STAFF REPORTS

NO. 888
MAY 2019

REVISED
MARCH 2021

Demographic Origins of the Startup Deficit

Fatih Karahan | Benjamin Pugsley | Ayşegül Şahin

FEDERAL RESERVE BANK *of* NEW YORK

Demographic Origins of the Startup Deficit
Fatih Karahan, Benjamin Pugsley, and Ayşegül Şahin
Federal Reserve Bank of New York Staff Reports, no. 888
May 2019; revised March 2021
JEL classification: J11, E24, D22

Abstract

We propose a simple explanation for the long-run decline in the U.S. startup rate. It originates from a slowdown in labor supply growth since the late 1970s, largely pre-determined by demographics. This channel can explain roughly 60 percent of the decline and why incumbent firm survival and average growth over the lifecycle have changed little. We show these results in a standard model of firm dynamics and test the mechanism using cross-state variation in labor supply growth. Finally, we show that a longer entry rate series imputed using historical establishment tabulations rises over the 1960-70s period of accelerating labor force growth.

Key words: firm dynamics, demographics, business dynamism, macroeconomics

---

Karahan: Federal Reserve Bank of New York (email: fatih.karahan@ny.frb.org) Pugsley: University of Notre Dame (benjamin.pugsley@nd.edu), Şahin: University of Texas Austin, NBER (aysegul.sahin@austin.utexas.edu). First version: July 2015. This paper was previously circulated under the title "Understanding the 30-Year Decline in Business Dynamism: A General Equilibrium Approach." The authors thank Rob Dent, Sam Kapon, Harry Wheeler, René Chalom, and Isabel Hansich for their superb research assistance. They are grateful for very helpful comments from discussants Petr Sedlacek (HEC 2015), Steve Davis (NBER Macro & Productivity 2015), Richard Rogerson (Atlanta Fed Employment 2015), John Haltiwanger (NBER EF&G 2018), and Chris Huckfeldt (Fordham Macro IF 2019), as well as from the editor and three anonymous referees, and for very helpful conversations with Mark Bils, Ryan Decker, Henry Hyatt, Bill Kerr, Pete Klenow, Joe Kaboski, Huiyu Li, Giuseppe Moscarini, Emi Nakamura, Pietro Peretto, Juan-Carlos Suárez Serrato, and numerous seminar and conference participants. The authors also thank Aaron Hedlund and two external reviewers for the Center for Growth and Opportunity for very helpful comments on an earlier version of the paper. They are also indebted to Felipe Schwartzman for his insightful comments on a previous paper and his suggestion to consider the effect of labor market shocks on firm entry. Pugsley and Şahin gratefully acknowledge financial support from the Center for Growth and Opportunity. Pugsley also gratefully acknowledges financial support from the Ewing Marion Kauffman Foundation.

To view the authors' disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr888.html.

1 Introduction

For over three decades, the pace of U.S. business creation has been slowing. In 1979, the startup rate—measured as the share of new employers as a fraction of all employers—was 13 percent. By 2007, *before* the onset of the Great Recession, it had fallen to roughly 10 percent, an almost 25 percent decline.[1] This "startup deficit" is also wide-ranging; it is found nearly universally within geographic areas and industries. Remarkably, the declining startup rate and its origins left little apparent effect on the survival or average growth rates of incumbent firms, which have remained steady conditional on firm age. Instead, many aggregate effects, such as the decline in the economy's exit rate and increase in average firm size, have followed from the implied changes in firm age composition.

We propose and evaluate a simple and novel explanation for these patterns. The startup rate is linked in equilibrium to the pace of labor supply growth, which slowed dramatically in the late 1970s. This explanation fits the data very well. The slowdown in U.S. labor supply growth can explain roughly 60 percent of the declining startup rate; it explains the widespread declines across markets since the labor force growth slowdown affected them nearly universally; *and* it explains the stability of average incumbent dynamics conditional on their age, with aggregate effects following from the resulting shift towards older firms. We show these results quantitatively in an equilibrium model of firm dynamics. Moreover, we validate the model's implications using variation across states and in the aggregate time series.

Why should labor supply growth affect the startup rate at all? Ultimately, growing labor supply requires growing labor demand, which can only arise from a decline in the real wage that allows incumbent labor demand to expand or through entry of new firms. A lesson from models of firm dynamics is that in the long run, free entry of new firms relieves any pressure on the real wage from shifts in labor supply. Without any real wage adjustment to move along the incumbent labor demand curve, shifts in labor supply are instead absorbed at the *entry* margin. This feature makes the startup rate highly responsive to changes in labor supply growth.

[1] See Davis, Haltiwanger, Jarmin and Miranda (2007) who first documented declines in measures of business dynamism, and Haltiwanger, Jarmin and Miranda (2011, 2012) who first noted the decline in the startup rate, in particular. More recently, work by Decker, Haltiwanger, Jarmin and Miranda (2014), Hathaway and Litan (2014a), and Pugsley and Şahin (2019) have shown the declines are both persistent and pervasive across markets.

To show this, we extend the general equilibrium model of Hopenhayn and Rogerson (1993) to include labor supply growth, η. The benefit of this well-known setting is its simplicity, but the mechanism is present in richer environments that share its key ingredients of firm-level diminishing returns and a linear free entry condition. Along a balanced growth path, new firms are needed to replace exiting firms and to expand labor demand. Without labor supply growth the equilibrium startup rate equals the economy's exit rate, x. With labor supply growth, the startup rate instead equals $(\eta + x)/(1 + \eta)$, which accounts for replacement and the growth needed to keep the number of firms per worker constant on the balanced growth path.

The equilibrium link between labor supply growth and the startup rate works both through the change in labor supply growth, η, *and* through an endogenous response of the exit rate, x. The direct effect of a change in η is the change in the growth rate in the number of firms needed to maintain balanced growth. The response of the aggregate exit rate creates an indirect amplification effect. Since changes in the startup rate shift the share of new and young firms, which are much more likely to exit, they alter the aggregate exit rate through a compositional effect. These changes in aggregate exit then feed back on the startup rate by changing its replacement component.

We calibrate the model to match U.S. firm dynamics and use it to quantify the effects of labor supply growth on the startup rate. The calibrated model has a long-run elasticity of the startup rate to the labor supply growth rate of 1.3, of which 0.9 is the direct effect and 0.4 is the indirect amplification from the endogenous aggregate exit rate. This amplification is strongest under the linear free entry condition, but we also show it remains sizeable for empirically plausible deviations from free entry. Viewed through the lens of the model, decades of declining labor supply growth reduced the startup rate, which through its compositional effects on aggregate exit, depressed the startup rate even further. To quantify these effects, we take into account the transitional dynamics in the firm age distribution by examining the equilibrium response of the startup rate to the actual path of two proxies for labor supply growth. The path of working-age population growth since 1948 can explain about 1/3 of the decline in the startup rate over the 1979–2007 period. The path of civilian labor force growth, which also includes the effects of secular changes in labor force participation, captures even more, explaining roughly 60 percent of the startup rate decline.

A further appeal of the mechanism is that it agrees with other salient features of the data. It is consistent with mostly stable incumbent survival and growth by firm age observed over the same time period. In the model, these are held fixed in the long run by free entry, but even over the transition, they change relatively little. With declines in entry and little offsetting change from incumbents conditional on firm age, the startup deficit accumulates in a shift towards older firms that mirrors the same shift in the data. This change in firm age composition explains declines in aggregate exit and increases in average firm size that are close to those in the data. When we explore other mechanisms that could also explain declines in entry, they are at odds with these empirical patterns. For example, generating comparable declines in the startup rate requires large changes in costs that would imply strong effects on firms' lifecycle dynamics inconsistent with the data.

The model also provides a rich set of predictions we can evaluate using detailed cross-sectional data. The first is the long-run elasticity between the pace of labor supply growth and the startup rate, and we find a strong and positive relationship across U.S. states that is close to the model. Further, the transitional dynamics of the model allow us to test the shorter-run relationship between labor supply growth and the startup rate using annual data. In the data, we use two separate IV strategies to generate plausibly exogenous variation in labor supply growth across U.S. states and compare the changes in state-level startup rates. The first IV predicts labor supply growth using 20-year lags of each state's birthrate, and the second IV exploits persistent historical migration patterns between states. Using both IVs, we estimate a cross-sectional startup rate elasticity of labor supply growth of 1.19, and we find this elasticity is little changed in a battery of robustness exercises. When we compare the elasticity from the data with its model counterpart computed from a simulated panel, the corresponding elasticity is 1.10.

As a final test of the demographic explanation, we test the time-series implication that the 1960-70s period of rising labor force growth should coincide with an *increasing* startup rate. Although data limitations preclude examining the startup rate prior to the late 1970s, we develop a method to impute an establishment entry rate using the Census Bureau's static County Business Patterns (CBP) data that extends back to 1965. Given that the establishment entry rate is highly correlated with the firm startup rate over the entire period where we can measure both, the imputed establishment entry rate will be informative about the historical firm startup rate. To do this, we estimate a statistical model for the aggregate exit rate as a function of the distribution of establishment characteristics measurable in the detailed CBP cross sections. We then add the estimated aggregate exit rate to the growth rate in the number of establishments, which is measurable across CBP years, to impute a historical entry rate series. The imputed establishment entry rate co-moves with the labor supply growth rates over their rise and fall, consistent with the demographic hypothesis.

Related literature. Our paper contributes to an emerging literature on the causes and consequences of declining entrepreneurship, first noted by Haltiwanger, Jarmin and Miranda (2011) and Haltiwanger, Jarmin and Miranda (2012). Early descriptive work on the trend decline in the startup rate had explored myriad potential explanations and documented cross-sectional correlations of the startup rate with changes in population growth (Hathaway and Litan, 2014b; Pugsley and Şahin, 2019), business consolidation (Hathaway and Litan, 2014b), and sectoral import competition (Pugsley and Şahin, 2019), as well as ruled out other possibilities such as changes in industrial composition (Decker, Haltiwanger, Jarmin and Miranda, 2014) and substitution from nonemployers (Pugsley and Şahin, 2019). We are the first, to our knowledge, to propose demographic changes in the growth rate of the labor supply as a causal explanation for the trend decline in the startup rate as well as the first to quantitatively and empirically evaluate the equilibrium mechanism.

Several recent papers have examined other potential causes of the decline in the startup rate using quantitative models. For example, Salgado (2017), Kozeniauskas (2017), and Jiang and Sohail (2017) consider effects of skill biased technological change that raises the opportunity cost of entrepreneurship for high-skilled workers. Beyond labor supply growth, demographics can also affect the startup rate through the age composition of the population. Liang, Wang and Lazear (2018) and Engbom (2019) find that incentives to start new businesses are diminished with an older workforce. Bornstein (2018) shows that population aging depresses firm entry because older consumers are less likely to demand new varieties.

The macroeconomic consequences of the startup deficit are far reaching. The most direct effect is through its effect on the firm age distribution which shifted towards older firms as a consequence of lower firm entry. Haltiwanger, Jarmin and Miranda (2013) and Decker, Haltiwanger, Jarmin and Miranda (2014) show the importance of young firms, including startups, for aggregate job creation and Pugsley and Şahin

(2019) show firm age composition affects aggregate employment dynamics and link the emergence of jobless recoveries to declining firm entry. Alon, Berger, Dent and Pugsley (2018) show that the aging of firms has been a drag on productivity growth as older firms exhibit lower productivity growth (see also Bartelsman and Doms, 2000). Cloyne, Ferreira, Froemel and Surico (2018) show the effects of shifting firm age composition also extend to the aggregate investment dynamics and the efficacy of monetary policy. In addition, changes in entry and firm age composition can have follow-on equilibrium effects that affect firm behavior *within* firm age group. For example, declines in entry can affect market competitiveness and incentives for incumbents to innovate that by affecting the rate of innovation can have substantial long-run effects (Lentz and Mortensen, 2008; Atkeson, Burstein and Chatzikonstantinou, 2018; Peters and Walsh, 2019; Klenow and Li, 2020; Akcigit and Ates, 2021).

The paper is organized as follows. Section 2 presents the trends in demographics and startups and a simple framework to connect them. Section 3 evaluates the demographic channel in a dynamic general equilibrium model. Section 4 uses cross-state variation, and Section 5 uses an imputation method to further evaluate the demographic channel. Section 6 concludes.

2 The startup deficit and demographic change

In this section, after first describing our data, we document the decline in the U.S. startup rate that coincides with a significant demographic slowdown in the growth rate of the labor supply. We then introduce a simple framework to motivate why parallel trends in firm dynamics and labor supply are related economically.

2.1 Trends in U.S. firm dynamics and labor supply

Throughout the paper we use confidential data on employer businesses from the U.S. Census Bureau's Longitudinal Business Database (LBD). The LBD covers nearly all U.S. nonfarm private-sector establishments with employees starting in 1976.[2] We construct firm-level measures by aggregating across each firm's one or more establish-

[2]The LBD is derived from the Census Bureau's Business Register of all private-sector establishments with paid employees that is then longitudinally-linked at the establishment level. Jarmin and Miranda (2002) provide a detailed description of the linking procedure and LBD construction.

ments.[3] To facilitate replication, whenever possible, we rely on the public-use tabulations available in the U.S. Census Bureau's Business Dynamics Statistics (BDS).

