Appendix to Demographic Origins of the Startup Deficit

Benjamin Pugsley   Fatih Karahan   Ayşegül Şahin

March 2021

FOR ONLINE PUBLICATION

Contents

A Data Appendix

Our analysis uses data on firms and labor markets, both national and local, from several sources. We provide additional details on these sources and the methodology to replicate the estimates in the paper. Throughout the paper, we use data on firms from the restricted-access U.S. Census Bureau Longitudinal Business Database (LBD) and its public-use tabulations, the Business Dynamics Statistics (BDS).[1] We combine these data with measures of labor supply growth from the Census Bureau and Bureau of Labor Statistics. Finally we use tabulations from Census Bureau's County Business Patterns (CBP) for an independent historical estimate of an establishment startup rate.

A.1 Measuring firm dynamics in the LBD and BDS

A.1.1 Firm-level measures

We measure firm-level employment (our measure of firm size), which we then aggregate by state, firm-age, group, and 4-digit NAICS industry. The LBD consists of annual establishment-level data, which are linked longitudinally at the physical establishment level. These linkages span changes in ownership or other reorganizations. For each establishment, the dataset contains employment reported for the week containing March 12 of each calendar year. Since employment is measured at the EIN level via payroll taxes, in the case of multi-unit firms, establishment-level employment is sometimes imputed across establishments within an EIN. Since we will aggregate all employment to the firm level, this imputation has no affect on our measure of firm-level employment. See Jarmin and Miranda (2002) for additional details on the construction and limitations of the LBD. The choice of LBD by construction limits the universe to firms with employees and thus omits any self-employment activity where the proprietor does not have any employees.[2]

To measure firm-level employment (our measure of firm size) and exit, we aggregate employment across all establishments within a firm.[3] 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 (or would belong) to an age group in a particular year. An age group's year t employment growth is calculated by first aggregating employment across all firms, E_t^a, currently within the age group cohort $a \in \{y, m\}$ and then calculating the growth relative to the total employment of this cohort in the previous year, E_{t-1}^a, where previous year employment is measured for all firms which, if operating, would be in group a in year t including those who exit.[4] We measure firm exit in year t when all of a firm's year $t-1$

[1]Specially Sworn researchers with an approved project may request the replication files from CES Project 1731 if the corresponding years of the LBD and SSL are within the project scope.

[2]The "nonemployer" universe, that includes this activity is much larger but counts many casual businesses or independent contractors, e.g., Uber drivers. For a discussion of trends in employer and nonemployer businesses see Pugsley and Şahin (2019) and references therein.

[3]The Census Bureau defines a firm as the highest level of operational control over establishments and this is ascertained during the quinquennial Economic Census or the Annual Company Organization Survey.

[4]The change, $E_t^a - E_{t-1}^a$, between current and previous year employment for an age group cohort a corresponds

establishments have 0 employment and are reported closed in year t. This measure of exit would not count exit through mergers or other reorganizations since establishments at these firms would still report activity in year t.

We identify startups and distinguish incumbents based on a measure of firm age. To be consistent with the BDS and the prior literature, we calculate firm age as the age of the firm's oldest establishment measured by the first year of positive employment for the March 12 pay period.[5] An establishment "enters" in the year it first reports employment and ages naturally thereafter (regardless of any ownership changes). Startups are age 0 firms and they are bona fide new firms, since they are composed entirely of age 0 establishments. The startup rate measures the number of startups as a fraction of the total number of employer firms. Incumbents are firms which are age 1 and higher, and we further split these into young (age 1-10) and mature (age 11+) age groups. Our measures of incumbent dynamics by age group start in 1987.[6] Using the LBD or BDS the startup rate can be computed as early as 1977, but it would include true entrants and those firms which may have existed (even with employees) but who did not record any payroll in 1976. Starting in 1979 ensures we look back at least 3 years for any payroll activity before labeling it a startup. Our measure of young firms in 1987-1988 may include some firms mistakenly classified as entrants in 1977 and 1978, but this effect should be very small. We then tabulate these firm-level measures by age group at the national, state, and state by 4-digit NAICS levels.

A.1.2 Geography

In the case when a firm operates multiple establishments we assign its location (state) as the state with the greatest employment share. For state-level tabulations, following the BDS, our state-level measures count firms separately for every state in which they operate establishments (implying the sum of state firm counts may exceed the total number of U.S. firms). A firm operating in two states will be counted twice, summing only across the respective establishments within each state. Only employment will vary across firm-states for the same firm. Other firm characteristics, such as total size, and industry will be identical.

A.1.3 Industry assignment

We assign firm-level measures of detailed industry using the NAICS 2002 industry classifications. There are two challenges to constructing firm-level measures of industry. The first is that industry classification naturally evolves over time and we need to construct a longitudinally consistent coding

to the BDS measure of the age group's net job creation. See the discussion below in A.1.4.

[5] This measure was first popularized by Davis, Haltiwanger, Jarmin, and Miranda (2007) and Haltiwanger, Jarmin, and Miranda (2013). The LBD also includes firms that recorded payroll tax during the calendar year (IRS Form 941) but did not have employment for the pay period including March 12, e.g., seasonal businesses. One could also define a measure of firm age based purely on first annual payroll activity. This measure would also include firms who "enter" after Q1 but exit before the following Q1, which can occur but is not changing systematically over time.

[6] Because birth year is left censored for any extant firms in the first year of the LBD (1976), we cannot measure the young and mature age groups until 1987 since this is the first year we can identify age 1-10 firms and thus the residual 11+ age group.

of detailed industry. The second is that industry is assigned at the establishment-, rather than firm-level. For example, the headquarters location of a large manufacturing firm may be classified within management of professional enterprises, while the plants may be classified within the manufacturing sector.

To address the periodic reorganization of industry codes over time, we assign a longitudinally consistent measure of industry developed in Fort and Klimek (2016). They use a concordance of SIC to 2002 NAICS coding to "backcast" NAICS codes at the establishment level in years in which only a 6 digit SIC code was assigned. This is straightforward for industries where there is a one to one mapping, however there are many SIC industries that map across multiple NAICS industries and vice versa. In these cases, they assign industry stochastically drawing a NAICS code from the empirical distribution of NAICS codes that map to a specific SIC code (for years in which the standard industry assignment overlapped). They also make some ad-hoc corrections, which are described in the appendix to Fort and Klimek (2016).

Having assigned a Fort-Klimek NAICS code to every establishment year, we then in cases of multi-unit firms assign a firm-level NAICS code. We assign this using an activity (payroll) weighted mode across establishments, but follow a hierarchical procedure to ensure that the NAICS code would be assigned consistently at each level of aggregation. That is, we first assign a 2-digit NAICS code as the modal 2-digit code across establishments within a firm (excluding any management of professional enterprises coding). Then within the firm's 2-digit NAICS industry, we assign a modal 3-digit NAICS industry across those establishments, and so on. This method ultimately assigns a NAICS-2002 6-digit industry to every firmid within every year.

A.1.4 Consistent measures in Business Dynamics Statistics (BDS)

Finally, in order to ensure our main results can be easily replicated, we use the BDS tabulations where possible. This requires some small adjustments in order to ensure consistent measures between LBD and BDS. The BDS report the stock of firms and their employment in each year, but because firms may go temporarily out of scope, measures such as within cohort exit and employment growth cannot be reliably measured only from the change in stocks. Moreover, for multi-age groups the previous year stock is not reported. We follow the procedure from Pugsley and Şahin (2019) to recover lagged cohort employment and number of firms in order to produce more accurate measures of exit and employment growth. We summarize that procedure here.

Stock measures Let EMP_t^a measures within a firm age group a cell the total stock of March 12 employment across all establishments (within the firm age group) in year t, and let $DENOM_t^a$ measure the average of EMP_t^a and the total stock of employment for that same cohort of firms in the previous year, $\widetilde{EMP}_{t-1}^{a-1}$.[7] This imputed previous year employment is computed from the BDS

[7]For some age groups the previous year's employment may not be directly observable in the BDS. For example, the 6 to 10 age group in year t cannot be observed directly in year $t-1$.

variables $DENOM_t^a$ and EMP_t^a as

$$\widetilde{EMP}_{t-1}^{a-1} = 2 \times DENOM_t^a - EMP_t^a .$$

For firms, let the BDS variable $FIRMS_t^a$ measure within an age group a the total number of firms with positive employment on March 12 of that year, and let variable $DEATHS_t^a$ measure the number of firms in the current age group a cohort that were active in $t-1$, but are now permanently shut down in year t. A shut down requires that all establishments within the firm in the previous year exit by the current year. Then we construct

$$\widetilde{FIRMS}_{t-1}^{a-1} = FIRMS_t^a + DEATHS_t^a \ .$$

We define year t age group a **number of firms** F_t^a and its lagged value F_{t-1}^{a-1} for the same cohort as

$$F_t^a \equiv FIRMS_t^a \quad F_{t-1}^{a-1} \equiv \widetilde{FIRMS}_{t-1}^{a-1} \ .$$

Next, we define **average employment size** and its lagged value for the same cohort as

$$N_t^a \equiv \frac{EMP_t^a}{FIRMS_t^a} \quad N_{t-1}^{a-1} \equiv \frac{\widetilde{EMP}_{t-1}^{a-1}}{\widetilde{FIRMS}_{t-1}^{a-1}} \ .$$

Flow Variables Using our above defintions for E_t, F_t, and N_t, we compute the dynamic measure defined in the paper. The **exit rate** is

$$x_t^a \equiv \frac{DEATHS_t^a}{\widetilde{FIRMS}_{t-1}^{a-1}} = 1 - \frac{F_t^a}{F_{t-1}^{a-1}} \ .$$

Note that this is a restrictive definition of exit, since firms that are reorganized are not counted as exits. The growth rate in average size or **conditional growth rate** is [8]

$$n_t^a \equiv \frac{N_t^a - N_{t-1}^{a-1}}{N_{t-1}^{a-1}} \ .$$

The age group **unconditional employment growth rate** is

$$g_t^a \equiv \frac{E_t^a - E_{t-1}^{a-1}}{E_{t-1}^{a-1}} \ .$$

[8]The conditional growth rate will only equal the growth rate "conditional on survival" when the average size of exiting and surviving firms is identical. Since exiting firms are typically smaller than surviving firms, the conditional growth rate measured in this way would be greater than the growth rate of surviving firms. In general, the growth rate of surviving firms is given by $1+g_t^{aS} = (1+n_t^a)(1-x_t^a(1-N_{t-1}^{aS}/N_{t-1}^{aX}))$, where $N_{t-1}^{aS}/N_{t-1}^{aX}$ is the ratio of exiting firm to surviving firm size.

Then also by construction

$$1 + g_t^a = (1 - x_t^a)(1 + n_t^a) \ .$$

Unconditional Growth Rate and Net Job Creation Rate The definition of g_t^a will differ slightly from the age group a net job creation rate $NJCR_t^a$ from the BDS where

$$1 + NJCR_t^a = 1 + \frac{JC_t^a - JD_t^a}{\frac{1}{2}\left(EMP_t^a + \widetilde{EMP}_{t-1}^{a-1}\right)} \ .$$

The growth rate differs both because of the denominator and because (until the September 2014 release) $JC_t^a - JD_t^a \neq EMP_t^a - \widetilde{EMP}_{t-1}^{a-1}$.[9]

A.2 Measuring labor supply growth rates

Our demographic measures include national- and state-level measures of the working age population and the civilian labor force. We use population data from the Census Bureau's decennial census and annual American Community Survey. We use the Current Population Survey (CPS) to measure the size of the civilian labor force. We define the working age population as the non-institutional population between the ages of 20 and 65 and the civilian labor force as the non-institutional population age 20 or older that are currently employed or actively searching for a job.

A.2.1 Working age population-based estimate

We construct the growth rate of the working-age population using annual Census Bureau intercensal population estimates by age group. These annual data are based on the decennial population census and intercensal estimates formed using data on births, deaths, and migration.[10] We sum the annual estimates by age group to estimate the population ages 20-64 and then take the one year growth rates. This is a benchmark measure of the growth rate of the working age population. This range 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. We have experimented with both narrower and wider definitions of the working-age with little effect on the aggregate patterns or cross-state results.

A.2.2 Civilian labor force-based estimate

At the national level the CLF is estimated monthly by the BLS using the Current Population Survey (CPS), and for states the labor force is estimated as part of the Local Area Unemployment Statistics program, which combines the CPS with information from state-level unemployment insurance programs, the BLS establishment survey, as well as local population estimates from the

[9] Starting in the September 2014 release of the BDS $JC_t^a - JD_t^a = EMP_t^a - \widetilde{EMP}_{t-1}^{a-1}$ nearly exactly.

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

Census Bureau. We average the monthly estimates of the CLF by year, and then take one year growth rates. Since, even at an annual frequency the CLF is procyclical, see for example Elsby, Hobijn, and Şahin (2015), we also use a version of the CLF growth rates purged of business cycle fluctuations using an HP filter with a smoothing parameter of 6.25 as recommended by Ravn and Uhlig (2002).

A.3 Cross-state sample construction

A.3.1 Sample description

We measure the startup rate, average startup size, exit rate for firms ages 1-10, and conditional growth rate (growth in average firm size) for firms ages 1-10 for each state and year in the BDS and LBD using the procedure described above in Sections A.1. We do this for years 1979 to 2007 for the startup rate and average startup size variables, and for years 1987 to 2007 for the exit rate and conditional growth rate variables defined for young incumbents. Because birth year is left censored in 1977, the year 1987 is the first year where we can identify all firms ages 1-10. We restrict the sample to states in the contiguous US plus the District of Columbia since Alaska and Hawaii were granted statehood in 1959 and consistent natality and population data are not available before 1960.

To these data defined by state and year, we merge the state-level counterparts for working age population growth and the civilian labor force described above in Section A.2. Then we add the birthrate and migration instruments, which are described in the next section. For the 48 contiguous US states plus DC, Table A.1 reports the sample statistics for the full 1979-2007 (1,421 state-year observations) and shorter 1987-2007 (1,029 state-year observations) samples before and after removing state and year fixed effects from each variable.

