrument, reported in column (3) of Table 5, and with both the mi-

[18]While it is useful to use the coefficients in Table 5 in this way and provide estimates of how much of the decline demographic shifts can account for, there is an important caveat to this approach. Most importantly, our identification does not necessarily identify the long run effects of labor supply shocks as it uses year-to-year variation. We come back to this issue later.

gration and lagged birthrate instrument, reported in column (4). The results for both specifications point to somewhat larger effects: We estimate a semi-elasticity above 1.

We explore the robustness of this result to several choices. In particular, we report three main robustness checks: allowing for detailed industry controls, using the civilian labor force growth rate instead of working age population growth rate as labor supply measure, and allowing for state-specific trends.

Industry Controls Our benchmark results are based on variation across-states and over time to measure the responsiveness of the start-up rate to demographic shocks. Demographic shocks, in theory, could affect industry composition of economic activity in a state and change its startup rate through a compositional shift. This channel is consistent with our empirical findings if positive demographic shocks shift economic activity towards high-entry industries. This alternative interpretation of our findings is different from the mechanism outlined in our theoretical framework, which predicts an across the board decline in entry in response to a demographic shock. To address this concern, we use data from the LBD, which provides detailed industry information. In Table 6, we estimate the response of the start-up rate to demographic shocks by including 4-digit industry dummies to the already included state and year fixed effects. The estimates are very similar to the ones without detailed industry controls in line with the implications of our theoretical framework.

Different Measure of Labor Supply Our proxy for labor supply is the working-age population. While demographics is an important factor in driving long-run changes in labor supply, another important margin is the labor force participation. More importantly, over the period 1980–2007, there has been a significant rise in the share of females that participate in the labor force. Our state-level analysis might yield biased results if female labor force participation correlates with the change in working age population, and if the former has a different effect on business formation. To evaluate this possibility, Table 6 reproduces the same analysis using the growth rate of the civilian labor force as the main dependent variable. The estimates using the IV strategy are remarkable similar.

State-specific trends Another concern is regarding the specification of trends. Our baseline results assume that states face similar trends in business dynamics, other than those driven by demographic shifts. We report estimates of equation (12) with state-specific linear trends in Table 6. Estimated effects are still significant and positive and somewhat larger.

These robustness analysis results in estimates for the effect of labor supply growth on startup rates that range from 0.77 to 1.40. The upper bound of our estimates attributes about 1.4 percentage points of the 3 percentage point decline in startup rates to demographics.[19]

[19]In the Appendix, we repeat our robustness checks using civilian labor force: Table A.2 allows for state-specific linear trends and Table A.3 allows for industry controls. Results are remarkably similar.

Table 6: Robustness exercises: effect of labor supply shocks on the startup rate.

A. Detailed industry controls				
	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.622***	0.893***	1.142***	1.055***
	(0.06)	(0.25)	(0.20)	(0.20)
Observations	300,000	300,000	300,000	300,000
R^2	0.51	0.51	0.50	0.51
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	25.48	15.26	13.63
J-stat	-	-	-	1.34
B. Alternative labor supply measure				
	(1)	(2)	(3)	(4)
CLF Growth (%)	0.236***	0.766***	1.111***	0.999***
	(0.037)	(0.263)	(0.250)	(0.214)
Observations	1316	1316	1316	1316
R^2	0.87	0.78	0.62	0.68
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	20.79	10.64	10.97
J-stat	-	-	-	1.64
C. State-specific trends				
	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.563***	1.176***	1.400***	1.317***
	(0.052)	(0.250)	(0.219)	(0.187)
Observations	1316	1316	1316	1316
R^2	1.00	1.00	0.99	1.00
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	31.71	11.84	16.86
J-stat	-	-	-	0.69

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state and year fixed effects as well as state time trends, and cover the period 1980-2007.