We identify startups and distinguish incumbents based on firm age, which, to be consistent with prior literature, we calculate as the age of the firm's oldest establishment. To measure firm-level employment (our measure of firm size), we aggregate employment across all establishments within a firm. Firm-level employment growth, g_{it}, is measured as the employment-weighted average of establishment-level employment growth across all of firm i's year t establishments. We also consider measures defined at the age group cohort-, rather than firm-, level, where an age group cohort is the set of firms that belong to an age group in a particular year.

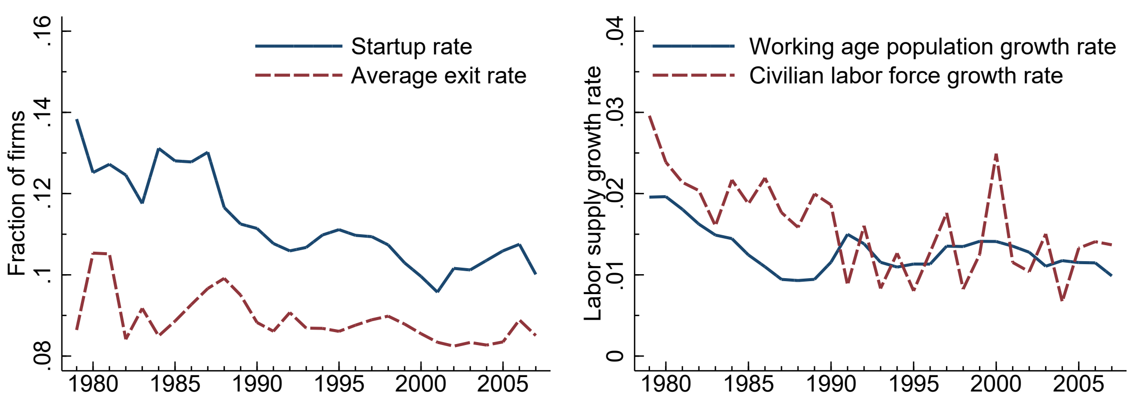
We measure firm exit in year t when all of a firm's year $t-1$ establishments have 0 employment and are reported closed in year t. Similar to an age group's employment growth, we measure an age group's exit rate, x_t^a, as the number of exits in year t for the age group a as a fraction of the year $t-1$ number of firms. While this measure is defined by age-group, we also compute the economy-wide exit rate, X_t, as the total number of exits as a fraction of the total number of (operating) firms in year $t-1$. Following Pugsley and Şahin (2019), we refer to the measure of an age group's employment growth $1+g_t^a = E_t^a/E_{t-1}^a$ as the *unconditional* growth rate, which we decompose as the product of a survival and conditional growth rate: $1+g_t^a \equiv (1-x_t^a)(1+n_t^a)$. For brevity, we provide further details in Appendix A.

Changes in aggregate entry, exit and labor market growth. We first examine the trends in the startup and exit rates over the 1979-2007 period, which pre-dates any effects from the Great Recession.[4] The startup rate has fallen almost 25 percent since

[3]Our unit of analysis throughout the paper is a firm, which is composed of one or more establishments (physical locations). The Census Bureau defines firm boundaries by the highest level of operational control, and this is well-suited to our analysis since it is the level where decisions about scale and labor demand are ultimately made. A mature incumbent may expand by opening a new establishment, but we do not want to measure this in our startup rate since it reflects an *incumbent* decision. Moreover, the aggregate establishment entry rate is also declining, but *only* from the establishments of new firms. Figure C.8 shows that among incumbents with multiple establishments, there is *no decline* in their establishment entry rate (Appendix C.1.4). In practice, since most firms are single establishments, the distinction between firm and establishments makes little difference for our results. For reference, we provide establishment counterparts of the model results in Appendix B.1 and the cross-state results in Section 4.3.

[4]While early work on the changes in the startup rate and job creation from new and young firms, such as Haltiwanger, Jarmin and Miranda (2011), Gourio, Messer and Siemer (2014), and Davis and Haltiwanger (2016), focused primarily on the Great Recession and its slow recovery, we focus on the trend and examine the 1979 to 2007 period, *before* any influence from the Great Recession.

1979, from an average of roughly 13 percent to around 10 percent in 2007 (Figure 1a). While not apparent from the aggregate data, the decline in startup activity was widespread across sectors and locations, ruling out compositional shifts as the main cause—a finding that resonates with Decker, Haltiwanger, Jarmin and Miranda (2014).[5] The declining startup rate also coincides with a decline in the economy-wide exit rate, which falls about 1 percentage point over the same period (Figure 1a). The lower exit rate is explained entirely by compositional change in *firm age*. As we show below, *within* age groups, the exit rate shows no comparable declines.



(a) Startup and average exit rates

(b) Labor supply growth rates

Figure 1: Declines of startup, exit, and labor supply growth rates from 1979 to 2007

Over the same period, there are also declines in measures of labor supply growth (Figure 1b). The growth rate of the working-age population (WAP) surges in the 1960s as the first post-WWII baby boom birth cohorts enter adulthood, and it peaks in the late 1970s.[6] It then declines from almost 2 percent in 1979 to just over 1 percent by 2007. Importantly, this trend in labor supply growth was mostly predetermined by historical changes in fertility that were plausibly unrelated to the business environment over the 1979-2007 period.[7] Similarly, the growth rate of the

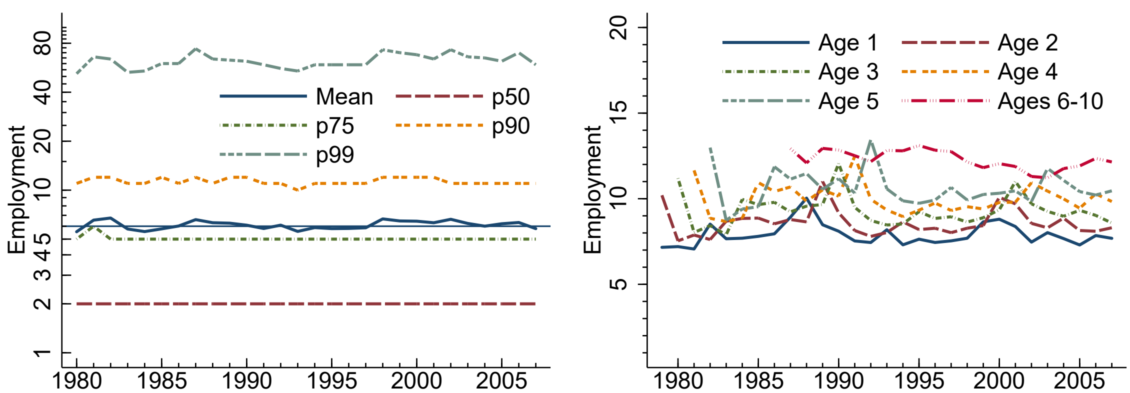
[5]We show that within-sector and within-geography changes account for more than 100 percent of the aggregate decline (Appendix C.1.2). We also show that entry declines are not specific to our measure of the startup rate (Appendix C.1.3).

[6]We define working age as 20-64, which is slightly more expansive than the 25-54 "prime-age" range. Participation among ages 20-24 and 55-64 is somewhat lower than prime-age, but it falls off steeply outside of ages 20-64.

[7]Net international migration also contributes to working-age population growth and could vary with contemporary business conditions, but over the 1979-07 period (except for a spike in the late 1990s) the net migration rate remained close to 0.3 percent.

civilian labor force (CLF), which includes all individuals age 20 or older currently employed or actively searching for a job, slows from about 3 percent in 1979 to 1 percent by 2007. The peak CLF growth rate is roughly 1 percentage point higher than the WAP growth rate since it adds the effects of rising female participation. After female participation levels off in the late 1990s, both measures grow at roughly 1 percent per year.

Stability of other margins by firm age. In contrast to the economy-wide entry and exit margins described above, other margins of firm dynamics are little changed once conditioned on firm age. We now examine this stability.



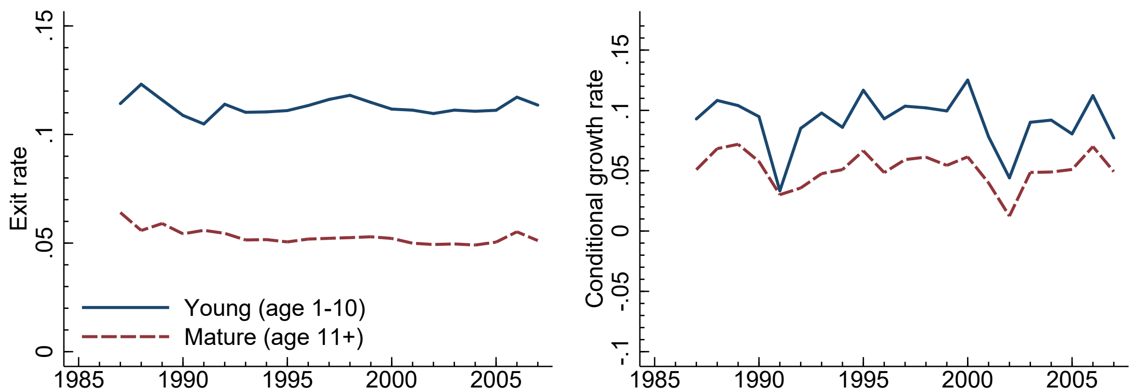
(a) Distribution of startup employment size (b) Young incumbent average employment size

Figure 2: Startup and young incumbent employment size

Note: Left panel uses log scale, and thin reference line plots 1979-2007 average of 6 employees.

First, among startups, the distribution of employment size has been stable (Figure 2a). Average startup size is procyclical but nevertheless remains centered around 6 employees. This stability appears throughout the distribution. The median new firm had 2 employees from 1979 through 2007, while the 75th percentile has remained 5 employees, except for 1981. Even the 90th and 99th percentiles have remained close to their long-run averages of 11 and roughly 60 employees, respectively. Average firm size also remained stable for young firms of ages 1, 2, 3, 4, 5, and 6-10 (Figure 2b).

Next, we examine exit and the average growth of incumbent firms by age. The exit rate for firms age 1-10 (young) and firms 11 or more years old (mature) have been stable (Figure 3a). This may seem puzzling given the decline in the aggregate exit rate, but this decline reflects changes in composition; because mature firms have significantly lower exit hazards, the increasing share of old firms has reduced the economy-wide exit rate.



(a) Exit rates

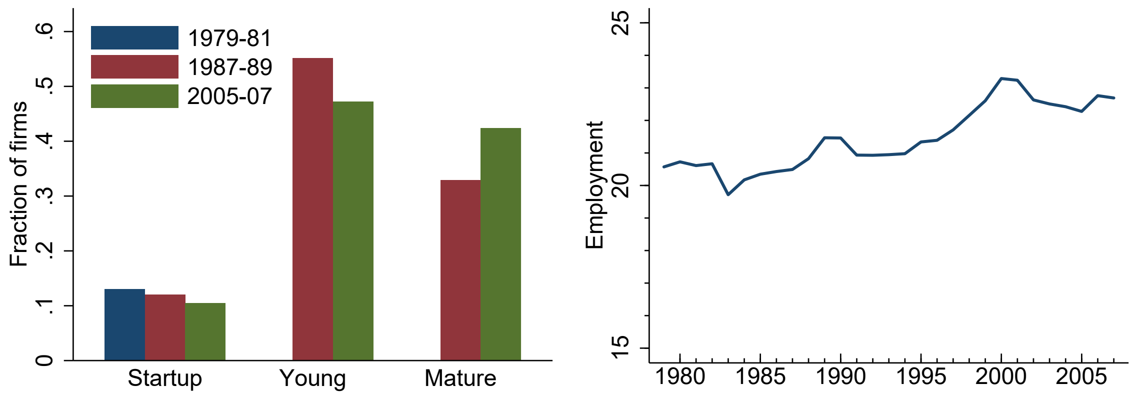
(b) Conditional growth rates

Figure 3: Incumbent exit and conditional growth rates for young and mature firms

Finally, we note similar stability in average employment growth (Figure 3b). For both age groups, this conditional growth rate is procyclical, but there is no evidence that it is systematically drifting over time. These patterns in exit and in employment growth rates are robust even when grouping by size and age (Appendix C.1.5). Taken together, the stability of entrant size and growth and survival conditional on entry, argue against a change in entrant selection, where startups are fewer but better.

Shift in firm age composition and aggregate effects. The shortage of entrants, or startup deficit, with no change in firm exit or growth to offset it gradually accumulates in a compositional shift toward older firms. Figure 4a shows the age distribution of firms in 1979-81, 1987-89 and 2005-07. Note that since birth years are left-censored, we can only show the share of startups for 1979-81. But even comparing 1987-89 to 2005-07, we see a clear shift towards older firms in the U.S. economy with the share of mature firms (ages 11+) increasing from 33 percent to 42 percent.