A.3.2 Instrumental variables construction

Measuring state birthrates from Department of Health data We tabulate historical state births from the Natality Data from the National Vital Statistics System of the National Center for Health Statistics. These public-use microdata are available for download from the CDC (`https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm`) for 1968 through the present. For each state and year we measure the number of births per 1000 adults (measured from the Decennial Census and inter-censal estimates), which is known as a "crude birth rate. We are grateful to Rob Shimer for providing us with his birthrate data constructed from the Statistical Abstracts for the period 1940–91. Data are unavailable for Hawaii and Alaska prior to 1960, and we drop these states entirely from the analysis.

We then use the state-level birthrates lagged 20 years to predict future labor supply growth, measured both by the working age population and the civilian labor force, conditional on state and time fixed effects. Even after removing these fixed effects, there is considerable cross state variation that remains. Table A.1 reports the descriptive statistics for the annual data. The standard

Table A.1: Cross-state sample statistics

	Actual values					Residualized			
	mean	sd	p10	p50	p90	sd	p10	p50	p90
Panel A. 1979 to 2007									
Startup rate	10.75	2.09	8.36	10.44	13.62	0.80	−0.87	−0.01	0.78
Startup size	5.96	1.05	4.97	5.84	7.02	0.78	−0.60	−0.06	0.54
WAP growth rate	1.27	1.17	0.16	1.09	2.68	0.72	−0.76	0.00	0.76
CLF growth rate	1.38	1.54	−0.34	1.26	3.23	1.19	−1.34	0.01	1.34
Birthrate (20 yr lag)	18.00	3.69	14.10	17.20	23.60	1.10	−1.25	0.00	1.22
N	1,421								
Panel B. 1987 to 2007									
Startup rate	10.16	1.73	8.11	9.91	12.56	0.59	−0.61	0.01	0.59
Startup size	6.00	1.03	4.95	5.90	7.12	0.71	−0.52	−0.05	0.51
Firm exit 1-10	10.93	1.23	9.48	10.87	12.45	0.67	−0.65	−0.05	0.66
Cdtl. growth 1-10	8.82	3.57	4.92	9.01	12.31	2.78	−2.78	0.01	2.61
WAP growth rate	1.18	1.09	0.17	1.01	2.45	0.64	−0.66	0.00	0.67
CLF growth rate	1.22	1.45	−0.39	1.17	2.89	1.14	−1.31	0.03	1.30
Birthrate (20 yr lag)	16.34	2.40	13.80	16.00	19.10	0.98	−1.11	−0.04	1.17
N	1,029								

Note: Sample used for estimation. Residualized columns report the statistics for each variable after removing state and year fixed effects.

deviation of birth rate across states and years is approximately 3.7 (births per 1000 adults) and falls by roughly 2/3 to 1.1 after conditioning on state and year fixed effects.

Measuring inter-state migration using Census data The long form of the Decennial Census (until replaced by the annual American Community Survey in 2001) asks respondents for the place of birth (U.S. state or country) of each person in the household. We use the 5% microdata samples for 1990 and 1980 Decennial Censuses. In 1970, we use the 1% Form 1 metro sample. These public-use samples are available from IPUMS (Ruggles, Genadek, Goeken, Grover, and Sobeck, 2017). We then aggregate over all native-born persons in that state to estimate the distribution of birth states. For each state k we then condition on all birth states that are not part of the same Census division to form the distribution of intra-division birth states.[11] When constructing the instrument we use the lagged distribution of birth states from at least 1 Census ago, so that there is a minimum of 10 years between current year and the year in which birth states are measured. The questionnaire also reports the state of residence 5 years ago, which can also be used to construct weights. We find similar estimates using these weights to construct the migration instrument, but the first stage regression is weaker.

[11]There are 9 Census divisions: New England, Mid-Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain and Pacific.

A.4 Census County Business Patterns (CBP) and startup rate imputation

We use data from the County Business Patterns (CBP) and Business Dynamics Statistics (BDS) jointly to impute the establishment entry rate for the pre-1979 period. Below is a detailed description of the CBP dataset. See also `http://www.census.gov/programs-surveys/cbp.html`.

Historical CBP data. The Census Bureau's County Business Patterns (CBP) program counts the number of establishments in each U.S. county. It has published tabulations of establishments by geographic area and employment size class annually since 1964.[12] These data are publicly available and downloadable from the Census Bureau (`https://www.census.gov/programs-surveys/cbp/data/datasets.html`). Years prior to 1986 are available in the National archives as well as digitized versions in ICPSR.[13] Because of slight changes in the CBP over time size categories depend on year:

Years 1964 to 1973 We use state level data binned by year and size category: Number of Employees: 1-3, 4-7, 8-19, 20-49, 50-99, 100-249, 250-499, 500+. For each cell, we measure the establishment count and employment.

Years 1974 to 2014 We use state level data binned by year and slightly different size categories: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+ employees. For each cell, we again measure the establishment count and employment.

These size categories are chosen to correspond as closely as possible to size categories in the publicly available BDS.

Baseline imputation method.

1. Iterating over establishment size categories and state geographies, estimate equation 14. Then take predicted values for years 1979-2007 using CBP data from that sample period.
2. For years 1966-1978, assume that for each size category and each state geography, the exit rate was constant at the (fitted) level it was in 1979.[14]
3. Using equation (13) and the predicted annual exit rate, $\hat{x}_t^{sj}$, for each state and size group, compute the establishment startup rate. The aggregate change in establishments Δe_t can be measured directly in the CBP by aggregating the changes by state and size category. Note that size transition flows need not be estimated in order to estimate the aggregate startup rate, as summing across all size categories nets out inflows and outflows.
4. Imputed entry rates for years 1974 and 1983 are dropped because of changes in methodology for tabulating establishments in the CBP. Imputed entry rates for those years are replaced as the midpoints of data from years 1973,1975 and 1982,1984, respectively.

[12] The CBP program provides data as early as 1946 at roughly triennial frequencies. In these early years, multi-unit establishments are often be combined within county and detailed industry. See `https://www.census.gov/programs-surveys/cbp/technical-documentation/methodology.html`

[13] See `https://research.archives.gov/id/613576` and `http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/25984`

[14] We consider alternative assumptions in Appendix C.3.

B Model appendix

This section provides further details on the model and its extensions as well as additional model results deferred from the main text because of space constraints.

B.1 Results with establishment calibration

In the paper we provided results from a calibration that targeted stationary moments of the firm distribution in the U.S. For a discussion of why firms are the preferred unit of analysis see Section C.1.4. We now use an alternative calibration based on establishment data and show that the substantive conclusions of our analysis still hold.

B.1.1 Establishment calibration

The establishment calibration is a close analog of the main calibration. Specifically, we keep the parameters set outside the model the same. That is, we keep the time discount rate β fixed at 0.96 and the curvature parameter of the production function θ at 0.64. The remaining eight parameters are calibrated by targeting the same 23 moments in Section 3.2, this time computed for U.S. establishments, by minimizing the weighted squared distance between the data and model counterparts. Table B.1 shows the calibrated parameter values, and Figure B.1 shows the fit of the model on the targeted moments. Like the fit of the main calibration for firms in Figure 5, the model does a very good job in fitting 23 moments for establishments with eight parameters.

Table B.1: Values for internally calibrated parameters

c_e	c_f	δ	ρ	σ_ε	μ_0	σ_0	σ_a
6.189	2.472	0.018	0.962	0.035	-0.056	0.146	0.041

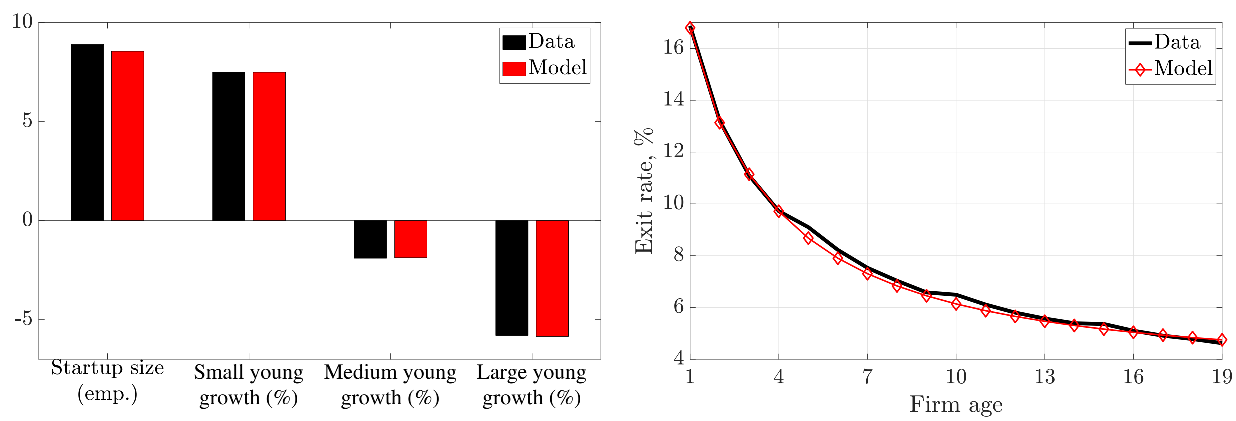


Figure B.1: Establishment calibration: Model fit to targeted moments

B.1.2 Comparative statics wrt η

Figure B.2 is the establishment counterpart to main text Figure 6. The effects of η are very similar when the model is calibrated to match establishment dynamics.

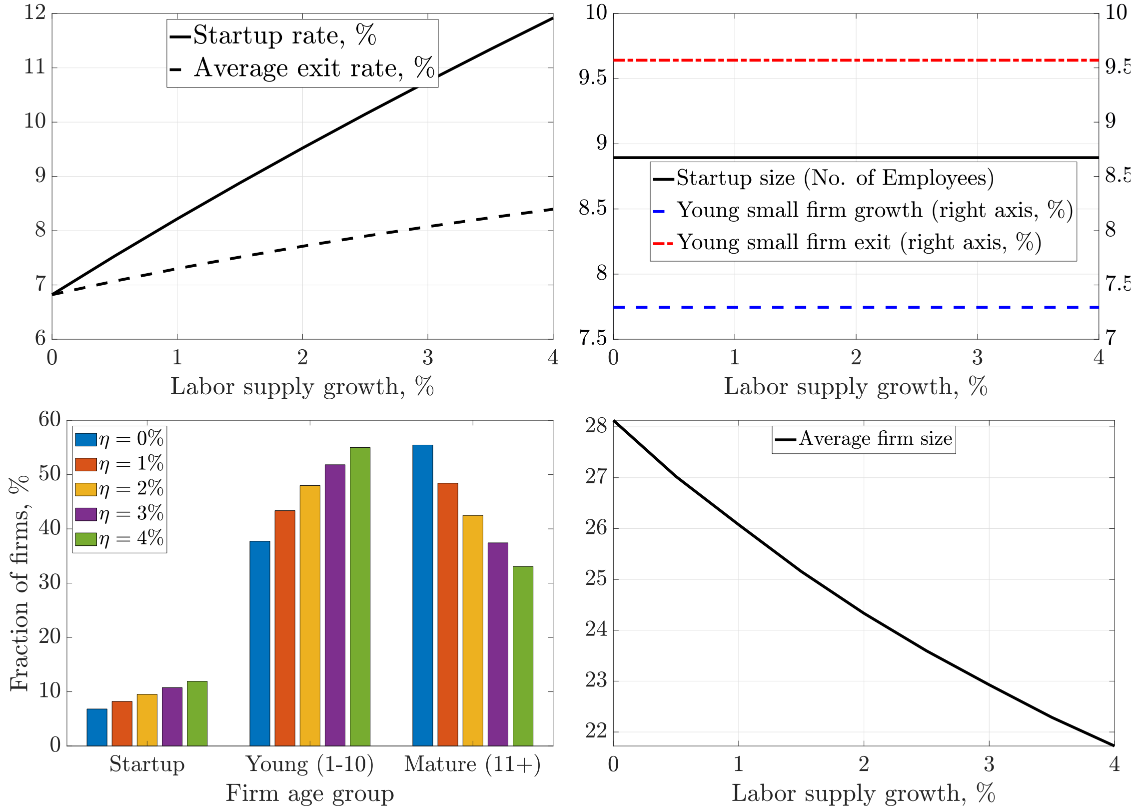


Figure B.2: Labor supply growth and firm dynamics

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

B.1.3 Effects of Demographics over 1979-2007 period

Table B.2 is the establishment counterpart to main text Table 5. The observed path of WAP (CLF) growth for η_t explains 45% (82%) of the 2.2 p.p. decline in the establishment entry rate.

Table B.2: Actual and predicted declines in the establishment entry rate

	Labor Supply Growth (%)		Establishment Entry Rate (%)			Establishment Economy Exit Rate (%)		
			Actual	Model		Actual	Model	
	WAP	CLF		WAP	CLF		WAP	CLF
1979-1981	1.9	2.5	12.6	9.2	9.9	9.4	7.8	8.0
2005-2007	1.1	1.1	10.3	8.1	8.1	8.2	7.6	7.6
Change	−0.8	−1.4	−2.2	−1.0	−1.8	−1.2	−0.2	−0.4

B.2 Model with transitional dynamics

In the paper we characterize the balanced growth path of the economy and how its entry rate varies in response to permanent shifts in labor supply (population) growth rate. We now consider the transitional dynamics of an economy on a balanced growth path in period $t = 0$ that receives a perfect foresight shock at time $t = 1$ of a new future path of the labor supply growth rate, η_t.