5.3 Labor supply growth, startup size, and incumbent dynamics

Recall that a unique implication of our hypothesis is that labor supply shifts do not affect the size of new firms or the growth and survival rates of incumbent firms. We now test these implications on cross-state data. To this end, we test if these margins respond to labor supply shifts by estimating (12).

Besides testing our hypothesis, measuring how other margins respond is crucial for getting a complete picture of how demographic shifts affect firm dynamics. The responses of these margins provide important restrictions for various models of firm dynamics and can be useful in estimating such models. For example, demographic shifts may affect the selection margin of entrants and the survival and growth rates of incumbents through their effects on the relative price of labor. Instead, if the only long run impact is on the entry margin, as predicted by workhorse models of firm dynamics with free entry of firms, as in Hopenhayn (1992), one would expect all other margins to be unaffected in the long run.

Motivated by these considerations, we turn to the response of other entrant and incumbent margins. Table 7 reports the effect of the growth rate of working age population on the size of startups. The estimated coefficients are negative but economically and statistically insignificant. For example, with a one percentage point decline in the growth rate of working age population over our sample period, these estimates imply a rise in startup size of about 0.3 workers. Given an average startup size around 6, this effect corresponds to an elasticity of -0.08.

Turning to the effects of labor supply shifts on incumbents, table 7 shows the response of young incumbents' survival and conditional employment growth rates. Interestingly, the estimated effects are positive and significant for the OLS. However, between 1980 and 2007, the estimated coefficients imply an increase in the survival rates of young incumbents by less than 0.1 percentage point. With a baseline survival rate of young incumbents about 89 percent, this effect implies a negligible elasticity (-0.001). Similarly, the effects on young incumbent conditional growth rates are economically small, with an elasticity of roughly 0.2.

Table 7: Start-up size, young firms' survival and growth rates in response to demographic shocks—BDS

A. Start-up size				
	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.0333	-0.358	-0.0818	-0.178
	(0.036)	(0.216)	(0.137)	(0.145)
Observations	1316	1316	1316	1316
R^2	0.54	0.44	0.53	0.51
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	24.86	14.98	13.38
J-stat	-	-	-	1.88
B. Young firm survival				
	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.374***	-0.0930	0.252	0.104
	(0.047)	(0.242)	(0.169)	(0.136)
Observations	987	987	987	987
R^2	0.73	0.67	0.72	0.71
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	36.27	9.06	17.09
J-stat	-	-	-	1.06
C. Young firm growth				
	(1)	(2)	(3)	(4)
WAP Growth (20-64, %)	0.614***	-1.092	0.0417	-0.445
	(0.168)	(0.716)	(0.894)	(0.748)
Observations	987	987	987	987
R^2	0.43	0.33	0.42	0.39
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	36.27	9.06	17.09
J-stat	-	-	-	1.96

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state and year fixed effects, and cover the time period 1980-2007.

6 Evaluating the mechanism in the time-series

The U.S. working age population has been growing 1960-80 period as we have seen in Figure 1. An immediate implication for our analysis is that the start-up rate must have been increasing to accommodate the new entrants into the labor force. The main difficulty of testing this implication is the lack of longitudinal firm-level data for that time period since the Census Business Register (from which the LBD is constructed) is not reliable prior to 1976. However, by imposing some structure on the data, we can estimate the behavior of entrants using the cross sectional establishment data from the Census Bureau's much older County Business Patterns (CBP) program.

To test the important time-series implication of our analysis we use data from the CBP and impute the start-up rate for the earlier time period. CBP has been published annually since 1964, which allows us to impute the establishment entry rate from cross-sectional observations on establishments in the U.S. CBP provide annual series on the number of establishments and employment during the week of March 12. Data for establishments are presented by geographic area and employment size class.

Our imputation methodology relies on a simple law of motion. Let e_t be the number of establishments in year t. The law of motion for the evolution of e_t can be expressed as

$$e_t = e_{t-1} - \sum_{k,j} \delta_t^{kj} e_{t-1}^{kj} + s_t, \tag{13}$$

where δ_t^{kj} is the exit rate for establishments of size k in state j and e_t^{kj} is the number of establishments of size k in state j in year t. This equation allows us to impute the number of startups, s_t, given the number of establishments and exit rates over time.