For any firm-level outcome that varies by firm age, there is scope for an aggregate compositional effect. We earlier discussed how age composition explains the decline in the economy-wide exit rate, and there are more examples. Average firm size by firm age has been mostly stable (Figure 2b), but overall average firm size has been increasing (Figure 4b). The explanation is again age composition: firms are not any



(a) Firm age distribution

(b) Overall average firm size

Figure 4: Change in firm age composition its effect on average firm size

Note: Young (ages 1-10) and mature (ages 11+) incumbent age groups cannot be measured in 1979-81 period.

bigger by age, but there are more old firms, and they tend to be larger.

Taking stock. Over almost three decades, economy-wide entry and exit rates have declined. These changes have been broad-based across sectors and geographic markets. Yet, conditional on firm age, exit and average employment growth have changed little, and the size distribution of startups has been stable. These patterns naturally restrict the set of explanations for the startup deficit. For example, while rising entry costs would reduce the startup rate, they would also alter the size and profitability of new entrants and by extension their expected growth and survival.

In contrast, declines in labor supply growth can explain these patterns. They generate broad-based declines in the equilibrium startup and exit rates, and they do so without altering other margins, except through changes in firm age composition. We show these effects first in a simple framework that satisfies the stability restrictions by assumption. Later, in Section 3, we show the same results in a model of firm dynamics, where stability by firm age emerges *endogenously* as a result of free entry.

2.2 A flow balance framework for the equilibrium startup rate

To see why changes in the growth rate of the labor supply might affect the firm entry rate, it is helpful to start with a simple model. Consider an economy with μ_t identically sized firms in year t and where M_t of those firms are new entrants. The startup rate for this economy is $SR_t \equiv M_t/\mu_t$. Exit is *exogenous* at rate x. Labor supply, N_t, is growing exogenously at rate η. With these assumptions, and letting $\bar{\cdot}$ denote per worker, the number of firms per worker, $\bar{\mu}_t \equiv \mu_t/N_t$, evolves as:

$$\bar{\mu}_{t+1} = \frac{1-x}{1+\eta}\bar{\mu}_t + \bar{M}_{t+1}. \tag{1}$$

A balanced growth path with a constant measure of firms per worker, $\bar{\mu}$, requires that the number of firms grow so that labor demand can meet the increase in labor supply. In this case, from (1), we can determine that the startup rate must equal:

$$SR = \frac{\eta + x}{1+\eta}. \tag{2}$$

This "flow-balance" startup rate equates the inflow of new firms per worker with the outflows per worker, where outflows are increasing in the exit rate, x, and the growth rate, η, in the number of workers.[8] The faster the labor supply expands ($\uparrow \eta$) the higher the required startup rate, since as labor supply expands, the additional labor demand to clear markets must come from *new* firms. This feature follows immediately because of the assumptions of constant firm size and the exogenous exit rate, but it extends to the more realistic equilibrium setting in Section 3 where we relax these assumptions on size and exit.

In spite of its simplicity, this stylized framework is useful is assessing the quantitative importance of the demographic channel. We use equation (2) to compute the flow balance startup rates for the 1979-81 and 2005-07 periods using the average exit rate and the average WAP (CLF) measure of labor force growth from each period (Table 1).[9] The flow-balance startup rate indicates the startup rate we would expect in the long run if labor supply growth and exit remained fixed at these values. The actual startup rate declined 2.9 percentage points from its 1979-81 average of 13.0 percent to an average of 10.1 percent for 2005-07. For the same period, the flow

[8]This calculation is conceptually similar to the flow-balance employment calculation implemented in Elsby, Michaels and Ratner (2017). They impose balance of inflows and outflows for each employment level in the establishment firm size distribution while we impose balance of the inflow of new firms per worker to the outflow of firms per worker. Equation (2) would apply equally well to any model with a stationary number of firms per worker.

[9]Average of HP filtered trends ($\lambda = 6.25$). Results are nearly identical if we instead predict annual flow balance startup rates and average over 3-year periods (Table C.1). We plot the entire time series for annual flow balance startup rates against the actual startup rate in Figure C.3.

balance startup rate from equation (2) using WAP (CLF) declined from 11.2 (11.7) percent to 9.7 (9.7) percent, a decline of 1.6 (2.0) percentage points. Taking the changes in η and x as exogenous, according to the flow balance calculations, they explain roughly two-thirds of the decline in the startup rate. Importantly, accounting for the declining startup rate requires exogenous changes in both η and x. Changing only η and keeping x at its 1979-81 average, the predicted startup rate declines 0.7 (1.0) percentage points—about half of the decline when both are changed.

Table 1: Actual and predicted flow balance startup rates

	Labor Supply Growth (%)		Exit Rate (%)	Startup Rate (%)		
				Actual	Predicted	
	WAP	CLF			η =WAP	η =CLF
1979-1981	1.9	2.5	9.5	13.0	11.2	11.7
2005-2007	1.1	1.1	8.7	10.1	9.7	9.7
Change	−0.8	−1.4	−0.9	−2.9	−1.5	−2.1

Note: Trend component of annual rates using HP filter and smoothing parameter 6.25 averaged over 3-year time periods. Startup rate, exit rate, and labor supply growth rates for working-age population (WAP) and civilian labor force (CLF) measured as described in text. Predicted startup rates use flow balance equation (2) with 3-year averages for η and exit.

While our simple framework quantifies the determinants of the declining startup rate, it has several limitations. First, exit is exogenous, and empirically, the declines in the exit rate are nearly as important as the declines in labor force growth in the flow balance calculation. While the demographic driven decline in labor force growth is plausibly exogenous, changes in exit are certainly not. Second, the constant firm size rules out incumbent employment growth and, by implication, any heterogeneity in age and size—all of which may interact with the changes in labor supply. Finally, it relies on a long-run comparison. Motivated by the findings of our simple framework, we extend a standard Hopenhayn and Rogerson (1993) model of equilibrium firm dynamics that relaxes these assumptions by endogenizing the exit and employment growth margins as well as allows us to compute transitional dynamics.

3 Firm dynamics with labor supply growth

In this section, we construct and calibrate an equilibrium model of firm dynamics with endogenous entry and exit. We then examine the balanced growth path of the model to illustrate the economic significance of the labor supply growth channel and its alternatives. Finally, allowing for transitional dynamics, we quantify the effects of the decline in labor supply growth over the 1979-2007 period.

3.1 Model

Our model economy consists of a continuum of firms owned by a growing representative household that inelastically supplies its labor to the firms. Time is discrete, and there is no aggregate uncertainty. Markets are competitive, and the prices and firm dynamics are determined in general equilibrium. We characterize the balanced growth path that grows at the same constant rate as labor supply.[10]

Households. Each member of the representative household is endowed with one unit of labor. To account for labor supply growth, the number of members (labor supply), H, grows at rate η.[11] We characterize the household's problem recursively and in per capita terms, which we denote with $\bar{\cdot}$. Letting $\bar{W}(\bar{b})$ represent the per capita household value in terms of per capita savings, $\bar{b}$, then $\bar{W}(\bar{b})$ satisfies:

$$\bar{W}(\bar{b}) = \max_{\bar{c},\bar{b}'} \left\{ \log(\bar{c}) + \beta(1+\eta)\bar{W}(\bar{b}') \right\}, \tag{3}$$

subject to the budget constraint $\bar{c} + \bar{b}'(1+\eta) = (1+r)\bar{b} + w + \bar{\pi}$. Here, β is the time discount factor, η accounts for the growth in the number of household members, w is the real wage, r is the real interest rate, and $\bar{\pi}$ the per capita dividend from aggregate firm profits. The consumption good is the numeraire.

Incumbent firms. A continuum of firms use labor, n, as the only input to produce output. Each firm has a decreasing returns to scale technology, $f(a, s, n) = asn^{\theta}$ with $\theta < 1$. A firm's productivity, as, is characterized by a permanent component, a, and

[10] Along the BGP prices and firm dynamics are constant and we drop any dependence on time t from the notation. We defer to Appendix B.2 the details of the non-stationary economy.

[11] As is customary, we model population growth as a growing representative dynastic household rather than a growing population of fixed-size households. With an appropriately chosen discount rate, this choice makes no difference along a balanced growth path.

a stochastic component, s, evolving in logs as $\log s' = \rho \log s + \sigma_\varepsilon \varepsilon$, with ε standard normal. For simplicity, we abstract from aggregate productivity growth.

Firms pay a fixed operating cost, c_f, in units of output each period.[12] After production, some firms exit exogenously with probability δ, or endogenously if the value of continuing becomes negative. Any exit occurs before s' is realized.

We characterize the incumbent firm's problem recursively. Given wage w and interest rate r, the value, V, of a firm with productivity as satisfies:

$$V(a,s) = \max_n \left\{ asn^\theta - c_f - wn + \frac{1-\delta}{1+r} \max_{X \in \{0,1\}} \{E_s V(a,s'), 0]\} \right\}. \tag{4}$$

The inner maximization reflects the end of period exit decision given the expected continuation value, where $X(\cdot)$ is 1 if the firm decides to exit and 0 otherwise. We let $h(\cdot)$ denote the optimal demand for labor, n, from the outer maximization.

Entry. A perfectly elastic supply of potential new firms can enter the economy by paying an entry cost, c_e, specified in units of output. After paying the entry cost, firms draw the fixed component of their productivity, a, and the initial stochastic component, s, from log-normal distributions $F(\cdot)$ and $G(\cdot)$, respectively, and produce within the same period. In general, free entry requires:

$$c_e \geq \iint V(a,s)\, G(ds)\, F(da), \tag{5}$$

where the inequality is strict in an equilibrium without entry. The free entry condition is satisfied with equality for a balanced growth path with constant prices when labor supply is growing, $\eta > 0$; since without entry, clearing the labor market through only incumbent labor demand would require a continually declining real wage.

Entrant and incumbent dynamics. Let M be the measure of entrants, and let $\mu(A,S)$ be the measure of all firms (entrants and incumbents) with productivity components $a \in A$ (permanent) and $s \in S$ (stochastic). Define $\bar{M} = M/H$ and $\bar{\mu} = \mu/H$ as per capita (equivalently, per worker) versions. Then, next period's

[12]Because we abstract from aggregate productivity growth the choice of costs in terms of output rather than labor makes no difference along the balanced growth path.

measure of firms per capita with $a \in A$ and $s' \in S'$ is determined by law of motion

$$\bar{\mu}'(A, S') = \iint_{(a,s)\in A\times\mathbb{R}_{>0}} (1-\delta)(1-X(a,s))P(S'|s)\frac{\bar{\mu}(da,ds)}{1+\eta} + \bar{M}'G(S')F(A), \tag{6}$$

where $P(S'|s)$ is the conditional probability of $s' \in S'$ implied by the process for $\log s'$, and $\bar{\mu}$ is scaled by $1/(1+\eta)$ to express it in terms of next period's labor supply, H'.

Closing the model. Aggregate profits per capita paid as a dividend to each member of the representative household must equal total production per capita net of fixed and entry costs:

$$\bar{\pi} = \iint \left(ash(a,s)^{\theta} - wh(a,s) - c_f\right)\bar{\mu}(da,ds) - \bar{M}c_e. \tag{7}$$

Balanced growth path. A balanced growth path equilibrium growing at rate η consists of a constant real wage w, interest rate r, consumption $\bar{c}$, savings $\bar{b}$, and profits $\bar{\pi}$, all per capita, labor demand and exit policy functions $h(\cdot)$ and $X(\cdot)$, an invariant measure of firms $\bar{\mu}$ and entrants $\bar{M} > 0$, both per capita, such that: i) given w, r and $\bar{\pi}$ that $\bar{c}$ and $\bar{b}$ solve (3); ii) firm labor demand, $h(\cdot)$, and exit, $X(\cdot)$, solve (4); iii) free-entry condition (5) is satisfied with equality; iv) measure of firms, $\bar{\mu}(\cdot)$, and entrants, $\bar{M}$, satisfy law of motion (6); and v) markets clear so that (7) is satisfied.

3.2 Calibration

We set the model period to be a year and calibrate the model to match a rich set of annual firm dynamics statistics. Because the overall firm age composition is shifting throughout the 1979-2007 period, there is no specific year or years that would correspond to a balanced growth path through the lens of the model. So we cannot target moments such as the startup rate, average firm size, or even labor supply growth that are all shifting over time. Instead, our strategy is guided by the findings of Section 2; we target a set of moments, all conditional on firm age, that have remained stable throughout the 1979 to 2007 period amid the decline in labor supply growth.

To implement this approach, we first calibrate two parameters outside the model: the time discount rate to 0.96 (real interest rate of $\beta^{-1} - 1 = 4.2\%$) and the curvature parameter of the production function θ to 0.64, which matches the labor share of total revenue. The remaining 8 parameters, which govern firm dynamics, are calibrated internally by matching a set of data moments. These eight parameters are: the entry cost c_e, the operating cost c_f, the exogenous exit rate δ, the persistence and standard deviation of the idiosyncratic shocks ρ and σ_ε, the parameters μ_0 and σ_0 governing the log-normal distribution G of its initial condition, and finally the dispersion σ_a of the log-normal distribution F of permanent types (its mean is normalized so $E[a] = 1$).