Given the now time-varying path of η_t, preferences over per capita consumption paths $\{\bar{c}_t\}_{t=1}^{\infty}$ are ordered according to:

$$U = \sum_{t=1}^{\infty} \beta^{t-1} \prod_{k=1}^{t} (1 + \eta_k) \log \bar{c}_t \tag{B.1}$$

The problem of the household is now subject to a time-varying interest rate, r_t, real wage, w_t, and per capita dividend, π_t. In per capita terms, the budget constraint for $t \geq 1$ is:

$$\bar{c}_t + (1 + \eta_{t+1})\bar{b}_{t+1} = \bar{\pi}_t + (1 + r_t)\,\bar{b}_t + w_t. \tag{B.2}$$

Given the time-varying wages and interest rates the value of a firm is no longer stationary. Let $V_t\,(a, s_t)$ be the value of a firm in period t with permanent productivity a and stochastic productivity s_t. For $t \geq 1$, V_t satisfies the following sequence of Bellman equations:

$$V_t\,(a, s_t) = \max_{n_t} \left\{ a s_t n_t^{\theta} - c_f - w_t n_t + \frac{1}{1 + r_t} \max_{X_t \in \{0,1\}} \{E_t V_{t+1}\,(a, s_{t+1})\,, 0\} \right\}. \tag{B.3}$$

The t subscript on V_t captures its dependence on the path of wages and real interest rates $\{w_t, r_t\}$. We let $h_t\,(a, s_t)$ be the policy function for labor demand n_t, and $X_t\,(a, s_t)$ be the optimal exit decision that satisfy the sequence of Bellman equations.

Let $\bar{\mu}_t\,(A, S)$ be the measure of firms per-capita with productivity $(a, s) \in A \times S$. Then the sequence $\{\bar{\mu}_t\}_{t=0}$ satisfies

$$\bar{\mu}_{t+1}(A, S_{t+1}) = \iint_{(a,s_t) \in A \times \mathbb{R}_{>0}} (1 - \delta)(1 - X(a, s))P(S_{t+1}|s_t)\frac{\bar{\mu}_t(da, ds_t)}{1 + \eta_{t+1}} + \bar{M}_{t+1}G(S_{t+1})F(A), \tag{B.4}$$

Athough the value of firm is now time varying, the path of equilibrium prices must still ensure for all $t \geq 1$ that the expected value of a potential entrant remains equal to the constant entry costs, c_e, to satisfy the free entry condition:

$$c_e = \iint V_t(a, s_t)G(s_t)F(da). \tag{B.5}$$

Finally, market clearing requires profits per-capita satisfy:

$$\bar{\pi}_t = \iint \Big(a s_t h_t\,(a, s_t)^{\theta} - c_f - w_t h_t\,(a, s_t) \Big)\, \mu_t\,(da, ds_t) - \bar{M}_t c_e. \tag{B.6}$$

Demographic shock and the equilibrium response. Let V_0 be the (stationary) value of a firm in an initial balanced growth path equilibrium with constant labor supply growth η_0. At the beginning of period $t = 1$ there is an unanticipated shift in the path $\{\eta_t\}_{t=1}^{\infty}$ of labor supply growth: it varies from $t = 1 \ldots, T$, and remains constant thereafter ($\eta_t = \eta_T$ for $t \geq T$).[15] The transition path to the new balanced growth path with constant $\eta = \eta_T$ consists of a sequence of prices $\{w_t, r_t\}_{t=1}^{\infty}$, per-capita profits $\{\bar{\pi}_t\}_{t=1}^{\infty}$, per-capita consumption $\{\bar{c}_t\}_{t=1}^{\infty}$, value functions and policy rules $\{V_t, X_t, h_t\}_{t=1}^{\infty}$, measures of firms $\{\bar{\mu}_t\}_{t=1}^{\infty}$ and entrants per capita, $\left\{\bar{M}_t\right\}_{t=1}^{\infty}$ that satisfy the following conditions:

1. Given prices and profits, $\{w_t, r_t, \pi_t\}_{t=1}^{\infty}$, $\{c_t\}_{t=1}^{\infty}$ maximizes household utility (B.1) subject to a sequence of budget constraints (B.2).
2. Value functions and policy rules satisfy (B.3).
3. Measures and entrants per capita $\left\{\bar{\mu}_t, \bar{M}_t\right\}_{t=1}^{\infty}$ satisfy the law of motion (B.4).
4. Free entry condition (B.5) holds with equality in each period.
5. Per capita dividends satisfy (B.6) every period ensuring market clearing.
6. There exists a large enough S, where $T < S < \infty$, such that for any $t \geq S$, a constant set of prices w', r', per capita quantities c' and π', firm value V', per capita $\bar{\mu}'$ and $\bar{M}'$ that satisfy the conditions for a balanced growth path described in Section 3.1.

B.3 Proof of simple formula in the full model

The formula given by equation (2) from Section 2 also applies in the full model from Section 3 where x is replaced by the endogenously determined aggregate exit rate. To see this start with the law of motion in equation (6). Along the balanced growth path, the the law of motion is satisfied with a stationary measure of firms per capita, $\bar{\mu}$. Integrating both sides, then:

[15] This specification allows for anticipation effects from news of future, $s > t$, changes in labor supply growth, $\eta_t = \eta_0 >$ for $t \leq s$.

$$
\begin{aligned}
\iint_{s',a} \bar{\mu}\left(ds', da\right) &= \iint_{s',a} \int_{s} \frac{1 - x\left(s, a\right)}{1+\eta} F\left(ds'|s\right) \bar{\mu}\left(ds, da\right) + \bar{M} \\
1 &= \iint_{s',a} \int_{s} \frac{1 - x\left(s, a\right)}{1+\eta} F\left(ds'|s\right) \frac{\bar{\mu}\left(ds, da\right)}{\iint_{s',a} \bar{\mu}\left(ds', da\right)} + SR \\
1 &= \iint_{s,a} \frac{1 - x\left(s, a\right)}{1+\eta} \frac{\bar{\mu}\left(da, ds\right)}{\iint_{s',a} \bar{\mu}\left(ds', da\right)} \int_{s'} F\left(ds'|s\right) + SR \\
1 &= \frac{1}{1+\eta} \iint_{s,a} \left(1 - x\left(s\right)\right) \frac{\bar{\mu}\left(ds, da\right)}{\iint_{s',a} \bar{\mu}\left(ds', da\right)} + SR \\
&= \frac{1 - \iint_{s,a} x\left(s, a\right) \frac{\bar{\mu}(ds)}{\int_{s'} \bar{\mu}(ds')}}{1+\eta} + SR \\
SR &= \frac{\eta + \int_{s} x\left(s\right) \frac{\bar{\mu}(ds)}{\int_{s'} \bar{\mu}(ds')}}{1+\eta} = \frac{\eta + x}{1+\eta}.
\end{aligned}
$$

The 2nd equality comes from dividing through by the total normalized mass of firms so that the $\bar{M}$ term becomes the startup rate (total normalized mass of startups over total normalized mass of firms). The 3rd equality comes from changing the orders of integration. The 4th equality comes from integrating out the conditional density, which is 1 for every s. The 5th equality comes from noting that $\frac{\bar{\mu}(ds)}{\int_{s'} \bar{\mu}(ds')}$ is the firm density and thus integrates for 1, and thus $x = \int_{s} x\left(s\right) \frac{\bar{\mu}(ds)}{\int_{s'} \bar{\mu}(ds')}$ is the economy's average exit rate.

B.4 Model with convex aggregate entry costs

In Section 3.3.1 we extend the model so that *aggregate* entry costs are convex in the measure of entrants per capita, $\bar{M}$. This change replaces the constant (firm-level) entry cost, c_e, in the free entry condition (5) with a specification $c_e(\bar{M}) = c_{e0}(\bar{M}/\bar{M}_0)^{\psi-1}$ from (9) that is increasing in $\bar{M}$ so that aggregate entry costs per capita, $c_e(\bar{M})\bar{M}$ are convex in $\bar{M}$, strictly when $\psi > 1$. With this change, the free entry condition (5) is replaced by the following entry condition:

$$
c_{e0} \left(\frac{\bar{M}}{\bar{M}_0} \right)^{\psi - 1} = \iint V(a, s; w) G(ds) F(da). \tag{B.7}
$$

The notation makes explicit that expected value of a potential entrant on the right hand side depends (negatively) on the real wage. When $\psi = 1$, this collapses to the free entry condition (5), which could be satisfied at a particular real wage regardless of $\bar{M}$. When $\psi > 1$, a real wage will only satisfy equation (B.7) for a particular measure of entrants per capita, $\bar{M}$. In effect, the "supply" of potential entrants is now upward sloping.

B.4.1 Parameterizing the supply elasticity of potential entrants.

Entry condition (B.7) replaces an infinitely elastic supply of potential entrants with a less than perfectly elastic supply with the supply elasticity inversely proportional to $\psi - 1$. To see this, take logs of (B.7) and differentiate:

$$\frac{d\log \bar{M}}{d\log w} = \frac{1}{\psi - 1}\frac{d\log \iint V(a,s;w)G(ds)F(da)}{d\log w},$$

For a given real wage, $\frac{d\log \bar{M}}{d\log w} \propto \frac{1}{\psi-1}$. Under free entry with $\psi = 1$ (linear aggregate entry costs) supply is infinitely elastic. With $\psi > 1$ (strictly convex aggregate entry costs), supply is less than perfectly elastic and the economy deviates from free entry.

B.4.2 Effects on equilibrium response

With a less than perfectly elastic supply of potential entrants, the equilibrium response to a shift in labor supply requires an adjustment along both price (real wage) and quantity (entry) margins. One can think about the determination of equilibrium real wage, w, and entrants per worker, $\bar{M}$, as clearing a market for entrants. Figure B.3 plots the "demand" and "supply" for entrants. We do this as a function of the price of the final good ($p = 1/w$) rather than the real wage so that demand slopes down and supply up. Demand for entrants, $\bar{M}^d(p)$, is the quantity of entrants required to clear labor markets, i.e., labor supply less *incumbent* labor demand. As price level $p \uparrow$ (real wage $w \downarrow$), incumbent labor demand expands, and demand for entrants falls since fewer entrants are required to clear the labor market.

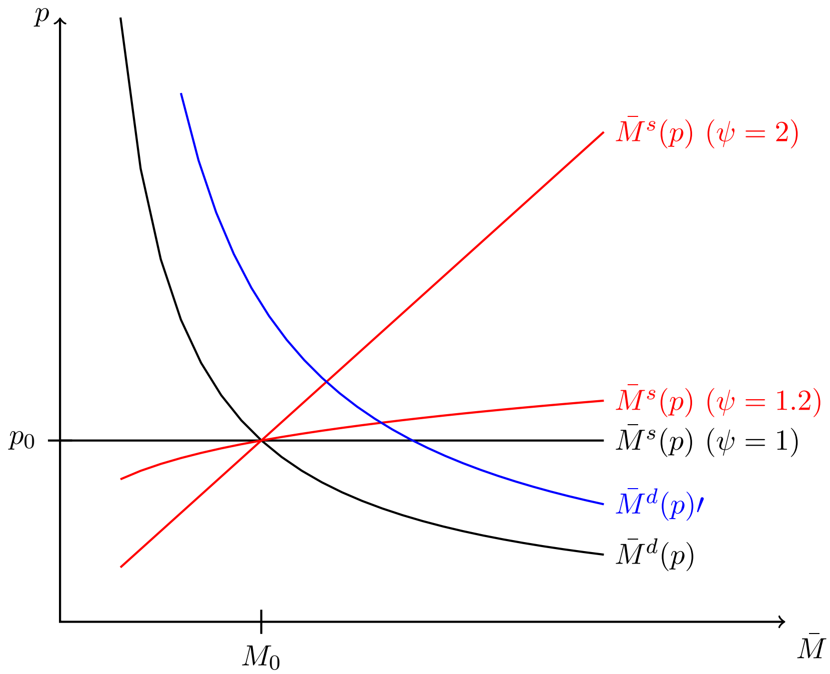


Figure B.3: Market clearing entrants per worker, $\bar{M}$, under deviations from free entry, $\psi = 1$

Supply for entrants, $\bar{M}^s(p)$ is generated by the entry condition, (B.7). Under free entry, $\psi = 1$, the supply of entrants is perfectly elastic, since (B.7) does not depend on $\bar{M}$. As ψ increases above 1 so the supply elasticity falls, the supply of entrants becomes upward sloping. A higher quantity of entry, which increases the entry costs faced by all potential entrants can only satisfy (B.7) with a higher price level (lower real wage).

An equilibrium is the quantity of entry at a given price level (inverse real wage) that simultaneously satisfies demand (clears labor markets) and supply (satisfies B.7). A shift in the demand for entrants, $\bar{M}^d(p)$, for example stemming from a shift in labor supply, is met under free entry by only a change in the quantity of entrants. But, when $\psi > 1$, it requires a change in prices and quantities, $p = 1/w$ and $\bar{M}$. It is too expensive to meet the increase in demand without a change in prices. The larger the parameter ψ, the more expensive it becomes for potential entrants, and the larger the equilibrium response of prices rather than entrants to any shift in demand.

Table 3 in Section 3.3.1 examines quantitatively, how deviations from free entry, $\psi = 1$, affect the relationship between labor supply growth and the startup rate, considering a range of ψ between 1 and 2 (quadratic aggregate entry costs). Gutiérrez, Jones, and Philippon (2019) estimate ψ using industry variation and find a range between 1.2 and 2, with a median value of 1.55. For each value of ψ we compute the entry rate on a grid of $\eta = 0$ to 4% and estimate the elasticity of the startup rate with respect to labor supply growth from equation (8) by regressing the startup rate on η. Table B.3 repeats the results for the elasticity of the startup rate from Table 3 and further includes the elasticity of the exit rate.

Table B.3: Sensitivity of labor supply growth elasticities 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
Elasticity of exit rate, $\frac{dx}{d\eta}$	0.39	0.30	0.24	0.06	−0.12

Relaxing free entry, by allowing the real wage (and thus incumbents) to respond to changes in labor supply growth, attenuates the effect on the startup rate. Recalling from equation (8) that the startup rate elasticity is the sum of a net growth, $1/(1+\eta)$, and a replacement, $\frac{dx}{d\eta}/(1+\eta)$, term. Under free entry, the response of aggregate exit, x, in the replacement term is determined entirely through changes in firm age composition. When free entry is relaxed, in response to an increase in labor supply growth, the real wage falls and incumbents respond: their exit rates conditional on age will decline. This works against the compositional change in determining the effect on the economy's exit rate. For a sufficiently large decline in the real wage, e.g. under quadratic aggregate entry costs, the within age group declines overcome the compositional effects and the contribution from exit is actually negative.