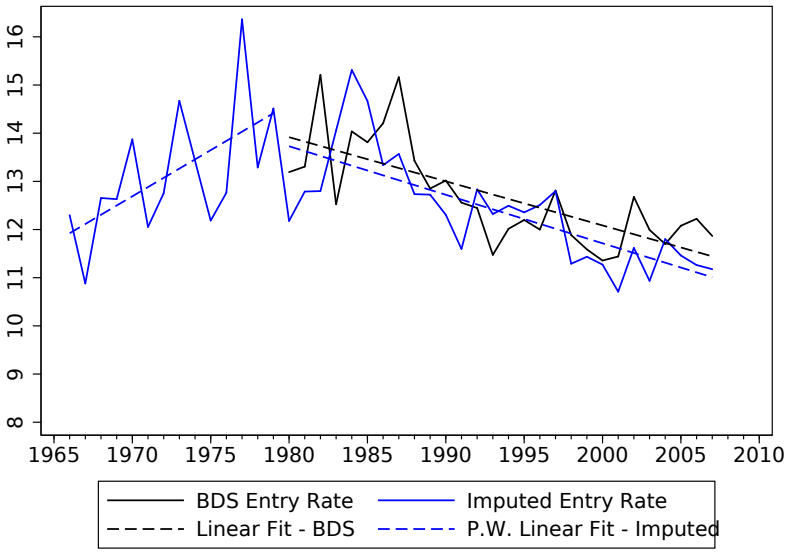


Figure 10: Imputed historical establishment entry rates and the BDS entry rate.

The CBP data provides us with measures of e^t_{kj} starting in 1964.[20] We use six size categories: $\{1-19, 20-49, 50-99, 100-249, 250-499, 500+\}$ and all 50 U.S. states and the District of Columbia.

Since the CBP is not a longitudinal data source, we can not use it compute the establishment exit rates, δ^{kj}_t. To the best of our knowledge, there is no other data source going back in time that allows for calculation of exit rates. What we do instead is to use the BDS starting in 1980 and compute the δ^{kj}_t starting in 1980. We find that these exit rates slightly trended downward. We estimate this downward trend and use it in our post-1980 imputation. For the pre-1980 imputation, we impose the restriction that exit rates equaled to their sample average in the earlier period, $\delta^{kj}_t = \delta^{kj}_{1980-2007}$.[21] With number of establishments and exit rates in hand, we use equation 13impute the number of entrants. Finally, we calculate the establishment entry rate by normalizing the number of entrants by the number of establishments in the previous period.

Figure 10 shows the imputed entry rate for 1965-2007 as well as the establishment entry rate from the BDS. First of all, the imputed entry rate using captures the evolution of the entry rate well in the time period that BDS and CBP overlap. Second, the imputed establishment entry rate exhibits an upward trend in the 1960-1980 period consistent with the time-series implication of our analysis.