Table 2: Values for internally calibrated parameters

c_e	c_f	δ	ρ	σ_ε	μ_0	σ_0	σ_a
7.594	2.296	0.018	0.972	0.032	0.142	0.138	0.078

For targets, as discussed above, we choose a set of 23 moments that remained stable throughout the 1979-2007 period. First, we target an average startup size of 6 employees, the average over our sample period. Second, we target the conditional growth rates of three-year old firms by their initial size. Specifically, we group three-year old firms into three size bins (0–49, 50–249, and 250+ employees), and compute the growth in average firm size for the firms in each initial size bin.[13] Averaged over the sample period, these growth rates are 8.1, 1.2, and -3.0 percentage points for the 0–49, 50–249, and 250+ initial size groups, respectively. Finally, we target the exit rate of firms by age out to age 19 from Sterk ⓡ Sedláček ⓡ Pugsley (2021), calculated on the LBD microdata for the 1979 to 1993 pooled birth cohorts of firms.

We then choose the 8 parameters by minimizing the weighted squared distance between the data and model counterparts for the 23 targets and report these parameters in Table 2.[14] Using these parameters, Figure 5 compares the model-implied statistics with the data. Despite targeting 23 moments with only eight parameters, the model captures the key features of the U.S. firm dynamics very well. For our application, matching the exit rate is particularly important since conditional on a growth rate η, the exit rate effectively determines the long-run firm age distribution and the startup rate. An excellent fit is feasible, even with few parameters, by allowing for permanent heterogeneity in long-run productivity through a and an unconditional probability of exogenous exit δ (Sterk ⓡ Sedláček ⓡ Pugsley, 2021).

Finally, we note that we do not need to calibrate the labor supply growth rate

[13] We use a specific age to control for any age composition changes within the broader young group.

[14] We weight the startup size and growth moments 10 percent each to ensure we fit these moments well, and we split the remaining 60 percent equally across the 19 exit rates (Appendix B.5).

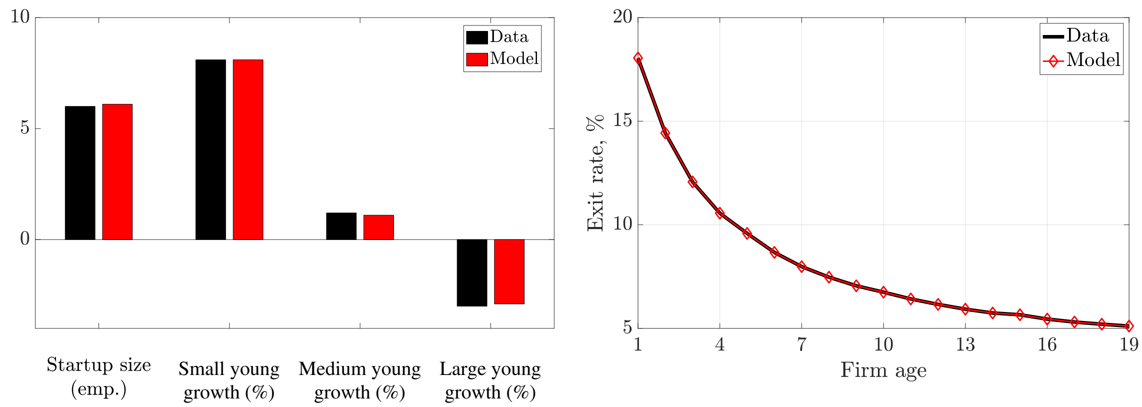


Figure 5: Model fit to targeted moments

parameter, η. A benefit of our calibration strategy, and a feature we further discuss below, is that the target moments, all conditional on firm age, are independent of the labor supply growth rate. This allows us to characterize the model's predictions for the startup rate and firm age composition for a range of labor supply growth rates while still matching the same set of calibration moments throughout.

3.3 Equilibrium effects of changing demographics

The model provides a natural quantitative framework to examine our hypothesis that the labor supply growth rate is a first order determinant of firm entry and exit. Using the same baseline calibration, we study the long-run equilibrium effects of changing labor supply growth, η, by comparing the balanced growth paths across a range of labor supply growth rates from 0 to 4%. The intuition is analogous to the simple example in Section 2.2 except now the effects on the startup rate of changes in labor supply growth are amplified by the equilibrium response of the endogenous exit rate.

Main effects. The upper left panel of Figure 6 shows the first order importance of labor supply growth for the economy's startup and exit rates. Starting from the case of zero labor supply growth, we increase η from 0 to 4%. In the calibrated model, this increases the startup rate along a balanced growth path from 7.8 to 13.1% and the exit rate from 7.8 to 9.6%.

To interpret these effects, it is helpful to refer back to the simple flow-balance formula in equation (2) from Section 2.2. In this formula, $\frac{\eta+x}{1+\eta}$, the startup rate is determined by two channels: labor supply growth rate, η, and aggregate exit rate, x.

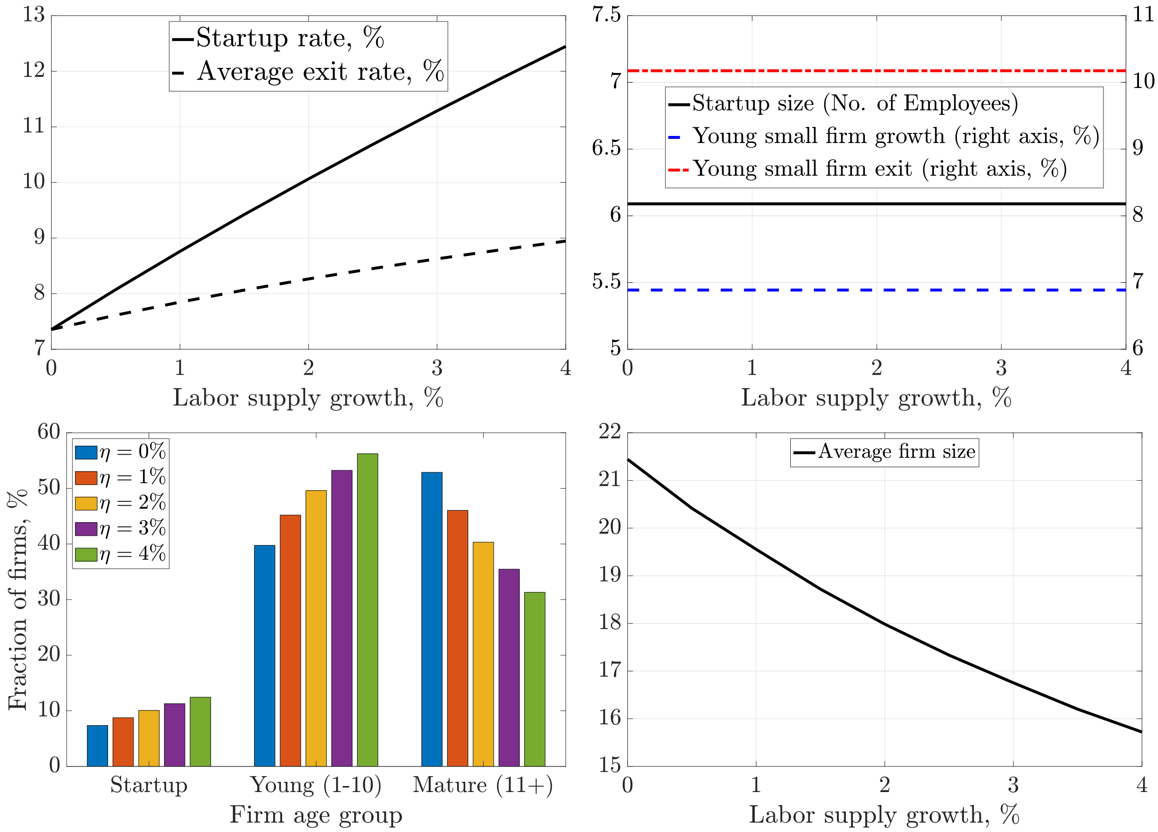


Figure 6: Labor supply growth and firm dynamics

Note: Young small in the upper right panel refers to age 3 firms with 1-50 employees.

These reflect the net growth and replacement components, respectively, of the startup rate. With $\eta = 0$, the startup rate equals the exit rate since new firms simply replace exiting firms. As η increases, the startup rate exceeds the exit rate since entering firms replace exiting firms and increase the net number of firms. This formula still holds in the full model, only x is now determined *endogenously*.[15]

Increases in the startup rate with labor supply growth arise because of effects on both channels. Totally differentiating (2) allowing for any effects on equilibrium exit gives an expression for the labor supply growth elasticity of the startup rate:

$$\frac{dSR}{d\eta} = \frac{1 - SR}{1 + \eta} + \frac{\frac{dx}{d\eta}}{1 + \eta}. \tag{8}$$

[15]This follows from integrating the law of motion in equation (6) for the balanced growth path over all firms. See Appendix B.3.

Its two terms capture the two sources of effects on the startup rate. The first term is the direct effect of labor supply growth. Increasing the growth rate of the workforce requires an increase in the growth rate of the number of firms to keep pace. The second term is an indirect amplification effect. Labor supply growth changes the replacement component of the startup rate vis-à-vis the startup rate's effects on firm age composition and the aggregate exit rate.

In the calibrated model, this elasticity (also the slope of the startup rate curve in Figure 6) is 1.3, which is the sum of direct effect 0.9 and indirect effect 0.4. While the direct effect is independent of any calibration, the indirect effect depends fundamentally on the response, $dx/d\eta$, of equilibrium exit to the change in labor supply growth. We now examine how this response is determined, first by firm age and then allowing for changes in the firm age composition.

Effects conditional on firm age. We find that all moments conditional on firm age are unaffected by η. The upper right panel of Figure 6 shows the average size of entrants, along with the exit rate and conditional growth rate of small young firms across the same 0 to 4% range of η. These margins are all constant.

This is consistent with our empirical findings in Figures 2 and 3 from Section 2. In the data, the exit, size, and growth margins are mostly stable when conditioned on firm age, despite the demographic changes and trend decline in the startup rate. In that sense, the constancy of these conditional moments in the model can be viewed as a set of testable implications for our demographic account of the decline in the startup rate, which the data strongly support.[16] With margins such as the exit rate by firm age unaffected by shifts in η, the aggregate response $dx/d\eta$ can only occur through changes in composition.

Aggregate effects through firm age composition. Changes in η shift the firm age distribution through their effects on the startup rate. The lower left panel of Figure 6 shows the share of firms by firm age for different levels of η. Higher levels of labor supply growth require a higher equilibrium entry rate, and with relatively more new firms becoming young firms, the share of older firms declines. Increasing η from 0 to 4% decreases the mature share of firms from 50 to 30 percent.

Aggregates are affected through these changes in firm age composition. For the

[16] Pugsley and Şahin (2019) show the slope of the trend component of these margins is statistically indistinguishable from 0.

aggregate exit rate, x, in the upper left panel of Figure 6, a higher η raises the aggregate exit rate by increasing the share of new and young firms, since these firms have higher exit rates independent of η. Thus the amplification effects of endogenous exit on the startup rate work entirely through shifts in firm age composition—the excellent fit on the observed exit rates (Figure 5) is paramount to the credibility of the amplification effect. The composition channel extends to effects on other economy-wide averages such as firm size (Figure 6 lower right). Since new and young firms are smaller, increases in their share resulting from higher η reduce average firm size.

The implications of a *decline* in labor supply growth for firm age composition are also consistent with the empirical patterns in Figure 4 from Section 2. In the more recent period of low labor supply growth, relative to the earlier period with high labor supply growth, the mature share and average firm size both increased significantly.

3.3.1 Role of free entry

Along the balanced growth path, the effects of η, operate entirely through the startup rate. The invariance of incumbent behavior is a consequence of assuming free entry, which holds the real wage constant along a balanced growth path.[17] We first discuss the role free entry plays in the mechanism and why we believe it is a reasonable assumption for the long run. We then show even for plausible deviations from free-entry that the response of the startup rate is still quite large.

Mechanism. As is well known from Hopenhayn (1992), the free entry condition in (5) makes labor demand infinitely elastic at the real wage that equates the expected value of a potential entrant to the entry cost. Without entry, firms' decreasing returns to scale technology would imply downward sloping aggregate labor demand, and any shock to labor supply would require a wage adjustment to clear labor markets. With entry, the *number* of entrants can expand or contract, relieving any pressure on the real wage to adjust.[18] Incumbents are unaffected, since the real wage remains fixed.

This property extends to changes in labor supply growth, η, because, shifts in η, like shocks to the level of labor supply, have no direct effects on the value of incumbent

[17]The invariance of incumbent dynamics result would still go through by allowing for aggregate productivity growth as long as all costs also grow with productivity. Bollard, Klenow and Li (2016) find ample evidence for costs rising with productivity. In this case, increases in productivity would scale the equilibrium real wage and profits as well as entry and overhead costs all in proportion, leaving firm-level equilibrium allocations unchanged.