B.5 Computational details

We describe how we compute the solution to the balanced growth path (B.5.1), the transitional dynamics in response to a perfect foresight shock to labor supply growth (B.5.2), and the solution fora balance growth path when entry costs are convex (B.5.3).

B.5.1 Solving for a BGP

We solve the model globally over a discretized state space using the following procedure.

Discretizing firm behavior We approximate firm decisions over a grid for permanent productivity a and the stochastic component of productivity s. We use 3 grid points to discretize the distribution of permanent productivity $F(a)$. Given the log normality in the calibration, we choose the middle grid to be the mean log productivity, which is zero, and the lowest and largest points to correspond to log productivities that are 2.5 standard deviations below and above zero. Once the grid points are chosen, the probability of drawing each productivity level is derived from F, which in our case is the CDF of the log normal distribution.

We use 71 grid points to discretize the stochastic productivity s. We do so using the Tauchen procedure, which gives us the discrete grid of productivities and the associated matrix of transition probabilities. We also use this grid to approximate the initial productivity distribution according to the distribution $G(s)$.

Solution algorithm The algorithm consists of two steps. In the first step, we find the equilibrium wage by solving the free entry condition, and in the second step we solve for equilibrium entry and the stationary distribution of firms, $\bar{\mu}$.

Step 1: Solving for the equilibrium wage. For a given level of wages, we solve firms' optimal size for each grid point. The optimal size decision also gives us the value to the firm of remaining in business, and thus the optimal exit decision. We find the expected value of entry as a function of the wage by integrating the value of a firm over the distribution of permanent and initial stochastic productivities. We use a golden search algorithm to find the value of the wage that satisfies the free entry condition in (5) with sufficient precision.[16] Given the equilibrium wage, we store firms' optimal size and exit decisions.

Step 2: Solving for equilibrium entry. To compute the equilibrium measure of firm entry and the resulting stationary distribution of firms, we make use of equation (6). Note that given the linearity in this equation, it suffices to solve the stationary distribution once for $\bar{M}_t = 1$. We calculate the stationary distribution of firms that correspond to a mass one of new entrants ($\bar{M}_t = 1$) by iterating on an initial guess $\bar{\mu}^0$ using the updating rule defined by equation (6). We stop when the sup norm of the percentage deviation between the guess and the updated

[16] We stop when the percentage deviation between the fixed cost of entry and the expected value of starting a firm is less than 10^{-6}.

distributions is less than 10^{-10}. We then solve for the equilibrium entry $\bar{M}$ by clearing the goods market. Given the linearity in equation (6), the stationary distribution corresponding to any given measure of new entrants $\bar{M}$ is obtained simply by multiplying this distribution with $\bar{M}$.

B.5.2 Solving for the transition path following a perfect foresight shock

We follow the procedure above to solve for the BGP of the initial $\eta = \eta_0$ economy and the terminal (prime) economy where $\eta' = \eta_T$. To solve for the sequence of equilibrium prices and quantities that satisfy the conditions above in B.2 and converge to the terminal BGP, we use the following procedure.

1. Choose a large number for $S > T$.

2. For $j = 0$, conjecture a sequence of per-capita consumption growth $\{g_t^j\}_{t=1}^S$ where
$$g_{t+1}^j = \frac{c_{t+1}}{c_t}.$$

3. Given V', let $V_t = V'$ for $t = S$.

4. Jointly solve (B.3), (B.5) and the Euler equation for r_{t-1}, V_{t-1} and w_{t-1}.
   - We do so using a golden search over the wage. The Euler equation dictates an interest rate given consumption growth.
   - We compute the value of a firm given w_{t-1} and the implied interest rate r_{t-1} and evaluate the value of starting a new firm against the entry cost c_e.
   - The golden search algorithm stops when the percentage deviation between the entry cost and the value of starting a firm is less than 10^{-5}.

5. Repeat step 4 backwards until t=1 and recover $\{w_t, r_t, V_t, h_t, X_t\}_{t=1}^S$.

6. Using the prices and policies recovered in step 4, solve forward on (B.4) from $t = 1$ to $t = S$ starting from $\bar{\mu}_0$, where $\bar{M}_t$ and $\bar{\mu}_t$ in each period jointly solve the law of motion (B.4) and labor market clearing.

7. Compute the implied sequence of consumption growth using $\{\bar{\mu}_t, \bar{M}_t\}_{t=1}^S$ and update $\{g_t^{j+1}\}_{t=1}^S$.

8. Repeat steps 3 to 7 until $\{g_t^{j+1}\}_{t=1}^S$ is sufficiently close to $\{g_t^j\}_{t=1}^S$.
   - We stop when the maximum absolute percentage deviation between the guess and the implied growth path falls below 0.0001.

9. Verify that $\bar{\mu}_S$ is sufficiently close to $\bar{\mu}'$ for the prime economy. If not, increase S and repeat steps 1-8.

B.5.3 Solving the model with imperfectly elastic entry

When entry costs depend on the measure of entrants per capita, the equilibrium wage can no longer be solved for in one step. We use a golden search routine to find the measure of entrants per capita such that the resulting equilibrium entry is consistent with it. The procedure is as follows:

- Conjecture a guess for the measure of entrants per capita, $\bar{M}$ and compute the implied entry cost c_e.
- Solve for the equilibrium wage using Step 1 of the algorithm for solving the BGP.
- Solve for the equilibrium entry per capita $\tilde{\bar{M}}$ that clears the markets following Step 2 above.
- Iterate until the percent deviation between $\bar{M}$ and $\tilde{\bar{M}}$ falls below 10^{-9}

B.6 Comparative statics of changes in costs

Here we supplement the main text discussion in Section 3.3 with additional comparative statics of the BGP to changes in cost. Specifically, we consider changes in entry costs c_e (Figure B.4) and operating costs c_f (Figure B.5).

Entry costs, c_e. When entry costs rise, wages must fall so that the free entry condition holds. Lower wages imply a lower exit rate and therefore a lower startup rate in equilibrium. Lower wages raise the optimal size of a new firm, resulting in fewer but larger startups.

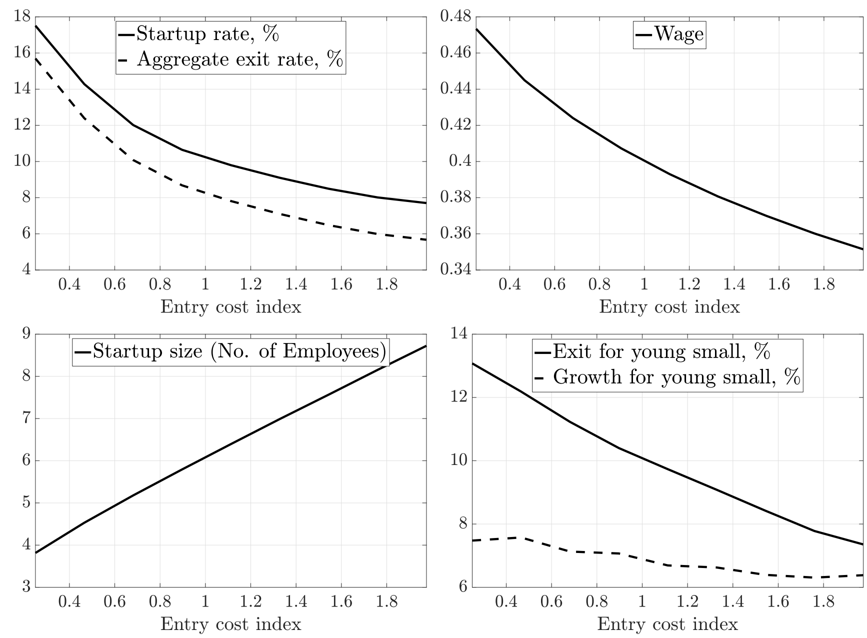


Figure B.4: Entry costs and firm dynamics

Note: The entry cost index is the ratio of the entry cost to its calibrated value so that a value of 1 corresponds to the baseline calibration. Young small refers to age 3 firms with 1-50 employees.

Fixed operating costs, c_f. Rising operating costs have similar effects on wages: By reducing the value of starting a firm, larger operating costs lower the equilibrium wage required by free entry. While lower wages push for less exit, this force is more than offset by the direct of the higher fixed cost, resulting in a lower exit rate and, consequently, also a higher startup rate.

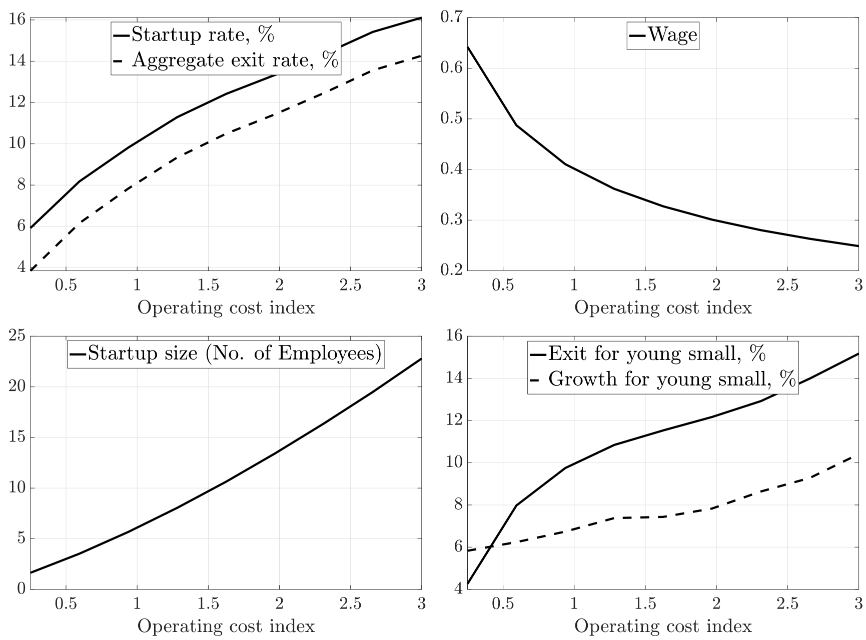


Figure B.5: Operating costs and firm dynamics

Note: The operating cost index is the ratio of the operating cost to its calibrated value so that a value of 1 corresponds to the baseline calibration. Young small refers to age 3 firms with 1-50 employees.

B.7 Change in additional margins over 1979-2007 period

Extending the results in Section 3.5, in Table B.4 we compute changes in several additional margins using the model and compare them against the same changes in the data.

Table B.4: Actual versus model changes for additional margins between 1979-81 and 2005-07 periods

	Actual	Predicted Change	
	Change	η_t =WAP	η_t =CLF
Startup rate (pp)	−2.94	−1.00	−1.7
Economy-wide exit rate (pp)	−0.85	−0.18	−0.6
Average firm size (emp)	2.00	0.78	2.3
Startup size (emp)	0.14	0.04	0.1
Young exit rate (pp)	0.03	0.06	−0.0
Young cond growth rate (pp)	−1.69	−0.54	−0.1

Note: Actual change computed from trend components of each series using HP filter with smoothing parameter of 6.25 averaged over 1979-81 and 2005-07 periods periods, respectively. Predicted change computed from transitional dynamics is response to observed path of WAP or CLF (See discussion in Section 3.5 and above in Aopendix B.5.2).

B.8 Dynamic response to labor supply growth shock

Figure B.6 shows the effect of a persistent shock to labor supply growth on the startup rate for the baseline calibration. The shock, shown in the left panel of Figure B.6, is a full percentage point increase in the labor supply growth rate, which follows by an AR(1) process with persistence 0.6356, estimated from the data. The right panel of Figure B.6 confirms that the elasticity of the startup rate with respect to labor supply growth is around 1. The dynamics of business formation is slower relative to the shock, as the firm age distribution converges is slow to converge to its steady state distribution.

As we discussed in the main text, the impact effect is muted somewhat because the transitory decline of the wage induces incumbent firms to soak up some of the labor supply. While the change in incumbent behavior is rather small relative to the baseline, it has large effects for the startup rate. To illustrate how much, we compute the impulse response of the startup rate in an economy where we let the intertemporal elasticity of substitution to go to infinity, which keeps the interest rate and the real wage constant along the transition path and shuts down the incumbent response entirely. The right panel of Figure B.6 shows that the impact effect on the startup rate is more than two times larger than the baseline.

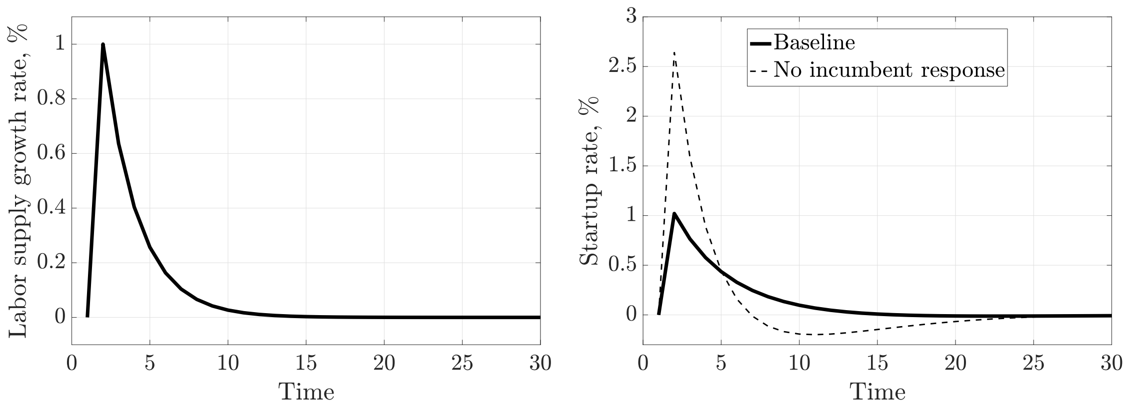


Figure B.6: Startup rate response to a labor supply growth shock

Note: Startup rate IRF to unit shock to AR(1) η_t with $\rho_\eta = 0.6356$ (See B.9); no incumbent response is IES$\rightarrow \infty$.