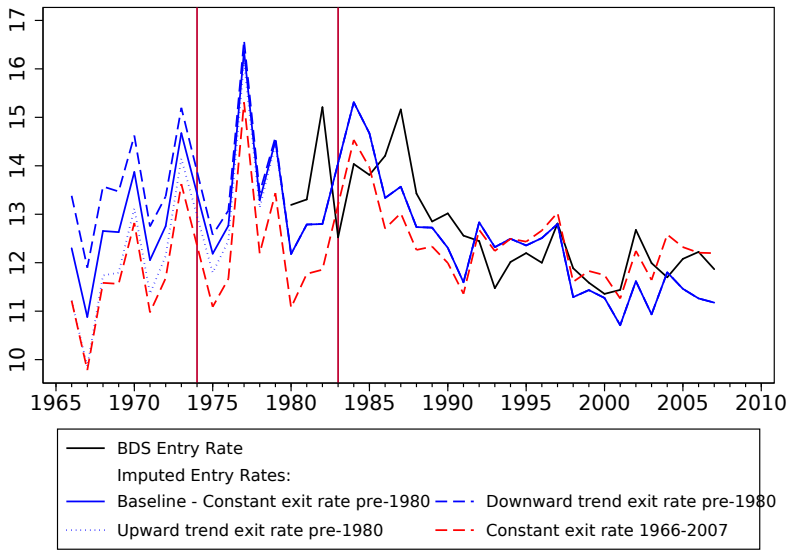


Figure 11: Imputed historical establishment entry rates under different assumptions for the behavior of exit rates before 1980.

We also relax our assumptions on the evolution of exit rates and repeat our imputation procedure under different specifications: (1) assume that the exit rates were constant during the entire 1965-

[20] Years 1964, 1974, and 1983 are dropped due to significant modifications in the records of establishments in the CBP.

[21] We later relax this assumption and repeat our imputation procedure under different assumptions for the evolution of exit rates.

2007 period at their average levels; (2) assume the downward time trend persisted from prior to 1980, (3)assume that the time trend prior to 1980 was opposite in direction but equal in magnitude. The resulting four imputed entry rate time series are plotted against the BDS entry rate in Figure 11. All imputed time series exhibit an upward trend in the pre-1980 period reinforcing the pattern suggested in the baseline specification.

Finally, one concern is that data restrictions allow us to impute the establishment entry rate while rest of our analysis focused on firm entry. However, as we show in the Appendix, establishment and firm entry followed a very similar trend in the 1980-2007 period.

7 Conclusion

Ongoing decline is business formation in the U.S. economy is an important development which has started to affect the macro economy. In this paper we proposed and evaluated a simple hypothesis to explain the decline in the entry rate: shifts in the growth rate of labor supply have a direct effect on firm entry. Put differently, when an economy goes through a demographic change, its startup rate adjusts to accommodate the change in its labor supply. We show that this insight—which is a basic implication of Hopenhayn (1992) style firm dynamics model—is the key to establish the link between worker and firm demographics.

We examine the quantitative relevance of this link using three different but complementary approaches. We first set up a firm dynamics model with population growth and evaluate the quantitative role of the decline in the growth rate of labor supply from around two percent in the early 80s to one percent by 2000. Our calibrated model implies that around two thirds of the decline in entry can be explained solely by this demographic shift. We then test this hypothesis using cross-sectional variation in the demographic component of labor supply growth and find that the data supports a big causal effect via the mechanism of the model. Lastly, we test the time-series implications of our hypothesis and examine the link between demographics and firm entry in the pre-1980 period by imputing data using an alternative data set. The earlier data support the relationship we have established in the post-1980 data: The increase in the growth rate of working-age population in the 1965-1980 period coincides with increasing firm entry.

A different strand of the literature suggests barriers to entry and increased regulations for running a firm as drivers of declining entry. We use our calibrated firm dynamics model to examine these explanations. We show that, while changes in the cost structure of the economy such as entry and operating costs could have potentially caused the decline in business formation, they are not likely to be the main drivers due to two reasons that we have shown: 1. These changes would have affected the incumbent firm dynamics profoundly, while in the data firm life cycle dynamics have not changed. 2. For changes in the cost structure of the economy to account for the decline in firm entry, these costs should have changed considerably relative to 1980s—a change for which there is little evidence in the data.

An open question remains whether this insight carries to cross-country data. Demographic shift

towards a lower population growth rate is a common trend in many developed countries. Our paper offers a common explanation for the decline in business formation in many developed countries. Exploring this link in cross-country data is a promising research question that we delegate to future research.

References

Bartelsman, E. J., and M. Doms (2000): "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, (3), 569–594.

Davis, S. J., and J. Haltiwanger (2014): "Labor Market Fluidity and Economic Performance," Discussion paper, National Bureau of Economic Research.