[18]See, e.g., Luttmer (2010) on the link between the number of firms and number of workers.

firms and thus no effect on the expected value of a potential entrant. Labor supply growth, η, can only affect firm behavior through its effects on the real wage, since constant consumption per capita already holds the real interest rate fixed. But, with expected value of a potential entrant unaffected, the real wage is also unaffected since it still satisfies the free entry condition (5). With no incumbent response, any changes in aggregate labor demand needed to clear markets must originate from changes in the entry margin, making the startup rate highly responsive to η.[19]

While entry is not literally *free* since entrants must pay the entry cost c_e, this cost is independent of the number of entrants. Implicitly, any amount potential entrants are willing to enter so long as the expected value of entering covers the entry costs. We believe this is a reasonable assumption for the long run since if the value of an expected entrant were to exceed the entry cost, given enough time, entrepreneurs would take advantage of these business opportunities. However, anything that restricted the supply of potential entrants, e.g. incumbent market power or the opportunity cost of would be entrepreneurs, would limit the responsiveness of the startup rate, forcing instead an adjustment of the equilibrium real wage that allows incumbent margins to adjust and help clear labor markets.

Deviations from free entry. We now relax the assumption of a perfectly elastic supply of potential entrants and study quantitatively how it dampens the responsiveness of the startup rate to changes in labor supply growth. To do this, we allow the previously constant entry cost of a potential entrant, c_e, to instead depend on measure of entrants per capita as follows:

$$c_e(\bar{M}) = c_{e0}\left(\frac{\bar{M}}{\bar{M}_0}\right)^{\psi-1} \qquad \psi \geq 1. \tag{9}$$

Parameter c_{e0} is the entry cost in our baseline calibration, and $\bar{M}_0$ is the measure of entrants per capita from the baseline economy with free entry and $\eta = 2.2\%$.[20] Parameter ψ controls the convexity of aggregate entry costs.

A benefit of this formulation is the convexity, ψ, provides a simple way to adjust

[19] There is an important distinction between changes to the level and growth rate of labor supply. Whereas free entry assures shifts in the level of labor supply can only scale up or down the size of economy leaving the distribution of firms (and startup rate) unaffected, shifts to the *growth* rate of labor supply, affect the startup rate and the distribution of firms.

[20] This assures for $\eta = 2.2\%$ (average of WAP and CLF for the 1979-81 period) that equilibrium entry costs will be the same regardless of the value of ψ; they will deviate as η changes.

the elasticity of supply of potential entrants. Under free entry, $\psi = 1$, $c_e = c_{e0}$ is constant, aggregate entry costs per capita, $c_e(\bar{M})$, are linear in the measure of entrants per worker, and there is a perfectly elastic supply of potential entrants. When $\psi > 1$, entry costs, $c_e(\bar{M})$ are increasing in the number of entrants, and aggregate entry costs, $c_e(\bar{M})\bar{M}$ become a convex function of $\bar{M}$, making the supply of entrants less elastic.[21]

As a comparison to our baseline, we consider several alternative values of ψ where aggregate entry costs become progressively more convex in the number of entrants, i.e, entry becomes less and less free. We choose this range in line with Gutiérrez, Jones and Philippon (2019) who estimate ψ using industry variation and find a range between 1.2 and 2, with a median value of 1.55. For each ψ, we compute the elasticity of startup rate with respect to labor supply growth from equation (8). Table 3 reports the results and shows that at their median estimate of convexity, ψ, the deviation from free entry attenuates the elasticity of the startup rate by about 25% (13% to 40% between their 10th and 90th percentiles).

Table 3: Sensitivity of labor supply growth elasticity to deviations from free entry

	Convexity of aggregate entry costs, ψ				
	Free entry	Less than free entry			
	1 (Linear)	1.1	1.2	1.5	2
Elasticity of startup rate, $\frac{dSR}{d\eta}$	1.27	1.18	1.11	0.93	0.77

Note: Baseline model with entry cost c_e replaced by equation (9). Elasticity estimated by regression of SR on η over $\eta \in [0, 4\%]$. Further details in Appendix B.4.

3.4 Considering alternative channels

Looking beyond demographics, changes in the cost structure could also depress the startup rate over time. A key question is how to evaluate the plausibility of potential changes in the economy's cost structure since unlike the labor supply growth rate, the parameters governing costs do not have obvious empirical counterparts that are independent of the model structure.

To explore these alternatives, we adopt the following approach. We ask in Table 4 how much of a change in each cost parameter individually is required to generate

[21] In general, the elasticity of supply is inversely proportional to $\psi - 1$. (Appendix B.4)

the 2.9 percentage point decline in the startup rate in the data between the 1979-81 to 2005-07 periods.[22] We then examine the implications for the aggregate exit rate, average firm size, startup size and young incumbent exit and growth rates, and we find along these margins that cost changes present some challenges to fit the data.

Table 4: Examining alternative channels for decline in startup rate

		Potential Channels		
	Actual Change	Entry cost, c_e	Operating cost, c_f	Labor supply growth, η
Panel A. Explaining the long-run decline in the startup rate				
Required parameter change	—	122.6%	−55.7%	−2.2(pp)
Panel B. Implied change in each margin				
Startup rate (pp)	−2.9	−2.9	−2.9	−2.9
Economy-wide exit rate (pp)	−0.9	−3.0	−3.0	−1.0
Average firm size (emp)	2.0	6.4	−10.7	3.7
Startup size (emp)	0.1	3.3	−3.4	0.0
Young small exit rate (pp)	−0.1	−3.2	−3.6	0.0
Young small growth rate (pp)	0.5	−0.4	−0.8	0.0

Note: Required change in each parameter relative to baseline calibration with $\eta = 2.2\%$ to reduce the startup rate by 2.9 ppts. Implied changes compares new model moments relative to their baseline values. Young small refers to age 3 firms with 1-50 employees as in Figure 6.

Changes in the entry cost. Stricter regulations and institutional constraints that make it harder for new businesses to enter could be responsible for the decline in the startup rate. The model captures this channel as a rise in the entry cost, c_e. The entry cost column of Table 4 shows that explaining a 2.9 percentage point decline in the start-up rate (Panel A) requires a 122.6% increase in the entry cost. The implied changes in the other margins (Panel B) are at odds with the data. Aggregate exit would have declined by 3 percentage points relative to a 0.9 percentage point drop in the data. Overall average firm size increases by 6.4 employees (increase 2.0 data) and average startup size increases by 3.3 employees (increase 0.1 data). Because the increase in the entry cost, through free entry, reduces the real wage, incumbent margins are also affected. Conditioning on young and small firms, exit declines by

[22] An earlier version of our paper also considered convex labor adjustment costs and found that the startup rate was not very sensitive to changes in their strength. Even if empirically adjustment costs have increased, as found by Decker, Haltiwanger, Jarmin and Miranda (2020), this will have had little effect, if any, on the startup rate.

3.2 percentage points (0.1 decline in data) and the conditional growth rate falls by 0.4 percentage points (0.5 *increase* in data).

Changes in the operating costs. Regulations or other frictions could also raise the operating costs for businesses, which we capture by a change in the fixed cost parameter, c_f. The operating cost column of Table 4 shows that to explain the 2.9 percentage point decline in the startup rate, c_f would have to *decline* by 55.7 percent; an increase would induce more exit among incumbents and thus increase the startup rate through its replacement component. The required decline in operating costs over time stands in contrast to evidence of increasing regulatory costs of running a business in the U.S. (Davis and Haltiwanger, 2014). Moreover, the implied changes in the other margins are well out of line with their data counterparts. For example, overall average firm size declines by 10.7 employees, whereas it increases by 2.0 employees in the data.

To sum up, cost-based channels are feasible explanations for the decline in firm entry, but they face challenges when matching other margins. As a point of comparison, the last column of Table 4 applies the same procedure to labor supply growth. Matching the 2.9 percentage point decline in the startup rate requires a 2.2 percentage point decline in η, and the implications for the other margins are much closer to the data. We now quantify this channel using the model's transitional dynamics.

3.5 Effects of Demographics over 1979-2007 period

Over the 1979-2007 period, demographics explain between roughly 1/3 and 60 percent of the 2.9 percentage point decline in the startup rate. To reach this conclusion, we examine the *transition* path of the startup rate over this period consistent with the observed path of labor supply growth. While the preceding comparative statics illustrate the long-run effects of a shift in labor supply growth, it is difficult to argue that either of the early or late periods in the data corresponds to a balanced growth path. Analyzing the model off its balanced growth path requires characterizing the non-stationary equilibrium, and we provide the details in Appendix B.2.

Using the non-stationary model, we compare the 1979-81 and 2005-07 periods along the transition path resulting from the following experiment. We assume the economy in 1948 is on a balanced growth path corresponding to a constant $\eta = 1.1\%$. Then, we consider a perfect foresight shock of the observed path of η_t from 1949 to 2019, followed by constant growth at 1% thereafter. For the observed path of η_t

we use the trend component of working-age population growth (also civilian labor force growth) to be consistent with the results in Table 1. For each, we solve for the equilibrium transition path that converges to a balanced growth path with $\eta = 1\%$.

Table 5: Actual and predicted declines in the startup rate

	Labor Supply Growth (%)		Startup Rate (%)			Economy Exit Rate (%)		
			Actual	Model		Actual	Model	
	WAP	CLF		WAP	CLF		WAP	CLF
1979-1981	1.9	2.5	13.0	10.0	10.6	9.5	8.6	8.9
2005-2007	1.1	1.1	10.1	9.0	8.9	8.7	8.5	8.3
Change	−0.8	−1.4	−2.9	−1.0	−1.7	−0.9	−0.2	−0.6

Note: Averages computed using transition path in response to perfect foresight shock of 1949–2019 WAP (CLF) growth rates. See text and Appendix B.5.2 for further details.

Examining the transition paths for the startup rate under our WAP and CLF proxies for labor supply growth, Table 5 reports the full model counterpart to Table 1. The path of WAP growth explains a 1.0 percentage point decline in the startup rate, or 34 percent of the decline. In the aggregate time series, changes in CLF growth provide an even better picture of shifting labor supply growth since they include the effects of the slowdown in female labor force participation. The 1.4 percentage point decline in CLF growth explains a 1.7 percentage point decline in the startup rate, or 59 percent of the decline in the data. Moreover, the changes in labor supply growth *endogenously* generate a decline in the aggregate exit rate. The path of CLF growth explains 75 percent of the observed decline in exit.

These results enhance the flow balance calculations from Table 1 in several ways. First, as we discuss above, the model endogenizes the exit rate. Second, the flow balance calculations, even when exit is endogenized, compare two steady states, each with its own constant firm age distribution. The transition allows firm age composition to evolve consistent with the endogenous fluctuations in the startup rate. Finally, along the transition, incumbents *can* respond to fluctuations in η_t.[23] In practice, we

[23] Even though free entry is satisfied in each period, the consumption smoothing of the representative household, through its equilibrium effects on r_t, functions as an adjustment cost along the transition, just as it would for investment in a neoclassical growth model. In effect, the startup rate cannot fully adjust to clear the labor market, allowing an incumbent response. For reference, Appendix B.8 compares the dynamic response of the startup rate to a labor supply shock in the calibrated model to a case with a fixed real wage and no incumbent response.

find these incumbent effects along the transition path to be small (Table B.4), consistent with the observed stability over time of margins conditional on firm age from Section 2. That is, transitional dynamics are primarily driven by firm age composition rather than significant changes in incumbent behavior.

4 Evaluating the mechanism across U.S. states

In this section, we confront the implications of our model with rich cross-sectional data. The model has a definitive long-run empirical implication: an economy with a higher labor supply growth rate should exhibit a higher startup rate than an economy with lower labor supply growth everything else being equal. As a starting point, Figure 7 shows that U.S. states which experienced a higher average working-age population growth rate also had a higher startup rate over 1979-2007. The relationship between the two is remarkably strong: a one percentage point higher WAP growth rate is associated with 1.6 percentage point higher startup rate.[24]

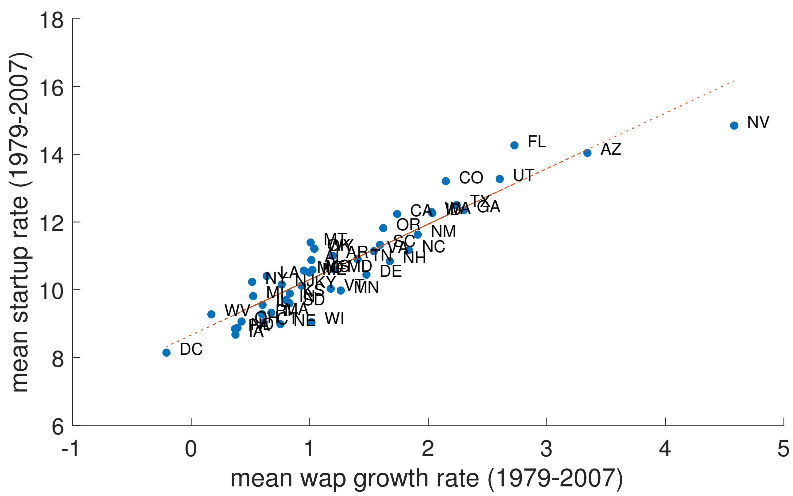


Figure 7: Average WAP growth and startup rates for U.S. states over 1979–2007.