B.9 Simulating a panel of states from the model

To compute the model counterpart of the startup rate elasticity in Table 6, we create 200 panels of 49 states for 29 periods. States in a panel are ex ante identical but face different shocks to labor supply growth. Specifically, at time $t = 0$, each state is in the balanced growth path equilibrium corresponding to a labor supply growth rate of $\eta = 1.1\%$. In period $t = 1$, each state draws an unexpected shock to labor supply growth from a normal distribution with mean zero and a standard deviation σ_η. Thereafter, no further shocks take place and labor supply growth in each state returns to its steady state value of 1.1% following an AR(1) process with persistence ρ_η. Given the path of labor supply growth η_t for each state, we then solve for the transition path. To discipline the parameters governing the dynamics of labor supply growth, we estimate an AR(1) process using state-level data on working age population growth rate. We find $\rho_\eta = 0.6356$ and $\sigma_\eta = 0.0053$.

C Robustness Appendix

C.1 Trends in firm and labor market dynamics

We supplement the analysis in main text Section 2 with these additional details on the aggregate trends in firm and labor market dynamics. Much of this draws on Pugsley and Şahin (2019) and its robustness appendix.

C.1.1 Aggregate labor supply growth and the startup rate

Figure C.1 plots a smoothed series for the startup rate and our two main proxies of labor supply growth: *working age population* growth surges in the 1960s as early "baby boomers" enter adulthood; *civilian labor force* growth accelerates even faster, because it combines the growth in the working age population with rapidly increasing female participation. The startup rate falls by roughly 3 percentage points exactly over the ensuing period of declining labor supply growth.

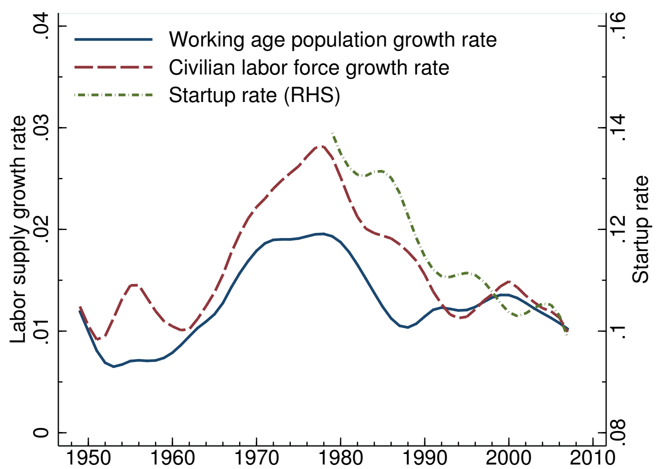


Figure C.1: Trend components of startup rate and labor supply growth rates

Note: Current Population Survey, Census Bureau annual population estimates, Business Dynamics Statistics. Annual data, HP filtered with smoothing parameter 6.25. See appendix Figure C.2 for unfiltered rates. Working age population is ages 20 to 64. Civilian labor force is measured for the adult (16+) civilian non institutional population. Startup rate is number of age 0 (employer) firms as share of the total number of firms within a year.

The unfiltered data, though noisier, show the same patterns. In Figure C.1 we had smoothed the data to remove the high frequency fluctuations primarily in the growth of the civilian labor force, which is both volatile and highly procyclical. However, even in the raw time series, the lower frequency comovement between both measures of labor supply growth and the startup rate is evident. Figure C.2 plots the unfiltered data.

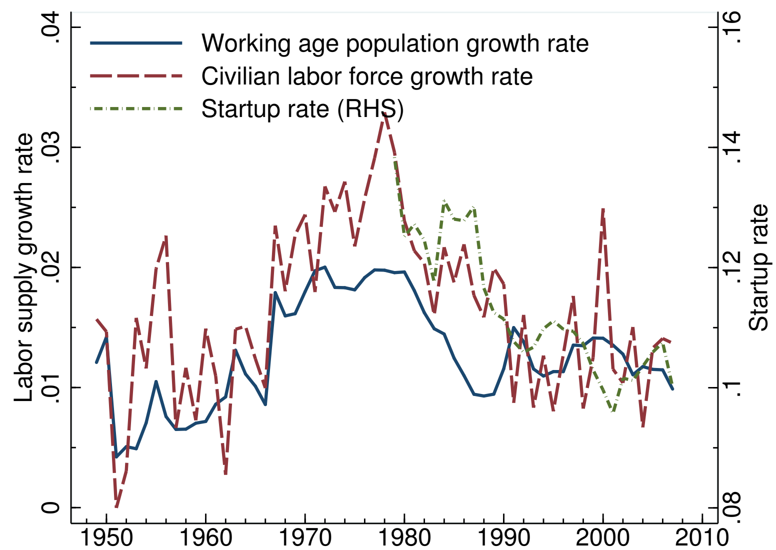


Figure C.2: Unsmoothed aggregate data on startup rate and labor supply growth

Actual and flow balance startup rates Table 1 in the main text reports the actual and flow balance predicted startup rates for the 1979-81 and 2005-07 3-year periods. To be consistent throughout the paper, the actual startup rate is the average within each 3 year period of a smoothed startup rate series estimated using an HP filter with a penalty parameter of 6.25 as suggested by Ravn and Uhlig (2002) for annual data. The economy-wide exit rate for each period is computed the same way. The respective flow balance startup rates are computed using the average labor supply growth rates and average exit rates for each 3-year period. The results are very similar when computed on the unfiltered data. We report below in Table C.1 a version of Table 1 instead computed using 3-year averages of the raw data for the startup rate and exit rate series.

Table C.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	1.91	2.49	9.90	13.03	11.59	12.09
2005-2007	1.10	1.37	8.58	10.45	9.57	9.82
Change	−0.81	−1.12	−1.32	−2.58	−2.02	−2.27

Note: Startup rate, exit rate, and labor supply growth rates for working age population (WAP) and civilian labor force (CLF) measured as 3 year averages of raw data. Predicted startup rates use flow balance equation (2) with 3-year averages for η and exit.

Next, we present the entire annual time series for the predicted flow-balance startup rate (Figure C.3). To form the time series, we calculate the flow-balance startup rate using equation (2) with the realized exit rate x_t and labor supply growth rate η_t for each year. That is, the predicted startup rate in each year is the one we would expect if the annual labor supply growth (Figure 1b and Figure C.2) and the actual average exit rate (Figure 1a) in each year were to prevail indefinitely. The largest declines in both flow-balance predicted startup rates occurs before the actual startup rate. This is to be expected since the flow-balance calculation in each year is based on the long-run effects of the change in labor supply growth and exit. It suggests that the period where the startup rate remains above the flow-balance predictions are part of a transition to the new BGP.

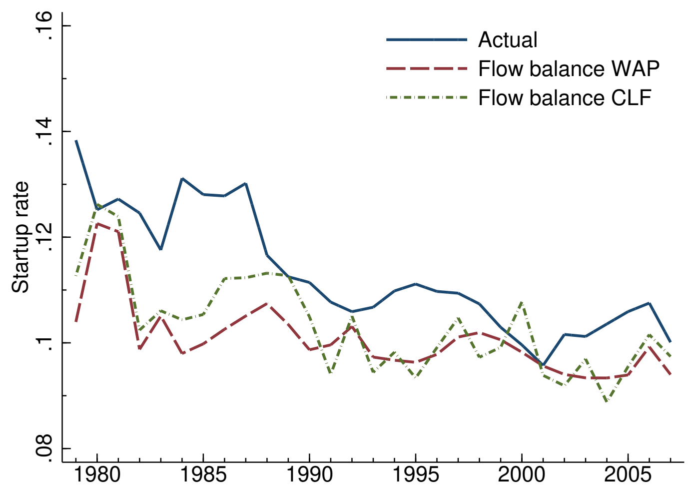


Figure C.3: Actual and flow balance startup rates for 1979 to 2007

C.1.2 Effects of compositional change on the aggregate startup rate

An immediate concern when interpreting the decline in the aggregate startup rate and the ensuing startup deficit, is that the aggregate decline may primarily reflect compositional changes in business sectors.[17] If sectors with lower startup activity are becoming more important because of ongoing structural change, this ongoing reallocation of employment may explain the declines in the aggregate startup rate even if startup rates by industry were unchanged. Here we replicate several exercises from Pugsley and Şahin (2019), who find little support for this hypothesis.

Compositional changes from structural transformation have, if anything, slowed the aggregate decline in the startup rate. Table C.2, reproduced from Pugsley and Şahin (2019) Table 1, reports the average startup rate by sector and by time period. Panel A reports the startup rates for each NAICS sector or supersector, and panel B reports a special aggregation of high-tech industries that draws from the manufacturing, information and professional services sectors as in Decker, Haltiwanger, Jarmin, and Miranda (2016). Two features are immediately apparent: First, relative to the early 1980s average, the startup rate has declined in all of these sectors. Even in the high-tech sector, containing firms in the 14 NAICS 4-digit industries with highest share of STEM workers

[17]The U.S. economy has been undergoing a significant structural transformation—the secular reallocation of employment across sectors—over the past several decades. See for example Duarte and Restuccia (2010) and Dent, Karahan, Pugsley, and Şahin (2016) for additional details.

Table C.2: Average sector startup rates by time period (Pugsley and Şahin 2019, Table 1)

Sectors	1980-1984	2003-2007	2008-2012
A. NAICS Sectors			
Mining (21)	0.182	0.097	0.095
Utilities (22)	0.067	0.053	0.039
Construction (23)	0.140	0.126	0.084
Manufacturing (31-33)	0.102	0.064	0.052
Wholesale Trade (42)	0.110	0.080	0.067
Retail Trade (44-45)	0.122	0.109	0.880
Transportation and warehousing (48-49)	0.146	0.136	0.116
Information (51)	0.160	0.118	0.098
Financial activities (52-53)	0.128	0.115	0.083
Professional and business services (54-56)	0.165	0.118	0.098
Education and healthcare (61-62)	0.101	0.085	0.072
Leisure and hospitality (71-72)	0.165	0.139	0.120
Other services (81)	0.118	0.076	0.064
B. Other Sectors			
High tech industries	0.173	0.120	0.100

Note: U.S. Census Bureau Longitudinal Business Database. Number of age 0 firms as fraction of total firms within each sector. 2-digit NAICS sectors listed in parentheses for each sector in panel A. In panel B. high tech sector is not mutually exclusive and is comprised of 14 NAICS 4-digit industries with highest share of STEM workers: 3341, 3342,3344, 3345, 5112, 5161, 5179, 5181, 5182, 5415, 3254, 3364, 5413, 5417. See Decker, Haltiwanger, Jarmin, and Miranda (2016) for additional details.

and in which entry rate increases in the late 1990s, the startup rate still declined from 17.3 percent in 1980-84 to 12 percent in 2003-2007 and further to 10 percent in the 2008-2012 period. Second, sectors with declining employment shares such as manufacturing already had among the lowest startup rates in the 1980s. Structural transformation, which reallocates employment away from manufacturing and into service providing sectors with higher startup rates, has weighed against the aggregate decline in the startup rate. Even at finer levels of disaggregation, more than 100 percent of the aggregate declines from since the 1980s are within industry.

To evaluate this explanation more formally we decompose the decline in the aggregate startup rate from the 1980-84 period to the 2008-2012 period into three components: within 4-digit NAICS industry changes, between industry changes and a covariance term (Table C.3, reproduced from Pugsley and Şahin 2019 Appendix Table B.1). Within industry declines account for more than 100 percent of the declines in the aggregate startup rate. This pattern is robust to an alternative period that does not include the financial crisis as well as when computed for startup employment shares. Ultimately, compositional shifts across industries, if anything, moderated the decline in startup formation. Startup deficits are also present even in narrowly defined geographic markets. In the center panel we present the same decomposition applied to U.S. counties instead of industries to investigate whether changes in geographic allocation of employment can explain the decline in startups. Similar to national industries, more than 100% of the aggregate decline is within county.

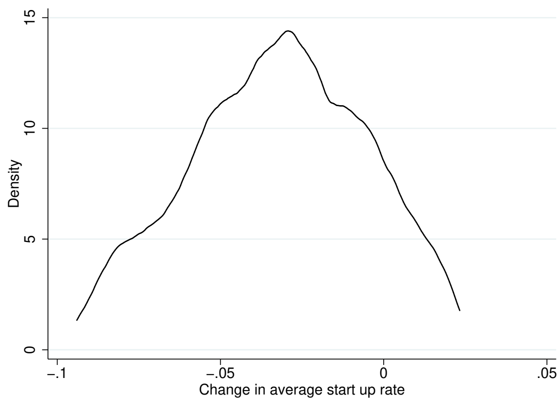
Allowing simultaneously for both industry and geographic shifts, we find the same pattern even within industry–geography submarkets. We evaluate the decline in the startup rate within 4-digit

Table C.3: Decomposition of the startup rate and startup employment share changes into between, between and covariance components (Pugsley and Şahin 2019, Appendix Table B.1).