Decker, R., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2014): "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *The Journal of Economic Perspectives*, 28(3), 3–24.

——— (2018): "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness," Working Paper No. 24236, National Bureau of Economic Research.

Elsby, M. W., B. Hobijn, and A. Şahin (2015): "On the importance of the participation margin for labor market fluctuations," *Journal of Monetary Economics*, 72, 64–82.

Engbom, N. (2017): "Firm and Worker Dynamics in an Aging Labor Market," Working paper.

Gourio, F., T. Messer, and M. Siemer (2014): "A Missing Generation of Firms? Aggregate Effects of the Decline in New Business Formation," mimeo.

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

Hathaway, I., and R. E. Litan (2014a): "Declining Business Dynamism in the United States: A Look at States and Metros," mimeo, Brookings Institution.

Hathaway, I., and R. E. Litan (2014b): "What's Driving the Decline in the Firm Formation Rate? A Partial Explanation," Discussion paper, Brookings Institution.

Hopenhayn, H., and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of Political Economy*, pp. 915–938.

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

Jarmin, R. S., and J. Miranda (2002): "The Longitudinal Business Database," Working Paper CES-02-17, US Census Bureau Center for Economic Studies.

Karahan, F., and S. Rhee (2014): "Population Aging, Migration Spillovers and the Decline in Interstate Migration," *FRB of New York Staff Report*.

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251.

Lentz, R., and D. T. Mortensen (2008): "An Empirical Model of Growth Through Product Innovation," *Econometrica*, 76(6), 1317–1373.

Liang, J., H. Wang, and E. P. Lazear (2014): "Demographics and Entrepreneurship," mimeo, Stanford University.

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

Pugsley, B. W., and A. Şahin (2014): "Grown-up business cycles," *FRB of New York Staff Report*, (707).

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

Reedy, E., and R. J. Strom (2012): "Starting Smaller; Staying Smaller: America's Slow Leak in Job Creation," in *Small Businesses in the Aftermath of the Crisis*, pp. 71–85. Springer.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobeck (2017): "Integrated Public Use Microdata Series: Version 7.0 [dataset]," Minneapolis: University of Minnesota.

Shimer, R. (2001): "The Impact of Young Workers on the Aggregate Labor Market," *Quarterly Journal of Economics*, Vol. 116, No: 3 (Aug., 2001), 969–1007.

A Additional Results

Table A.1: First-stage regressions

WAP Growth (20-64, %)	IV_1	IV_2
IV_1: Lagged Birthrate	0.153***	
	(0.03)	
IV_2: Migration		1.198***
		(0.31)
IV_1 & IV_2	0.106***	1.004***
	(0.03)	(0.29)

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All specifications are firm-weighted and all errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state, year, and NAICS 4 fixed effects, and cover the time period 1980-2007.

Table A.2: Start-up rate with clf and state-specific trends

	(1)	(2)	(3)	(4)
CLF Growth (%)	0.205***	1.089***	1.332***	1.238***
	(0.036)	(0.281)	(0.350)	(0.266)
Observations	1316	1316	1316	1316
R^2	1.00	0.99	0.98	0.99
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	16.54	7.37	8.40
J-stat	-	-	-	0.46

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state and year fixed effects as well as state time trends, and cover the period 1980-2007.

Table A.3: Start-up rate with clf and industry controls

	(1)	(2)	(3)	(4)
CLF Growth (%)	0.257***	0.882***	1.089***	1.022***
	(0.04)	(0.25)	(0.26)	(0.23)
Observations	300,000	300,000	300,000	300,000
R^2	0.51	0.49	0.48	0.49
Specification	OLS	IV_1	IV_2	IV_1 & IV_2
F-stat	-	21.20	10.81	11.18
J-stat	-	-	-	0.62

Notes: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All specifications are firm-weighted and all errors are clustered on state. Regressions exclude AK, HI, UT, and DC, contain state, year, and NAICS 4 fixed effects, and cover the time period 1980-2007.