This observation admittedly only establishes a simple correlation in the cross-state data and does not directly lend itself to a causal interpretation. The empirical challenge of testing the long-run prediction directly is to identify differences in labor supply growth that are together exogenous and permanent. To overcome this chal-

[24]We find the same slope using the 1979–2007 average CLF growth rate (Figure C.12). For brevity, we present CLF results, which are very similar, in Appendix C.2.2.

lenge, we note that even in the short run our model still predicts that the startup rate is highly responsive to shifts in labor supply growth.[25]

We focus on year-over-year changes in startup rates and labor supply across U.S. states. This strategy provides richer variation by using idiosyncratic changes in state-level labor supply growth, and it allows us to employ a powerful IV strategy to identify the *causal* effects of changes in labor supply growth on the startup rate. Clearly, the responses we estimate using this variation do not necessarily correspond to a long-run elasticity based on comparisons of steady-states. To reconcile our empirical strategy with the model, we estimate the same short-run elasticity on model-simulated data and find it lines up very well with its empirical counterpart.

4.1 Empirical strategy

We now turn to evaluating empirically the response to labor supply shocks using cross-state variation pooled over the 1979-2007 period.[26] Our main specification is the regression:

$$SR_{st} = \alpha_s + \gamma_t + \beta g_{st} + \varepsilon_{st}. \tag{10}$$

where SR_{st} is the startup rate in state s in year t, and g_{st} is a measure of the state's annual labor supply growth. The terms α_s and γ_t capture state s and year t fixed effects, respectively, and ε_{st} other unobserved determinants of the state's startup rate.

The coefficient β is the cross-sectional elasticity of the startup rate to labor supply growth g_{st}.[27] Because we use within-state variation in labor supply growth, coefficient β is not in general equal to the long-run labor supply growth elasticity of the startup rate from equation (8). Instead β captures the short-run response to a shift in labor supply growth in the non-stationary model, and we estimate this response in both

[25]Figure B.6 in Appendix B.8 examines the IRF of the startup rate to a one percentage point labor supply growth rate shock.

[26]Our choice of states as the unit of analysis requires mobility costs to be large enough to prevent geographic mobility from completely equating differences across these segmented labor markets. Drawing on Kennan and Walker (2011), who estimate an average moving cost of $312,000 (in 2010 dollars), we argue this is likely the case. 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. However, to the extent this is not true and mobility attenuates the local effects of labor supply shocks, our cross-state estimates will likely understate the effects of demographics on startups and incumbent dynamics.

[27]With the time-effect absorbing national changes, the coefficient captures only the relative responses across states within a year, and we refer to it as a cross-sectional elasticity. See the discussion of cross-sectional elasticities in Nakamura and Steinsson (2018) and references therein.

data and model. In the data, we estimate the cross-sectional elasticity using the same proxies for labor supply growth, i.e. the growth rate of the entire population ages 20-64 and the growth rate of the civilian labor force, also now measured at the state rather than national level (see Appendix A.3 for additional details).

The challenge with estimating this cross-sectional elasticity via OLS is that, even within a state, changes in the growth rate of labor supply are likely endogenous: The same unobserved time-varying features of the business environment within ε_{st} that affect firm dynamics may also influence labor supply. For example, states that become more profitable for incumbent firms may both attract new firms and new workers through internal and international migration. To generate variation in labor supply growth that is plausibly exogenous to labor demand or other related factors, we use two distinct instrumental variables identification strategies, one based on lagged birthrate patterns and the other based on patterns of inter-state migration.

Fertility instrument. Our first IV identification strategy uses within-state variation in labor supply growth rates predicted by past birthrates. In the spirit of Shimer (2001) and Karahan and Rhee (2014), we instrument each state's current labor supply growth using 20-year lags of its birthrate. To the extent that young adults enter the labor force in their state of birth, the state's lagged birthrate will affect future labor supply growth.[28] We compute these annual state-level birth rates for the years 1959-1987, which we use to instrument labor supply growth for the years 1979-2007, a 20-year lag. We show below that lagged birthrates are a strong predictor of future labor supply growth, even after conditioning on state and year fixed effects.

This strategy relies on the exclusion restriction that conditional on state and year fixed effects, higher fertility or its determinants have no effect on future firms except indirectly through their effect on labor supply growth. This requires that any lagged economic conditions that generated variation in past birthrates are uncorrelated with current business conditions. The assumption would be violated, for example, if people in a given state had a higher fertility rate 20 years ago in anticipation of persistently (but not permanently) stronger labor market conditions relative to other states.[29] We

[28] These birthrates are measured as births relative to the population and are available in the various Statistical Abstracts of the United States. We are grateful to Rob Shimer for providing us with his digitized data from the Statistical Abstracts for the period 1940–91. For additional details on the fertility instrument construction, please refer to our data appendix, section A.3.2

[29] Although for many individuals fertility is a forward looking choice variable, we believe this is a reasonable assumption for business conditions 20 years in the future. It may be violated over a

examine many potential confounding explanations in our robustness analysis.

Migration instrument. We also use an Altonji and Card (1991)-style instrument which utilizes differences across states in the historical patterns of inter-state migration. We rely on the lagged distribution of individuals' birthplaces within each state, which can be measured using historical Decennial Census data. These lagged shares then serve as weights for adding up "pushes" from other states' labor supply growth. Changes in the population growth of a state k will predict changes in the population growth of state $s \neq k$ if, historically, migrants out of state k tend to move to state s. We implement this idea as follows:

$$\hat{m}_{st} = \sum_{k \notin C(s)} \omega_{st^*}^{k} g_{kt}, \tag{11}$$

Here, $\omega_{st^*}^{k}$ is the share of residents of state s in time t^* that were born in state k and g_{kt} is the growth rate of the working-age population in k at time t. In computing $\hat{m}_{st}$, we exclude states in the same Census Bureau division $C(s)$ since the labor supply growth in neighboring states, g_{kt}, may be related to state s business conditions. To isolate the historical component of migration patterns, we use the birthplace shares $\omega_{st^*}^{k}$ from 2 censuses ago, t^*[30] Thus, if t is a census year (those ending in 0) then $t^* = t - 10$, and if t is an inter-censal year then t^* is the census year from the previous decade, e.g. for $t = 2005$, then $t^* = 1990$.

The identifying assumptions implicit in the exclusion restriction are two-fold. First, as with the fertility instrument, we assume that any changes in the lagged distribution of birth states, conditioned also on future time effects, are unrelated to differences across states in future business conditions and have no effect on firms except through their effects on labor supply. Second, we must assume that the "push" from source state k labor supply growth, g_{kt}, is also unrelated to the current business conditions in destination state s. [31] That is, labor supply growth rates in states outside of the same Census division are not driven in part by state s economic conditions.

shorter horizon. For example, at the national level Buckles, Hungerman and Lugauer (2018) find fertility to be a leading indicator of the business cycle for recent recessions and recoveries.

[30]We use the IPUMS microdata for the long form responses to the 1970, 1980, 1990 Decennial Censuses; see Ruggles, Genadek, Goeken, Grover and Sobeck (2017). In 1979, the lag is 9 years ($t^* = 1970$ instead of 1960). See Appendix A.3.2 for additional details.

[31]In Appendix C.2.1, we also consider alternative constructions of this instrument, which yield similar results.

First-stage regressions. We first examine the first-stage regressions for both instruments. The left panel of Figure 8 plots the variation in the working-age population growth rate against the fertility instrument, where each variable has already been purged of state and time fixed effects. The horizontal axis plots, for each year and state, the state's lagged fertility relative to its within-state average and the year's between-state average. Notably, even conditional on permanent differences across states, there is considerable variation in residual birthrates. The vertical axis plots the corresponding residual working-age population growth rate for each state and year. The positive correlation between these two measures confirms the relevance of the instrument. The slope is equal by construction to the coefficient on the instrument in the first-stage regression; a 10 percentage point increase in the lagged birthrate predicts a 1.4 percentage point increase in the growth rate of the working-age population (Column 1 of Table 6). The birthrate instrument is also strong, with an F statistic of roughly 33.

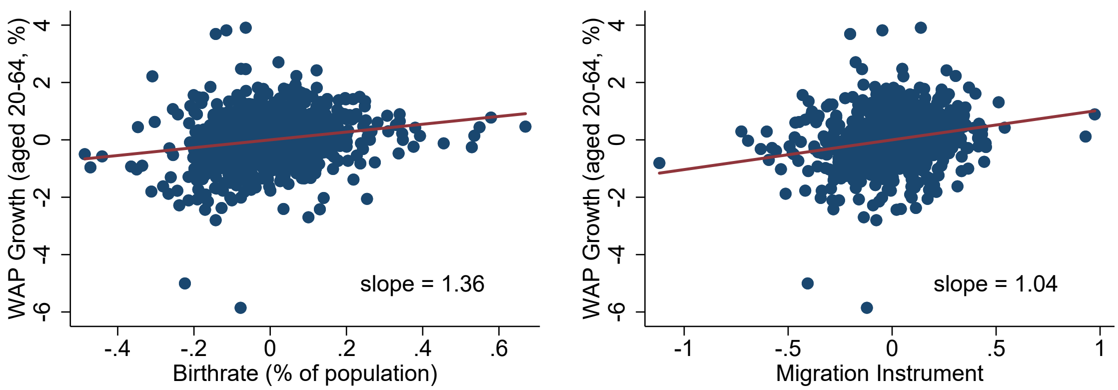


Figure 8: First-stage regressions of WAP growth rate on each instrument

Turning to the second instrument, the right panel of Figure 8 plots the variation in the WAP growth rate against the migration instrument, both residualized in a similar fashion by regressing on state and year dummies. The migration instrument, even after removing state and year fixed effects, varies significantly and predicts differences in states' WAP growth rates. A one percentage point increase in the migration "push" instrument (a weighted average of other-state WAP growth) predicts a 1.04 percentage point increase in own-state WAP growth (Column (2) of Table 6). To be conservative, our migration IV excludes "pushes" from states in the same Census division. Alternative formulations with even stronger first stage predictions have little effect on the second stage estimates (Appendix C.2.1). The migration IV also provides independent variation to the birthrate instrument. When used together, Column (3) of Table 6 shows that each IV has predictive power conditional on the other.

Table 6: Regression of startup rate on working-age population growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	First Stage			OLS	IV_1	IV_2	IV_1&IV_2	Model
WAP Growth (%)				0.61	1.09	1.27	1.19	1.10
				(0.05)	(0.32)	(0.22)	(0.22)	(0.00)
Birthrate IV	1.36		1.11					
	(0.24)		(0.24)					
Migration IV		1.04	0.87					
		(0.30)	(0.28)					
N	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R^2	0.64	0.64	0.65	0.90	0.87	0.85	0.86	0.90
F-test	32.63	11.82	17.46					
p-value of J test							0.55	

Note: Standard errors are clustered on state. Regressions contain state and year fixed effects and cover years 1979-2007 and 48 contiguous states plus D.C. Column (8) is OLS regression with state and period fixed effects on simulated panel of 49 states and 29 periods (Appendix B.9); results averaged over 100 simulated panels; standard error (0.0037) is standard deviation of estimates across panels.

4.2 Labor supply growth and the startup rate

We now estimate (10) on pooled state-level data for the period of 1979-2007 to evaluate the effect of changes in labor supply growth on the startup rate. Our benchmark estimates use the growth rate of the working-age population as a proxy for labor supply growth, but they are robust to the alternative civilian labor force growth measure (Appendix C.2.2). All specifications include state and year fixed effects. Where possible, we utilize the publicly-available data from the BDS to facilitate replication; we turn to the confidential micro data from the LBD only to include detailed industry controls. Throughout, standard errors are clustered by state.

Table 6 presents the results. The OLS estimate (Column 4) shows that a one percentage point increase in labor supply growth is associated with a 0.61 percentage point increase in the startup rate. Column (5) of Table 6 presents the results using the fertility instrument. According to this estimate, a one percentage point reduction in the working-age population growth rate leads to a 1.09 percentage point decline in the startup rate. Using the migration instrument in column (6), we find an elasticity of 1.27. When combined with the fertility IV, the elasticity is 1.19. Incorporating two instruments also permits a J test of over-identifying restrictions, which we would fail to reject at any level below 0.55. Given the similarity of the estimates, we take the results using both IVs, Column (7), as our benchmark.