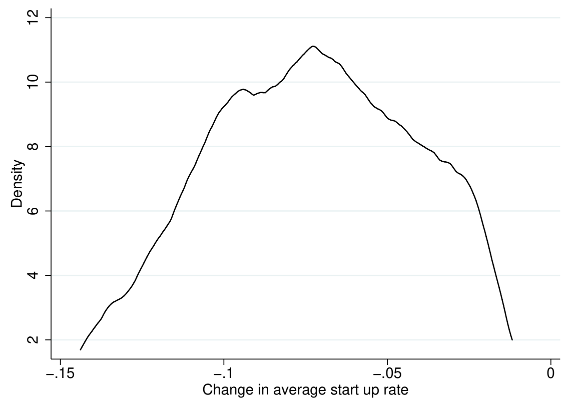
	Startup Firm Share			Startup Employment Share		
	Between	Within	Covariance	Between	Within	Covariance
			A. By Industry (NAICS4)			
1980-84;2003-07	1.19	-2.87	-0.59	1.16	-1.11	-0.82
	(-52.2%)	(126.4%)	(25.8%)	(-149.6%)	(143.4%)	(106.2%)
1980-84;2008-12	1.37	-4.98	-0.76	1.24	-1.80	-0.96
	(-31.4%)	(114.0%)	(17.4%)	(-81.1%)	(118.1%)	(63%)
			B. By County			
1980-84;2003-07	0.68	-2.87	-0.09	0.71	-0.93	-0.57
	(-29.9%)	(126.2%)	(3.73%)	(-90.8%)	(118.5%)	(72.3%)
1980-84;2008-12	0.80	-4.82	-0.32	0.81	-1.62	-0.72
	(-18.3%)	(110.9%)	(7.4%)	(-53.1%)	(105.8%)	(47.3%)
			C. By State × Industry			
1980-84;2003-07	1.44	-3.11	-0.59	1.75	-0.72	-1.80
	(-63.2%)	(137.1%)	(26.1%)	(-225.7%)	(92.9%	(232.8%)
1980-84;2008-12	1.58	-5.20	-0.75	1.83	-1.31	-2.05
	(-36.1%)	(118.9%)	(17.2%)	(-119.9%)	(85.7%)	(134.2%)

Note: U.S. Census Bureau LBD. Decomposes change in average startup firm (employment) share from 1980-1984 period to 2003-2007 or 2008-2012 period. See Pugsley and Şahin (2019) Appendix B.2 for exact decomposition.

NAICS industry and state pairs, which yields roughly 13,000 submarkets. Again, the declines are within these narrow submarkets. In all cases, the structural transformation captured by the between terms actually puts upward pressure on the aggregate startup rate. Another way to visualize the widespread nature of the declines in the startup rate is to examine a histogram of the within state and industry long-run changes. For each state and 4-digit NAICS industry, as above in Table C.3 Panel C, we compute the change in the state×industry startup rate since its 1980-84 average. Figure C.4, reproduced from Pugsley and Şahin Figure B.4, plots the histogram of the changes from 1980-84 to 2003-07 (left panel) and to 2009-11 (right panel). Over the period that does not include the Great Recession, almost 85 percent of state industry pairs have declines, a share that rises to nearly 100 percent when the change is computed over the longer time period.



(a) 1980-84 to 2003-07



(b) 1980-84 to 2009-11

Figure C.4: Density estimates of distribution of long run changes in startup rate and employment share over alternative time periods (Pugsley and Şahin 2019, Figure B.4)

C.1.3 Declines in additional measures of entry

In the main text, we focus primarily on the firm startup rate, but the declining entry rate is a robust feature of the data. Here, drawing on Pugsley and Şahin (2019) Appendix B.2, we consider several additional measures of entry activity and show declines in each of the 1979-2007 period.

First, one may worry that defining entry by considering only age 0 firms is too restrictive. To address this concern, we extend our definition of startups to age 0 and age 1 firms and define entry measures accordingly. As a share of all firms and of private payroll employment both broader measures of startups show similar declines (Figure C.5, Pugsley and Şahin 2019 Figure B.5). We can also define an entry measure as the number of new firms per capita, which we plot in Figure C.6, Pugsley and Şahin 2019, Figure B.6. This measure also declines over the 30-year period. This also closely tracks the model's prediction that the number of firms per worker should decrease in the labor-supply growth rate / startup rate. A final concern is that the decline the startup rate stems in part from our choice measuring firms rather than establishments. We plot the establishment entry rate and age 0 establishment employer share (Figure C.7, Pugsley and Şahin 2019 Figure B.7). Both measures show a similar decline to our preferred measure of the firm startup rate (Figure 1a).

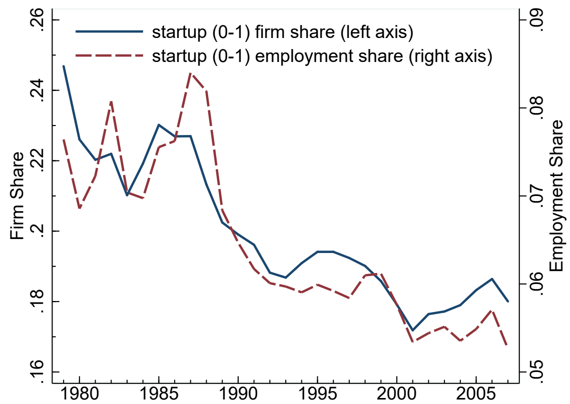


Figure C.5: Startups defined as firms ages 0-1: firm and employment share 1979–2007 (Pugsley and Şahin 2019, Figure B.5)

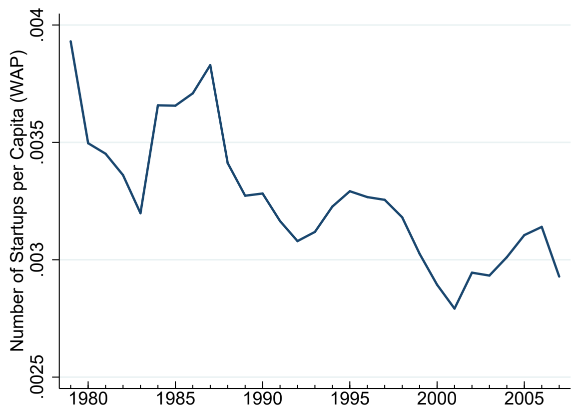


Figure C.6: Number of startups per working age population 1979–2007 (Pugsley and Şahin 2019, Figure B.6)

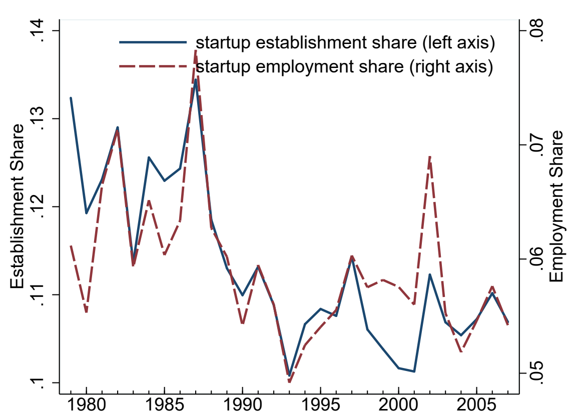


Figure C.7: Startup (age 0) establishment shares 1979–2007 (Pugsley and Şahin 2019, Figure B.7)

C.1.4 Firm age vs establishment age trends

As discussed in the main text Section 2 we focus on firms rather than establishments. While in practice, the distinction makes little difference (see Figure C.7 and discussion above) since most firms are single establishments, conceptually, the distinction is important. We take firm boundaries as defined only by the span of managerial control over inputs that ultimately limits each firm's scale, as in Lucas (1978). A firm may be composed of multiple establishments, whose boundaries are instead defined by their specific technologies and physical locations. We are interested in the firm startup rate since it captures the birth of a new organization. An establishment startup rate will capture expansions and reallocation of incumbent organizations' establishments in addition to those by new firms.

Empirically, the trend decline in entry is only found for new firms and their establishments. For establishments of incumbent firms, there is *no decline* in the entry rate. Figure C.8 compares the firm startup rate and establishment entry rate, against the establishment entry rate computed separately for new and incumbent firms. When measured only for establishments belonging to incumbent firms, the establishment entry rate is roughly flat over the 1979-2007 period. It is only declining when measured for establishments at new firms. Since the vast majority of new establishments are also new firms, this entry rate tracks the overall establishment entry rate and the firm startup rate closely. The lack of a trend decline in the entry rate of establishments created by incumbent firms further reinforces our choice of firms as the unit of analysis since it is tied to the decision-making rather than the physical technology.[18]

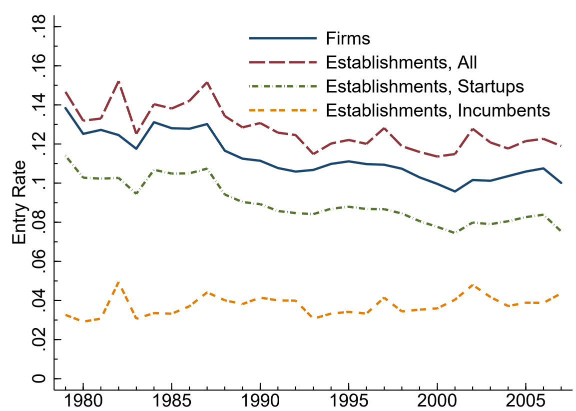


Figure C.8: Firm and Establishment Entry Rates 1979–2007

C.1.5 Trends in other margins

Table C.4 confirms the stability of exit and conditional employment growth rates for more detailed age groups. We consider three age groups within the young firm age category: 2-3, 4-5 and 6-10

[18]We condition on firm age throughout the paper for the same reason: it is the age of the decision-making organization. See also the discussion on firm age in Haltiwanger, Jarmin, and Miranda (2013).

years old firms and mature firms (11+ years) as well as three size categories: small (1-49 employees), medium (50-249 employees) and large (250+ employees) firms. We filter the exit and employment growth rates by firm age and size with H-P filter using smoothing parameter 6.25 to remove higher frequency fluctuations and report the estimated linear trend of the filtered component. Columns (1) to (8) report the estimated coefficient on the linear trend and show that the stability result still holds. For both young and mature firms—regardless of their sizes—the estimates are quantitatively insignificant. For example, the estimated trend implies that over thirty years, the exit rate of both young and old firms will have changed only by a fraction of 1%.[19]

Table C.4: Average slope of HP trend for exit rate and conditional growth rates, 1987-2007

	Exit Rate x_t				Conditional Employment Growth Rate n_t			
	All Sizes (1)	Small (2)	Medium (3)	Large (4)	All Sizes (5)	Small (6)	Medium (7)	Large (8)
	A. Firm Age 2-3 Years							
Trend	-0.0002*	-0.0002*	-0.0006***	0.0006***	0.0003	0.0002	0.0006	0.001
	(0.0001)	(0.0001)	(0.0002)	(0.00010)	(0.0004)	(0.0003)	(0.0005)	(0.0010)
R^2	0.26	0.25	0.45	0.54	0.03	0.04	0.12	0.08
RMSE	.0024	.0024	.004	.0034	.011	.0075	.011	.025
N	21	21	21	21	21	21	21	21
	B. Firm Age 4-5 Years							
Trend	-0.0002**	-0.0002**	-0.00010	0.0001	-0.00003	-0.00010	0.0002	0.004***
	(0.00008)	(0.00008)	(0.00007)	(0.00008)	(0.0002)	(0.0002)	(0.0002)	(0.0009)
R^2	0.31	0.31	0.12	0.16	0.00	0.01	0.03	0.60
RMSE	.0017	.0017	.0017	.0018	.0047	.0055	.007	.022
N	21	21	21	21	21	21	21	21
	C. Firm Age 6-10 Years							
Trend	0.00007	0.00006	0.0001***	0.0003***	-0.0005**	-0.0003*	-0.0003	-0.0006**
	(0.00008)	(0.00008)	(0.00001)	(0.00003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
R^2	0.06	0.04	0.82	0.86	0.19	0.16	0.11	0.13
RMSE	.0017	.0018	.00043	.00084	.0065	.0046	.0057	.01
N	21	21	21	21	21	21	21	21
	D. Firm Age 11+ Years							
Trend	-0.0004***	-0.0004***	0.00003	-0.00004**	-0.0006***	-0.0005***	-0.0001	-0.0002
	(0.00009)	(0.0001)	(0.00004)	(0.00001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
R^2	0.61	0.63	0.03	0.24	0.40	0.45	0.03	0.05
RMSE	.002	.0021	.001	.00041	.0044	.0037	.0045	.0041
N	21	21	21	21	21	21	21	21

We also plot the raw data, conditional on size and age. We pool ages 1-10 in the young category. Figure C.9 plots exit rates and Figure C.10 plots the conditional growth rates. These correspond

[19] This finding is robust to controlling for sectors and states. See extensive robustness exercises in Pugsley and Şahin (2019) Appendix B.3.

to Figure 3 in the paper, except now further conditioned on size.

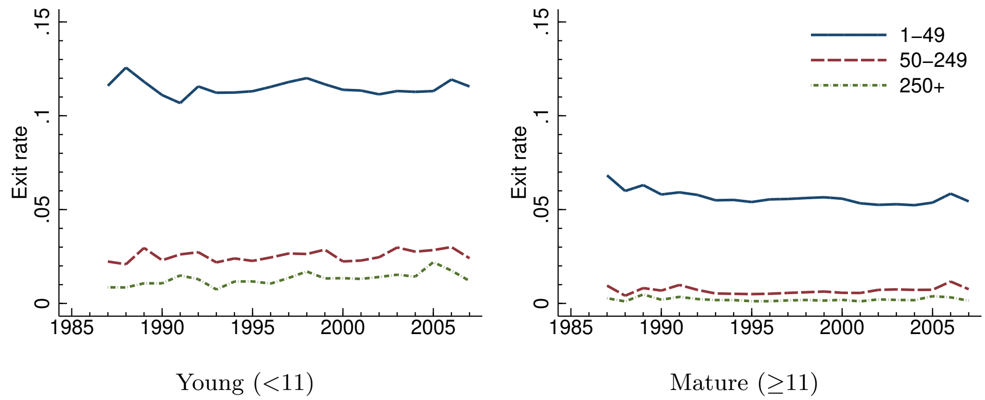


Figure C.9: Incumbent exit rates by firm size for young and mature firms

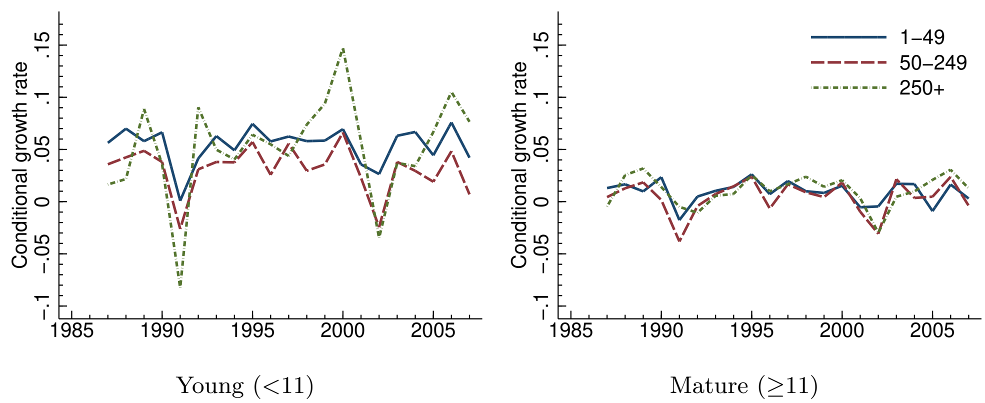


Figure C.10: Incumbent growth rates by firm size for young and mature firms

C.1.6 Trends in gross versus net flows

The stability of the employment growth rates by firm age may appear inconsistent with the body of work by Decker, Haltiwanger, Jarmin and Miranda (2014b, 2014a, 2016, 2020) emphasizing within-age group declines in job creation and destruction. There is no doubt that job creation and job destruction rates have fallen, especially since the early 2000s.[20] But, the relative stability we observe for the exit and conditional growth rate components of employment growth is a product of observing net (creation minus destruction) rather than gross flows. It turns out that within age groups, both gross flow measures have fallen by roughly the same amount. Job reallocation is of course reduced, but the effect on the net job creation is roughly a wash.[21]

[20]Changes in age *composition* still play a role in the overall declines in these measures, accounting for roughly 1/3.

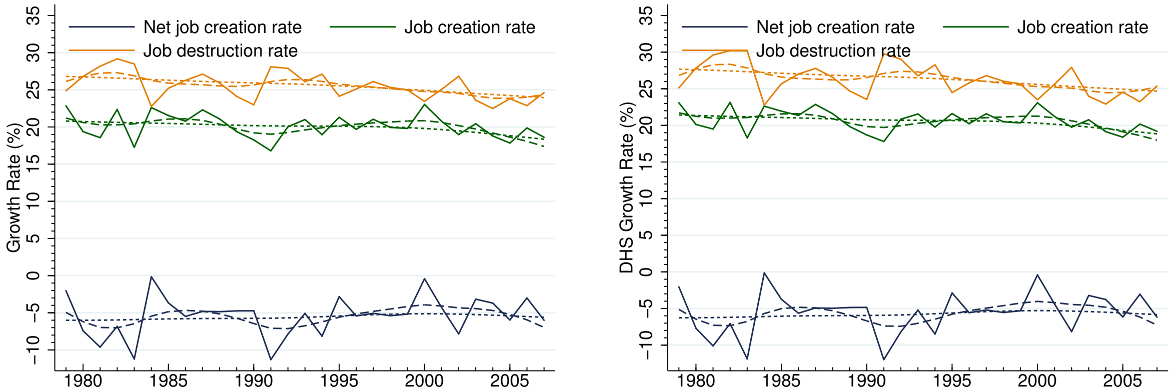
[21]Net job creation by age group, $1+g_t^a$ is the product of the survival rate, $1-x_t^a$, and the conditional growth rate, $1+n_t^a$. (See discussion above in Appendix A.1.4.) We look at these components separately, but the products (net job creation margins) by firm age are also stable.

To further examine this point we have calculated the gross job creation and job destruction rates, and the net job creation rates for several age groups of young incumbents. Figure C.11 shows the raw rates of job creation, destruction and net job creation for the 1979 to 2007 period.[22] The solid line is the raw data, and the dashed lines the the trend components from HP filtering using two different smoothing parameters.[23] In our paper we use traditional growth rates (left column), but we have also calculated the Davis, Haltiwanger and Schuh (1996) symmetric growth rates (right column)—the distinction makes little difference.

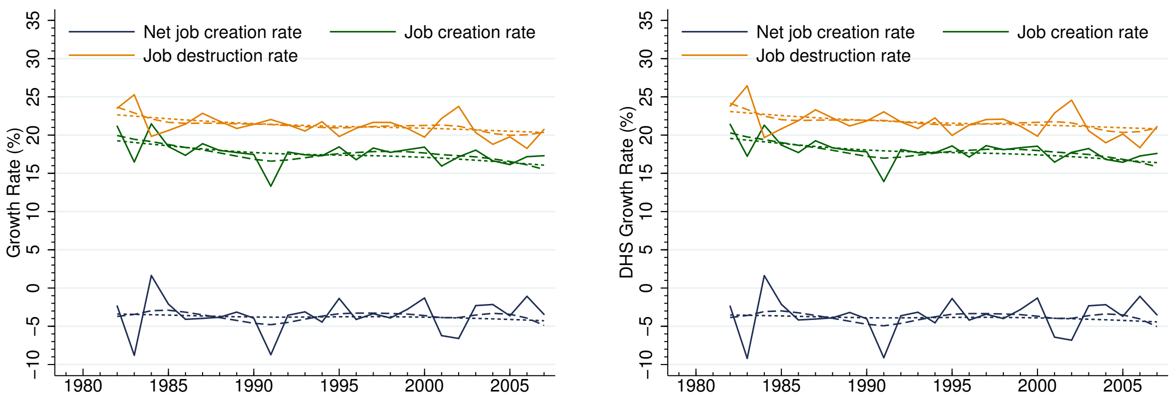
Inspecting the plots, several points are apparent: (i) the (conditional) job creation/destruction rates are volatile (much more than exit rates for example) making it challenging to definitively identify any structural breaks. Nevertheless, (ii) as first uncovered by Decker, Haltiwanger, Jarmin and Miranda it seems clear that starting sometime in the early 2000s the gross job creation and job destruction rates start tending down. Yet, over this period, (iii) the job creation and job destruction rates are tending to decline together, leaving little trace of a decline in the net job creation rate series. The heavily smoothed series ($\lambda = 400$) looks almost flat.

[22] The plots start at different years because of birth year censoring.

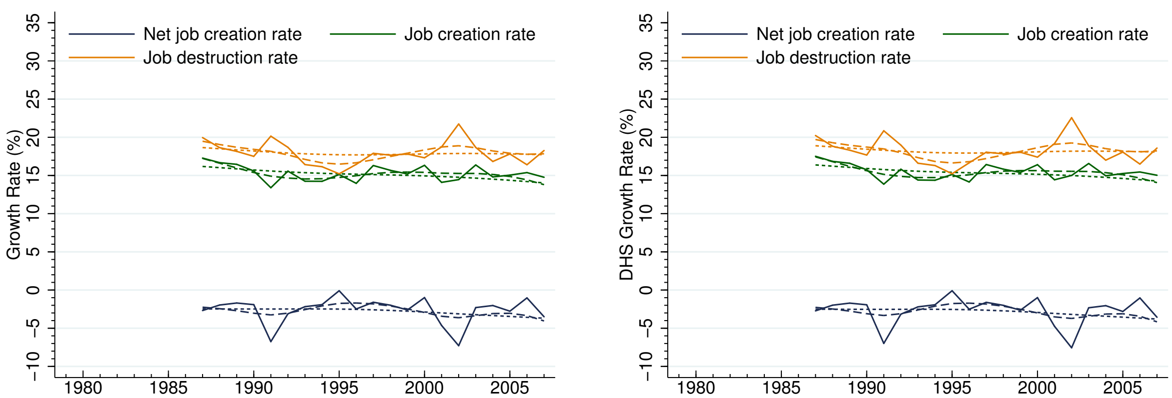
[23] In our paper, we use the HP smoothing parameter of 6.25 suggested by Ravn and Uhlig (2002) for annual data. The smoother series matches Decker, et al (JEP 2014) who use the a parameter of 400.



A. Firm ages 2-3



B. Firm ages 4-5



C. Firm ages 6-10

Figure C.11: Trends in annual gross and net job flows by firm age group

Note: Solid line is raw data, dashed line is the trend component of the HP filtered series with smoothing parameter 6.25, dash dot is trend component with smoothing parameter 400 as used in Decker, Haltiwanger, Jarmin and Miranda (2014). Left column is standard growth rates, right column is Davis, Haltiwanger and Schuh (1996) symmetric growth rates.

C.2 Cross-state results

Next, we supplement our analysis in main text Section 4 with additional cross-state tests of the mechanism and other supporting material.

C.2.1 IV Robustness

Alternative migration instrument construction We present here a set of alternatives to the migration IV in the main text. We adjust whether or not to exclude states from the same Census division and the source of the origin state "push". We first describe the baseline, $\hat{n}_{st}$, and its three alternatives: a,b,c. In each formulation, recall that $\omega^k_{st^*}$ is the share of residents of state s in year t^* that were born in state k.[24]

Baseline This is defined in the main text in equation (11). For reference:

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

Here, we exclude summing over pushes from states in the same Census division in case those states' working-age population growth is correlated with destination state business conditions.

Migration IV$_a$ This version is identical to above, except it includes states in the same Census division, i.e., it excludes only itself:

$$\hat{m}^a_{st} = \sum_{k \neq s} \omega^k_{st^*} g_{kt}$$

This version relies on a stronger exclusion restriction that neighboring origin state's working-age population growth is also uncorrelated with destination state business conditions.

Migration IV$_b$ This version replaces each origin state contemporary working-age population growth "push" with that state's 20 year lagged birthrate, b_{kt-20}, which predicts contemporary working-age population growth:

$$\hat{m}^a_{st} = \sum_{k \neq s} \omega^k_{st^*} b_{kt-20}$$

This IV is equivalent to the hybrid IV used in Section 4.3. Its exclusion restriction requires only that all other state's lagged birthrates are uncorrelated with the destination state's business conditions.

Migration IV$_c$ This version relaxes further the requirements of IV_b, by excluding neighboring state lagged birthrate pushes:

$$\hat{m}^c_{st} = \sum_{k \notin C(s)} \omega^k_{st^*} b_{kt-20}$$

Here, only lagged birthrates outside of the destination state's Census division need to be uncorrelated with its contemporary business conditions.

[24] To isolate the historical component of migration patterns, we use the birthplace shares $\omega^k_{st^*}$ from 2 censuses ago in the IPUMS microdata, see Ruggles, Genadek, Goeken, Grover, and Sobeck (2017), ong form responses to the 1970, 1980, 1990 Decennial Censuses. In 1979, the lag is 9 years, i.e., we set $t^* = 1970$ instead of 1960. Appendix A.3.2 provides additional details on the migration instrument construction.

Next, Table C.5 shows the first stage for each of the alternative migration IVs and the baseline when used individually and in conjunction with the birthrate IV. The first two columns match Columns (1) and (2) from Table 6 in the main text. The IVs in all specifications perform well; the alternatives have even stronger first stages than the baseline.

Table C.5: First stage regressions for WAP Growth using alternative migration IVs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	WAP Growth (%)							
Birthrate IV		1.11		0.41		0.66		1.05
		(0.24)		(0.26)		(0.36)		(0.27)
Migration IV	1.04	0.87						
	(0.30)	(0.28)						
Migration IV_a			1.69	1.62				
			(0.20)	(0.22)				
Migration IV_b					0.43	0.31		
					(0.07)	(0.11)		
Migration IV_c							0.48	0.29
							(0.12)	(0.13)
N	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R^2	0.64	0.65	0.71	0.71	0.64	0.65	0.64	0.64
F-test	11.82	17.46	68.96	44.35	37.01	32.45	15.68	21.68

Note: Standard errors clustered on state. State and year fixed effects, years 1979-2007 and lower 48 plus D.C.

Finally, we find that each alternative delivers very similar estimates of the main elasticity to the baseline. Table C.6 shows the 2SLS estimates of cross-state startup rate elasticity using each migration IV alone and together with the birthrate IV. Here the first two columns will match Columns (6) and (7) from Table 6 in the main text.

Table C.6: 2SLS estimates for startup rate and WAP growth using alternative migration IVs

	(1) IV2	(2) IV1&IV2	(3) $IV2_a$	(4) IV1&$IV2_a$	(5) $IV2_b$	(6) IV1&$IV2_b$	(7) $IV2_c$	(8) IV1&$IV2_c$
WAP Growth (%)	1.27	1.19	1.12	1.12	1.23	1.19	1.37	1.20
	(0.22)	(0.22)	(0.11)	(0.11)	(0.26)	(0.25)	(0.41)	(0.31)
N	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R^2	0.85	0.86	0.87	0.87	0.85	0.86	0.83	0.86
J test p-value		0.55		0.93		0.54		0.39

Note: Standard errors cluster on state. IV2 is the migration IV and its alternatives a,b,c. State and year fixed effects, years 1979-2007 and lower 48 plus D.C.

IV effects on age composition One concern, addressed in Section 4.2.1, is that one or both IVs may have effects on the share of young workers, which may itself be complementary with the startup rate. As we discuss, this turns out not to be the case. We can directly examine the systematic effects of lagged fertility on share of young people (age 20-34) in the working age population after removing state and year fixed effects. Table C.7 regresses this age 20-34 share on the birth rate and migration instruments individually and together. The partial R^2 row reports the incremental increase in fit from adding the instruments to a baseline regression that includes only state and year fixed effects. The IVs on their own and in conjunction have statistically negligible effects on the share of young workers.

Table C.7: Predicting the share of young workers

	(1) 20-34 Share	(2) 20-34 Share	(3) 20-34 Share
Birthrate IV	−0.32		−0.28
	(0.89)		(0.93)
Migration IV		−0.17	−0.13
		(0.25)	(0.31)
N	1,421	1,421	1,421
R^2	0.95	0.95	0.95
Partial R^2	0.00	0.00	0.00

Note: Standard errors clustered on state. Regression of 20-34 year old share of working age population on each instrument and state and year fixed effects. 48 contiguous states plus D.C., and years 1979 to 2007. Partial R^2 is increase in R^2 from adding instrument relative to regression with only fixed effects.

C.2.2 Civilian labor force results

We present the full set of results using civilian labor force (CLF) growth rather than working age population (WAP) growth as the measure of labor supply growth. Figure C.12 replicates using CLF growth the long-run scatter plot using WAP growth in the main text (Figure 7). Figure C.13 replicates Figure 8 from the main text, here predicting CLF growth with the instruments. Scatter plot points are first residualized on state and time fixed effects. Next, Table C.8 replicates Table 6 from the main text using the CLF measure of labor supply growth.