Comparison with the model. We next compare the estimated cross-sectional elasticity with its counterpart in the calibrated model. To do this, we subject our calibrated equilibrium model to exogenous shocks to labor supply growth, η_t, and create a simulated panel of 49 states for 29 periods to mimic the data (see Appendix B.9 for details). In Column (8) of Table 6, we estimate (10) on simulated data with $g_{st} = \eta_{st}$, also with state and time period fixed effects, and we report averages over 100 simulated panels. The model-implied elasticity is 1.10, which is very close to the data estimates (1.09 to 1.27) and well within their 95% confidence sets.

OLS versus IV estimates. Also of note, the estimated effects on the startup rate by OLS are roughly half as large as those estimated using the two instruments either individually or in conjunction. Our model provides an intuitive interpretation for this finding. Exogenous shifts in labor supply growth, η_t, have no direct effect on firm values and are largely absorbed by the entry margin, but this is not true in general for other determinants of WAP growth. For example, a change in productivity might also affect an area's WAP growth by attracting new workers from other areas. And, because of the change in productivity's effects on incumbent firm values and their behavior, it would require a smaller equilibrium response of the startup rate. The OLS estimate combines the effect of shifts in labor supply growth with all other drivers of WAP growth, and it implies a lower coefficient compared to IV estimates which isolate the response to changes in labor supply.

A corollary to this intuition implies that the opposite should be true for the incumbent margins. If specification (10) were estimated for incumbent margins, the OLS estimate of each response should capture a larger effect than the IV. This is indeed the case: OLS estimates for the response of incumbent margins are statistically and and economically larger than their IV estimates (Appendix Table C.11).

4.2.1 Labor supply growth versus age composition

Empirically, the slowdown in labor supply growth coincides with a shift towards older workers. One question is whether the elasticity we estimate reflects the combined effect of labor supply growth shifts with any related changes in worker-age composition.

At first glance, an aging of the workforce would tend to *increase* the startup rate. Azoulay, Jones, Kim and Miranda (2020) show that the propensity to start an employer business peaks between ages 35-44, and this group's share of the working-age population rose throughout the most rapid declines in the startup rate during the 1980s and 1990s. But there are other channels by which the age composition could affect the startup rate in general equilibrium: for example an effect could occur through rising consumer inertia favoring incumbents (Bornstein, 2018) or increased difficulty recruiting older workers (Engbom, 2017). In these examples, the share of young workers is complementary with the startup activity.

Importantly, our fertility and migration instruments are able to predict working-age population growth without disproportionately affecting the share of these young workers. The migration instrument generates growth contribution somewhat uniformly from each of the young, middle and older age groups. Even the lagged fertility IV does not systematically affect the share of young workers. This happens because past birth rates reflect both the size of the child-bearing-age population *and* their fertility behavior. While births 20 years ago contribute positively to the current share of young workers, their parents—who were in child-bearing age 20 years ago—contribute instead to the current share of middle and older-aged workers. These offsetting effects, leave the share of young workers little changed by the fertility IV. Appendix Table C.7 shows the IVs alone or together do not predict the share of young workers.

To examine the independent effects of the labor supply growth and age composition channels we supplement our main specification with the share of young workers. A challenge is that, like the WAP growth rate, the share of young workers in a state is also endogenous. To generate exogenous variation in the share of young workers, we supplement our set of IVs with an instrument proposed by Karahan and Rhee (2014), which extends our birthrate instrument by summing over the preceding 14 years.[32] In conjunction with the 20 year lag of the birthrate, the cumulative sum captures births between 21 and 34 years ago, to predict the contemporary age 20 to 34 share.

[32] Specifically, we define the IV as $IV_{st}^{BR15} = \sum_{k=1}^{14} BR_{t-20-k}$.

Using the expanded instrument set, Columns (1) and (2) of Table 7 show the first stage regressions for WAP growth and the 20-34 share, respectively. This new instrument is a strong predictor of youth share of working-age population (Column 2) even conditional on state and year fixed effects. Whereas, examining Column (1), the cumulative birthrate IV adds little predictive power for WAP growth conditional on our baseline IVs. Therefore, this instrument can isolate any additional effect of age composition in *excess* of effects from changes in the WAP growth rate.

Table 7: Effects of labor supply growth versus age composition

	(1)	(2)	(3)	(4)	(5)
	WAP Growth	20-34 Share	OLS	2SLS	2SLS
WAP Growth (%)			0.59	1.24	1.23
			(0.05)	(0.23)	(0.19)
20-34 share of WAP			0.17		0.22
			(0.04)		(0.10)
Birthrate IV	1.08	-1.38			
	(0.26)	(0.78)			
Migration IV	0.87	-0.37			
	(0.28)	(0.25)			
2-15 Year Birthrate Sum IV	0.01	0.53			
	(0.03)	(0.09)			
N	1,421	1,421	1,421	1,421	1,421
R^2	0.65	0.97	0.90	0.85	0.85
F-test	12.36	13.62		12.36	12.11
p-value of J test				0.05	0.57

Note: Standard errors are clustered on state. Regressions contain state and year fixed effects and cover years 1979-2007 and 48 contiguous states plus D.C.

We find the youth share of the working age indeed has an independent effect on the startup rate. In both OLS (3) and when instrumented (5), increasing the 20-34 share of the working-age population by one percentage point increases the startup rate by roughly 0.2 percentage points. This effect is consistent with the channels proposed by Bornstein (2018) and Engbom (2017). Importantly for our analysis, this effect is independent of our WAP growth elasticity. Our estimates with (Column 5) and without (Column 4) the youth share are very close to our baseline estimates in Column (7) of Table 6. Appendix C.2.4 shows our baseline estimate is also robust to controlling for other compositional shifts, e.g., in education and immigration.

4.3 Additional robustness

The statistical significance of the estimates and their quantitative stability across alternative sources of variation are a robust feature of the data. We consistently find an elasticity slightly larger than 1. We explore here the robustness of this result to several choices and report five main checks: an alternative IV, allowing for detailed industry controls, allowing for state-level trends, using the civilian labor force growth as the labor supply growth measure, and examining the elasticity of the establishment- rather than firm-startup rate.

Hybrid IV. Our first robustness exercise combines key features of the two instruments in a new "hybrid" IV. Specifically, we use each state s's *lagged* interstate migration patterns from all other states $k \neq s$, $\{\omega^k_{st^*}\}_{k\neq s}$ as in the migration IV and replace the "push" from other state contemporary labor supply growth, g_{kt} with those other states' *lagged* birthrates, b_{kt-20}. This follows equation (11), except now:

$$\hat{h}_{st} = \sum_{k\neq s} \omega^k_{st^*} b_{kt-20}. \tag{12}$$

This formulation relaxes the exclusion restrictions of the birthrate and migration IVs by requiring only that each state's *current* economic conditions be unrelated to *other* state's *lagged* birthrates.[33] For example, 1960 births in California should be unrelated to future economic conditions in New York in 1980. Instrumenting g_{st} using just the hybrid IV, $\hat{h}_{st}$, the first column of Table 8 shows the estimated elasticity of 1.23 is little changed from the baseline value of 1.19.

Industry controls. Demographic shocks could potentially affect the industry composition in a state, which might also have effects on the startup rate. For example, lower labor supply growth may shift economic activity away from high- towards low-entry industries. In contrast, our mechanism would predict an across the board decline in entry in response to a demographic shock, i.e., an effect within industries. To examine this possibility, we use detailed industry information from the confidential Census Bureau Longitudinal Business Database (LBD) and aggregate these data to form startup rates by year, state, and 4-digit NAICS industry.[34] Column (2) of Table

[33] This instrument's first stage is also very strong with an F-stat of 37.0 (Appendix Table C.5 Column 7). For further details and variations on this IV see Appendix C.2.1.

[34] We use the Fort and Klimek (2016) measure of NAICS industry (Appendix A.1.2).

Table 8: Robustness of effect of labor supply shocks on the startup rate.

	(1) Hybrid IV	(2) Detailed Industry	(3) State Trends	(4) CLF Growth	(5) Estab. Entry
WAP Growth (%)	1.23	1.06	1.32	1.21	1.05
	(0.26)	(0.20)	(0.19)	(0.28)	(0.20)
N	1,421	300,000	1,421	1,421	1,421
R^2	0.85	0.51	0.86	0.56	0.86
p-value of J test	—	0.25	0.10	0.70	0.38

Note: Standard errors are clustered on state. All regressions use benchmark specification from Column (7) of Table 6 with the following changes: (1) uses hybrid IV defined by (12) in place of fertility and migration IVs; (2) measures startup rate at state × NAICS4 digit industry level and adds NAICS4 fixed effects in LBD, exact N is suppressed; (3) adds linear state-level trends; (4) replaces WAP growth with CLF growth for g_{st}; (5) uses establishment entry rate in place of firm startup rate.

8 reports the estimated response of the startup rate to demographic shocks when including 4-digit industry dummies in addition to the state and year fixed effects. The estimate is very similar to the ones without detailed industry controls.

State-specific trends. The elasticity we estimate may be influenced by state-level secular trends in business dynamics unrelated to demographics affecting both the startup rate and the instruments. State fixed effects would control for the permanently low startup rates and fertility, but not for continued declines in both. However, allowing for state-specific linear trends in Column (3) our estimated elasticity of 1.32 is close to the baseline value of 1.19.

Different measure of labor supply. Our main proxy for labor supply, working-age population, counts people regardless of labor force status. While demographic change is one important factor in driving long-run changes in labor supply, another important margin is change in labor force participation. This distinction is important in the aggregate time series, but it should have little effect on cross-state results with state and year fixed effects. To evaluate the robustness of our results, we reproduce the same analysis using the growth rate of the civilian labor force. Column (4) shows this CLF-based estimate is nearly identical at 1.21. See Appendix C.2.2 for all main results using the CLF growth measure.

Establishment entry rate. Our measure of startups is based on new employer firms, which may be composed of one or more establishments. As we discuss in Section 2.1 the establishment entry rate has also declined (see also Appendix C.1.4). Column (5) re-estimates equation (10) using the state-level establishment entry rate and finds an elasticity of 1.05, just slightly below the 1.19 counterpart for firms.

Further robustness. For brevity, we provide further robustness of our cross-state results in Appendices C.2.1 (IVs), C.2.2 (CLF measure), C.2.3 (spatial correlation), and C.2.4 (labor supply composition). Collectively, these robustness analyses paint a consistent picture of a cross-state elasticity very close to our baseline.

5 Evaluating the mechanism in the time-series

As a final test of our hypothesis, we assess the main time-series implication of the mechanism. Namely, the startup rate should co-move with the labor supply growth rate, including the roughly two decades since 1960 marked by *increases* in both measures of labor supply growth. The difficulty with examining this implication directly is the absence of longitudinal firm-level data over this entire period.

To address this challenge, we provide a novel method to impute an aggregate establishment entry rate over a longer time span, which we take as a proxy the aggregate startup rate over the longer time span. Although we lack comprehensive longitudinal data prior to 1977, the Census Bureau has collected annual tabulations of establishments by county and size since 1964 as part of its County Business Patterns program.[35] These static data do not contain any measures of entry, but we propose a methodology to impute a historical aggregate entry rate series from these detailed cross sections. Essentially, we recover the number of new establishments or "gross entry" by summing the net change in the stock of establishments with the number of exiting establishments. Since the latter cannot be observed, we predict exits using a simple statistical model based on the establishment size distribution.

Imputing the establishment entry rate. Our imputation methodology begins with a simple law of motion for the total number of establishments e_t:

$$e_t = (1 - x_t)e_{t-1} + s_t,$$

[35] For further details on the CBP data, see Appendix A.4

where x_t is the average establishment exit rate and s_t is the number of new establishments. Given annual establishment counts and an average establishment exit rate, we could trivially recover the number of new establishments using the law of motion. While the establishment counts are readily available in the CBP, to our knowledge, there are no measures of U.S. establishment exit rates prior to the LBD.

Instead, we propose to predict exit using the annual state-level employment size distributions observable in the CBP. The aggregate exit rate can also be written as an establishment-share weighted average across states s and size groups j:

$$x_t = \sum_s \sum_j x_t^{sj} \frac{e_{t-1}^{sj}}{e_{t-1}}.$$

We condition on both size group and state, since both characteristics are also easily measured in the LBD and BDS data. If one had an estimate of each $\hat{x}_t^{sj}$ for the CBP, the aggregate establishment exit rate could be predicted by plugging in these estimates to the above expression. Then plugging in the predicted aggregate exit rate into the law of motion and normalizing by e_{t-1} to express in terms of rates, this procedure imputes the establishment entry rate as:

$$\frac{\hat{s}_t}{e_{t-1}} = \frac{\Delta e_t}{e_{t-1}} + \sum_s \sum_j \hat{x}_t^{sj} \frac{e_{t-1}^{sj}}{e_{t-1}}. \tag{13}$$

Summing the growth rate of the number of establishments with our predicted aggregate exit rate provides the imputed entry rate series.

To implement the imputation from equation (13) we use six size categories {1-19, 20-49, 50-99, 100-249, 250-499, 500+} and all 50 states plus the District of Columbia.[36] To estimate each $\hat{x}_t^{sj}$ to plug in, we estimate the following regression by size group and state in the BDS for the years 1980 to 2007 using a linear trend:

$$x_t^{sj} = \bar{x}^{sj} + \lambda^{sj} t + \varepsilon_t^{sj}. \tag{14}$$

We then predict $\hat{x}_t^{sj}$ for the CBP using the fitted values from 1980 to 2007 and holding exit rates constant at their predicted 1980 values, $\hat{x}_{1980}^{sj}$, for 1965 to 1979.[37] Appendix C.3 considers a number of alternatives to this simple procedure.

[36] Our choice of 500+ is not restrictive since average exit rates vary little among large firms.

[37] We drop years 1964, 1974, and 1983 because of significant data quality problems.

Historical startup rate. The imputed establishment entry rate confirms the main prediction of our mechanism. Figure 9, which plots the imputed establishment entry rate for 1965-2007 and the actual establishment entry rate for 1979 to 2007 measured from the BDS. We take this as a proxy for the historical startup rate since, in practice, most new establishments are also new firms, and the establishment entry rate is highly correlated with the firm startup rate (Appendix C.1.4).

The imputed establishment entry rate increases throughout the period of increasing labor supply growth. Although the annual data are noisy, a linear trend estimated from 1965 to 1979 shows a significant upward trend in the entry rate over the exact period of increases in our two measures of labor supply growth. This is exactly the time-series prediction of our mechanism where shifts in the growth rate of labor supply are accommodated primarily along the firm entry margin. Moreover, for the period in which they overlap, the imputed startup rate tracks on average the path of the BDS establishment startup rate reasonably well. Linear trends estimated over the overlapping 1979-2007 period for both series are very close.

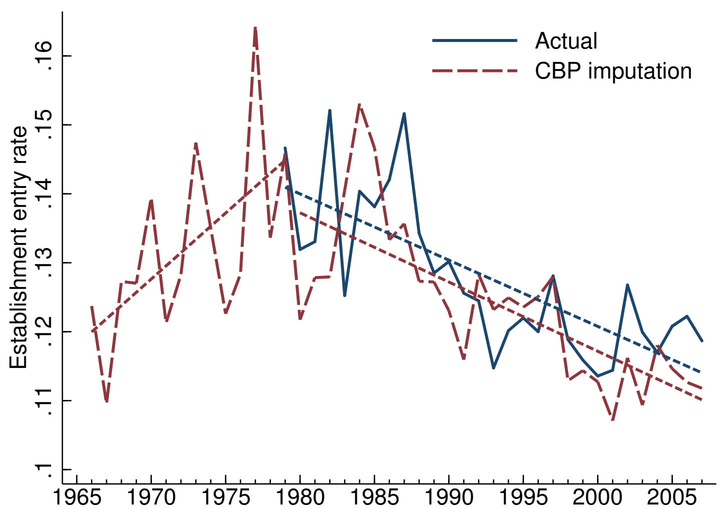


Figure 9: Imputed historical establishment entry rates and the BDS entry rate.

Note: Actual establishment entry rate measured in the BDS. CBP establishment startup rate imputed using equation (13) and interpolated across dropped years 1974 and 1983.

The hump shaped pattern in the imputed establishment entry rate is also robust to alternative assumptions on the path of establishment exit. Our baseline imputation holds the exit rate by size and state constant before 1979. The hump shape persists whether we use a constant average exit rate by size or even extrapolate the downward trend to the pre-1980 period (Appendix C.3). The trend controls for the gradual declines in exit rate within size-group from changes in age composition induced by changes the declines in entry. If anything, a higher establishment entry rate over the earlier period would increase average exit from the same age composition channel. If we were instead to extrapolate exit over the earlier period by reversing the sign on the trend term, the hump shaped pattern would be even more pronounced.

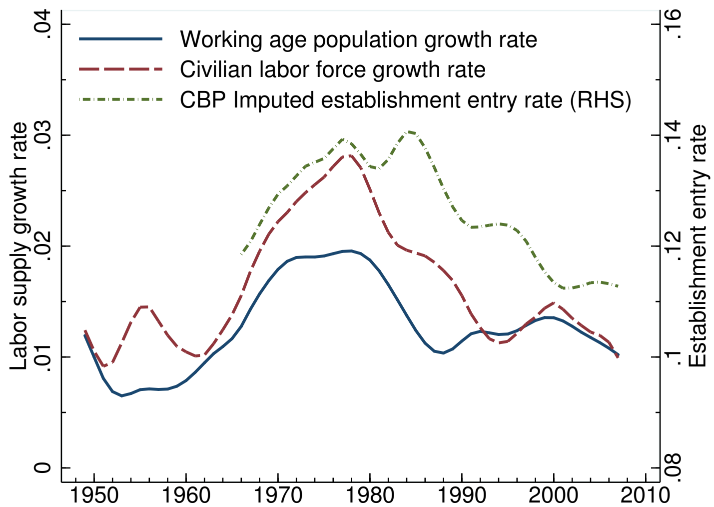


Figure 10: Measures of labor supply growth and historical entry rate

Note: Establishment entry rate is imputed from the CBP using equation (13). Trend components from HP filtered series with a smoothing parameter of 6.25, consistent with Ravn and Uhlig (2002) for annual data.

Collectively, the time series evidence strongly supports the labor supply growth mechanism. To see this most clearly, in Figure 10 we apply an HP filter to the annual data for both measures of labor supply growth and the imputed establishment startup rate and plot the trend components. Here, the comovement of the entry rate with the measures of labor supply growth is clear. The decline in entry is a reversal of its earlier increases, mirroring the hump-shaped patterns for labor supply growth.

6 Conclusion

The startup rate has been trending down since the series begins in the late 1970s, raising some alarm over the decline in U.S. entrepreneurship. In this paper, we identify declines in the growth rate of the labor supply–long preordained by demographics–as the leading cause the startup rate's decline, explaining between roughly 1/3 and 60 percent of the decline, depending on the measure of labor supply growth. In fact, the underlying equilibrium link between labor supply growth and the startup rate, although previously unexamined, is an inherent feature of standard models of firm dynamics. The explanation fits the data remarkably well, as we show in three independent ways: (i) quantitatively, examining the transitional dynamics of an equilibrium model of firm dynamics calibrated to match the U.S. economy, (ii) across U.S. states, using plausibly exogenous variation in labor supply growth, and (iii) in the aggregate time series, using a startup rate series extended to include the period of increasing labor supply growth. Moreover, this labor supply growth-based account of the startup deficit explains both the widespread nature of the declines in entry and the relative stability of incumbent exit and growth conditional on firm age.

The demographic origins of the decline in the startup rate should not diminish its macroeconomic significance. If anything, the decline in the startup rate and the implied shift in the age distribution of firms have far-reaching macroeconomic effects. An emerging literature has already started to explore these consequences. Glover and Short (2018) and Hopenhayn, Neira and Singhania (2018) link the demographic channel to the decline in the labor share; Pugsley and Şahin (2019) show that the emergence of jobless recoveries can be traced back to the decline in firm entry and subsequent aging of firms; Alon, Berger, Dent and Pugsley (2018), Atkeson, Burstein and Chatzikonstantinou (2018) and Peters and Walsh (2019) link the demographic driven declines in entry to the slowdown of aggregate productivity growth; and Crump, Eusepi, Giannoni and Şahin (2019) argue that the demographic channel is an important driver of the secular downward trend in unemployment. If present trends in fertility and immigration continue, the startup rate will remain near its current levels, and understanding the equilibrium implications of slowing labor force growth will be especially relevant.

References

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

Alon, Titan, David Berger, Robert C Dent, and Benjamin Pugsley (2018) "Older and Slower:

The Startup Deficit's Lasting Impact on Productivity Growth," *Journal of Monetary Economics*, 93, 68–85.

Altonji, Joseph G. and David Card (1991) *The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives*, 201–234: University of Chicago Press, http://www.nber.org/books/abow91-1.

Atkeson, Andrew, Ariel Burstein, and Manolis Chatzikonstantinou (2018) "Transitional Dynamics in Aggregate Models of Innovative Investment,"Technical report, National Bureau of Economic Research.

Azoulay, Pierre, Benjamin F Jones, J Daniel Kim, and Javier Miranda (2020) "Age and high-growth entrepreneurship," *American Economic Review: Insights*, 2 (1), 65–82.

Bartelsman, Eric J. and Mark Doms (2000) "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 38 (3), 569–594.

Bollard, Albert, Peter J Klenow, and Huiyu Li (2016) "Entry costs rise with development," *Unpublished manuscript, Stanford University*.

Bornstein, Gideon (2018) "Entry and Profits in an Aging Economy: The Role of Consumer Inertia," mimeo.

Buckles, Kasey, Daniel Hungerman, and Steven Lugauer (2018) "Is Fertility a Leading Economic Indicator,"Technical report, National Bureau of Economic Research.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico (2018) "Monetary policy, corporate finance and investment,"Technical report, National Bureau of Economic Research.

Crump, Richard, Stefano Eusepi, Marc Giannoni, and Ayşegül Şahin (2019) "A Unified Approach to Measuring u*," *Brookings Papers on Economic Activity*.

Davis, Steven and John Haltiwanger (2016) "Dynamism Diminished: The Role of Housing Markets and Credit Conditions," mimeo.

Davis, Steven J and John Haltiwanger (2014) "Labor Market Fluidity and Economic Performance,"Technical report, National Bureau of Economic Research.

Davis, Steven J, John Haltiwanger, Ron Jarmin, and Javier Miranda (2007) "Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms," *NBER Macroeconomics Annual 2006*, 21, 107–156.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda (2020) "Changing business dynamism and productivity: Shocks versus responsiveness," *American Economic Review*, 110 (12), 3952–90.

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

Elsby, Michael W.L., Ryan Michaels, and David Ratner (2017) "The Aggregate Effects of Labor Market Frictions," mimeo.

Engbom, Niklas (2017) "Firm and Worker Dynamics in an Aging Labor Market," mimeo, Princeton University.

——— (2019) "Firm and Worker Dynamics in an Aging Labor Market," mimeo.

Fort, Teresa C. and Shawn D. Klimek (2016) "The Effects of Industry Classification Changes on US Employment Composition," March, mimeo, Dartmouth University.

Glover, Andrew and Jacob Short (2018) "Demographic Origins of the Decline in Labor's Share," mimeo.

Gourio, Francois, Todd Messer, and Michael Siemer (2014) "A Missing Generation of Firms? Aggregate Effects of the Decline in New Business Formation," July, mimeo.

Gutiérrez, Germán, Callum Jones, and Thomas Philippon (2019) "Entry costs and the macroeconomy,"Technical report, IMF Working Paper 19-233.

Haltiwanger, John, Ron Jarmin, and Javier Miranda (2011) "Historically, Large Decline in Job Creation from Startup and Existing Firms in the 2008-2009 Recession," Business Dynamics Statistics Briefing 5, Kauffman Foundation.

——— (2012) "Where Have All the Young Firms Gone?" Business Dynamics Statistics Briefing 6, Kauffman Foundation.

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

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

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

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

Hopenhayn, Hugo, Julian Neira, and Rish Singhania (2018) "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share,"Technical report, National Bureau of Economic Research.

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

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

Jiang, Helu and Faisal Sohail (2017) "Skill biased entrepreneurial decline,"Technical report, Working paper, September 2017. 5.

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

Kennan, John and James R. Walker (2011) "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79 (1), 211–251, 10.3982/ECTA4657.

Klenow, Peter J and Huiyu Li (2020) "Innovative growth accounting," in *NBER Macroeconomics Annual 2020, volume 35*: University of Chicago Press.

Kozeniauskas, Nicholas (2017) "What's Driving the Decline in Entrepreneurship," mimeo, New York University.

Lentz, Rasmus and Dale T. Mortensen (2008) "An Empirical Model of Growth Through Product Innovation," *Econometrica*, 76 (6), 1317–1373, 10.3982/ECTA5997.

Liang, James, Hui Wang, and Edward P. Lazear (2018) "Demographics and Entrepreneurship," *Journal of Political Economy*, 126 (S1), S140–S196.

Luttmer, Erzo GJ (2010) "Models of growth and firm heterogeneity," *Annu. Rev. Econ.*, 2 (1), 547–576.

Nakamura, Emi and Jón Steinsson (2018) "Identification in macroeconomics," *Journal of Economic Perspectives*, 32 (3), 59–86.

Peters, Michael and Connor Walsh (2019) "Declining Dynamism, Increasing Markups and Missing Growth: The Role of The Labor Force," mimeo.

Pugsley, Benjamin Wild and Ayşegül Şahin (2019) "Grown-up Business Cycles," *The Review* *of Financial Studies*, 32 (3), 1102–1147.

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

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobeck (2017) "Integrated Public Use Microdata Series: Version 7.0 [dataset]," https://doi.org/10.18128/DO10.V7.0, Minneapolis: University of Minnesota.

Salgado, Sergio (2017) "Technical Change and Entrepreneurship," mimeo, University of Minnesota.

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

Sterk, Vincent ⓡ Petr Sedláček ⓡ Benjamin Pugsley (2021) "The nature of firm growth," *American Economic Review*, 111 (2), 547–79.