Figure C.12: Average CLF growth and startup rates for U.S. states over 1979–2007.

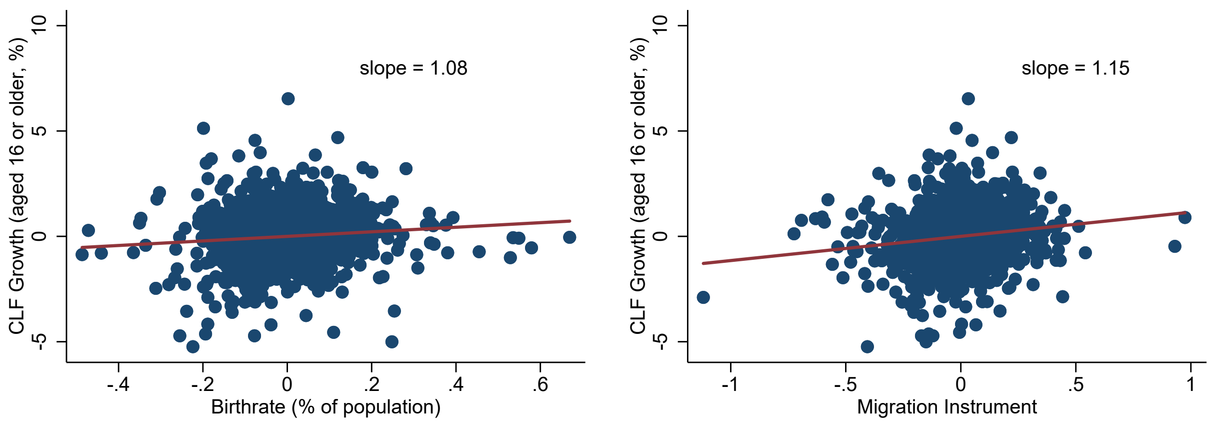


Figure C.13: First-stage regressions of CLF growth rate on fertility and migration instruments.

Table C.8: Startup rate and civilian labor force growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First Stage			OLS	IV_1	IV_2	IV_1 & IV_2
CLF Growth (%)				0.22	1.38	1.14	1.20
				(0.04)	(0.61)	(0.28)	(0.28)
Birthrate IV	1.08		0.78				
	(0.34)		(0.28)				
Migration IV		1.15	1.03				
		(0.36)	(0.35)				
N	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R^2	0.41	0.42	0.42	0.87	0.44	0.60	0.56
F-test	9.92	10.10	8.44				
p-value of J test							0.70

Note: Standard errors are clustered on state. State and year fixed effects, years 1979-2007 and lower 48 plus D.C.

C.2.3 Spatial correlation

Firm and labor market activity may be spatially correlated, e.g., across adjacent states. In the results from the main text we compute standard errors clustering on state, which allows for arbitrary serial correlation within a state but assumes that observations across states are uncorrelated. To the extent there is spatial correlation, the standard errors may be biased down.[25]

Table C.9: Adjusting standard errors for spatial correlation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		IV_1			IV_2			IV_1&IV_2	
	CGM	DK	THOM	CGM	DK	THOM	CGM	DK	THOM
WAP Growth (%)	1.09	1.09	1.09	1.27	1.27	1.27	1.19	1.19	1.19
	(0.33)	(0.22)	(0.32)	(0.26)	(0.27)	(0.29)	(0.25)	(0.22)	(0.26)
N	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421	1,421
R^2	0.87	0.87	0.87	0.85	0.85	0.85	0.86	0.86	0.86
p-value of J-test							0.58	0.55	0.61

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.

Our cross state results remain significant even allowing for spatial correlation. We consider in Table C.9 three corrections that have been proposed in the literature. In Columns (1), (4) and (7) we estimate the elasticity of the startup rate with birthrate, migration and joint IVs, respectively, where we compute standard errors using the two-way procedure developed by Cameron, Gelbach, and Miller (2011), which allows for arbitrary spatial correlation within a year and arbitrary serial correlation within a state. Next, in Columns (2), (5) and (8) we compute standard errors using the procedure from Driscoll and Kraay (1998), which allows for arbitrary spatial correlation within a

[25] See, for example, Foote (2007).

year and corrects for aggregate serial correlation using a Newey-West procedure with a bandwidth of 3 years. Finally, for columns (3), (6) and (9) we combine these two approaches using the method recommended by Thompson (2011), which allows for both two way clustering and aggregate serial correlation. Using any of these approaches, the results remain significant at reasonable levels. Overall, to the extent there is spatial correlation present in the data, it enlarges our estimated confidence sets, but those sets still lay far from zero.

C.2.4 Additional compositional channels

Recent work has identified a number of other factors not considered in the main text that are important for business formation. These are skill-biased technical change, changes in educational attainment and immigration.[26] In this section, we analyze whether the estimated effect of the labor supply growth is robust to controlling for these additional changes by exploiting various additional data sources. Overall, our analysis suggests that while some supply composition factors may have also affected firm entry over the period we study, they do not reduce or amplify the quantitative importance of the labor supply growth channel.

Table C.10: Additional labor supply composition channels

	(1)	(2)	(3)	(4)	(5)	(6)
	First Stage			2SLS	2SLS	2SLS
WAP Growth (%)				1.24	1.19	1.23
				(0.20)	(0.22)	(0.19)
College Share of WAP	0.00			0.06		
	(0.02)			(0.03)		
Mincer College Premium		−1.42			0.44	
		(0.78)			(0.64)	
Native-born Share of WAP			0.06			−0.08
			(0.02)			(0.02)
Birthrate IV	1.12	0.11	0.11			
	(0.25)	(2.40)	(1.83)			
Migration IV	0.87	0.88	0.97			
	(0.29)	(0.28)	(0.33)			
N	1,421	1,421	1,078	1,421	1,421	1,078
R^2	0.65	0.66	0.68	0.85	0.86	0.85
F-test	16.91	17.63	22.42	16.91	17.63	22.42
p-value of J test				0.86	0.54	0.84

Education and skill-biased technical change. An important change in our sample period is the rise in skill premium. Rising skill premium might depress entrepreneurship by making payroll employment more desirable than entrepreneurship for college graduates, thereby depressing

[26] See Salgado (2017), Kozeniauskas (2017), Jiang and Sohail (2017) and Burchardi, Chaney, Hassan, Tarquinio, and Terry (2020).

business formation, especially in locations with a higher share of college graduates. We examine whether this channel is a confounding factor for our estimated elasticities by including the share of college graduates and skill premium, both measured at the state-year level using data from the Current Population Survey, as additional controls in Table 8. Column (2) controls for the fraction of the working-age population with a college degree and column (3) controls for the skill premium, obtained by running Mincer regressions separately for each state year combination. The elasticity of the startup rate is essentially unchanged.

Immigration. Since immigrants typically have a higher propensity to form businesses (Kerr and Kerr, 2016), changes in share of native vs. foreign born in the working-age population could affect the start-up rate. We can control for the native born share of each state's working-age population by bringing in additional data from the BLS Current Population Survey (CPS), Survey of Income and Program Participation (SIPP), and the Decennial Census PUMS.[27] We show in column (4) of Table 8 that controlling for the immigrant share has essentially no effect on our results.

The sample is slightly different from the benchmark, because data limitations prevent us from constructing native-born shares for all state years between 1979 and 2007. The CPS includes a question on nativity starting with the 1994 redesign and we use the CPS question to construct native born shares for all states between 1994 and 2007. For the years 1986 to 1993 we are able to calculate nativity shares using the SIPP (omitting 1989, which was never released), and for years 1980 and 1990 we can calculate nativity using Census PUMS. So we use 1980 and 1990 from the Census, 1986-88 and 1991-93 from SIPP, and 1994-2007 from CPS. Some data sources overlap, e.g., the Census PUMS and the CPS and the results are little changed under other permutations.[28]

C.2.5 Effects on incumbent exit and growth

We also estimate equation (10) using several additional margins of firm dynamics as outcomes in place of the startup rate, and we present these results in Table C.11. Column (2), estimates the labor supply growth elasticity for each margin using differences across states in labor supply growth rates predicted only by past fertility. For average startup size (Panel A), the exit rate of young firms (Panel B) and their conditional growth rate (Panel C) the estimates are statistically indistinguishable from zero.[29]

Related to our discussion in Section 3.3 there is an important distinction between using the variation in labor supply growth generated by the instruments and all of the variation in WAP growth across states. The small, or non-existent, effects we find using exogenous shifts in labor supply growth stand in contrast to those estimated using all variation across states (Column 1). While the effect on startup size is still zero, the exit rate for young firms declines and the conditional

[27] Decennial Census PUMS data retrieved from IPUMS (Ruggles, Genadek, Goeken, Grover, and Sobeck, 2017).

[28] In some panels, the SIPP groups low population density states, e.g. Maine is pooled with Vermont. For affected years, we split these grouped states into separate states, assigning each member the group value when estimating and cluster standard errors by the 45 SIPP state groups in the entire sample.

[29] The conditional growth rate refers to the within cohort growth in average firm size (Section 2.1).

Table C.11: Effects of labor supply shocks on additional firm margins

	(1) OLS	(2) IV_1	(3) IV_2	(4) IV_1 & IV_2
Panel A. Average startup employment				
WAP Growth (%)	0.03	−0.27	−0.05	−0.15
	(0.04)	(0.18)	(0.15)	(0.14)
N	1,421	1,421	1,421	1,421
R^2	0.45	0.41	0.45	0.43
p-value of J test				0.20
Panel B. Young firm exit rate (%)				
WAP Growth (%)	−0.37	0.12	−0.21	−0.01
	(0.04)	(0.21)	(0.21)	(0.15)
N	1,029	1,029	1,029	1,029
R^2	0.74	0.68	0.73	0.71
J-test p value				0.34
Panel C. Young firm conditional growth rate (%)				
WAP Growth (%)	0.75	−0.84	−0.11	−0.54
	(0.19)	(0.65)	(1.07)	(0.75)
N	1,029	1,029	1,029	1,029
R^2	0.41	0.33	0.39	0.36
p-value of J test				0.45

Note: Standard errors clustered on state. Young firms (ages 1-10) do not include startups. All regressions use RHS specification and sample from Table 6 with the following additional LHS margins: Panel A is average employment size of startup firms; Panel B is the survival rate in percent of young firms; Panel C is the young firm within-cohort percent growth in average firm size. Panels B and C of young incumbents use years 1987 to 2007, because of birth year censoring.

growth rate increases. With OLS (Column 1) we find increases in WAP growth are associated with no difference in average startup employment (Panel A), but statistically significant declines in the exit rate of young firms (Panel B) and increase in the conditional growth rate of young firms (Panel C). In contrast, when just using variation induced by the labor supply growth instruments, these effects mostly vanish.

Since OLS uses variation in labor market growth that may also be determined by local economic conditions, these results are unsurprising. If labor markets are growing due to increases in local profitability, which directly affects the value of incumbent firms, incumbent firms would be less likely to exit (Panel B) and more likely to expand (Panel C). These patterns also help explain the consistently smaller elasticity of the startup rate estimated via OLS found throughout Tables 6 and 8; the additional increases in incumbent labor demand from less exit and more growth require a smaller equilibrium response along the entry margin.

C.3 CBP Startup Rate imputations

Finally, we consider our CBP-based establishment startup rate imputation against some alternatives. First, we show that the trend component of the CBP imputation and the actual establishment startup rate measured in the BDS track each other closely. (See Table C.14, which is the smoothed version of Figure 9 from the main text.)

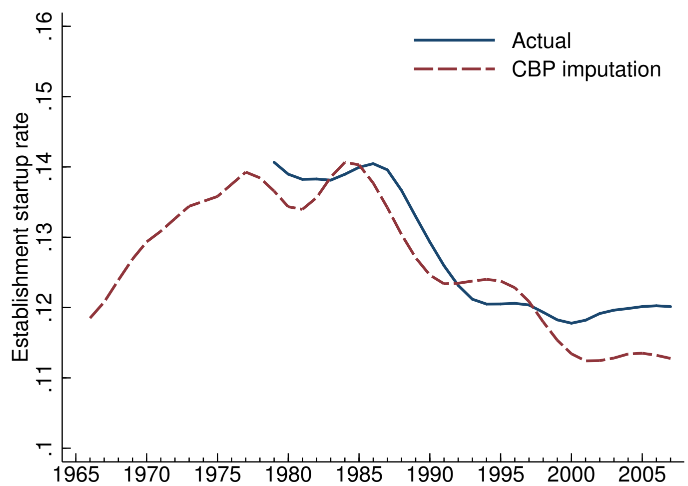


Figure C.14: Comparing trend component of actual and CBP imputed startup rates

Next, we consider the effects of alternative assumptions in the statistical model to predict the aggregate exit rate. In Section 5, we estimate for each state and establishment size category within the BDS the following regression for exit

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

The linear time trend coefficient λ^{sj} captures the slow movement in exit within size group because of implied changes in the age distribution. We allow this coefficient to differ for each state and size group. When predicting exit rates out of sample, we have to choose whether or not to extrapolate the time trend. In the main text, we keep the trend term fixed at its 1979 level when estimating exit prior to 1979. In Figure C.15, we consider several alternative choices: a symmetric trend, a continuuing trend, or holding exit fixed at its in-sample average. For each, we predict exit by state and size group and year $\hat{\delta}_t^{sj}$, and then compute the imputed CBP startup rate using equation (13).

The solid line plots the actual establishment startup rate measured in the BDS for the years 1979 to 2007. The broken red line plots the imputation used in the paper. As an alternative, the orange and gray lines consider imposing a symmetric trend pre 1979 and continuuing the in-sample time trend, respectively. By construction they are the same for the 1979-2007 period. As a final alternative, we eliminate the time trend entirely and just estimate exit using the average exit rate over the entire period. Regardless of the assumption used to predict exit, the hump shaped pattern in the CBP-based imputed establishment startup rate remains the same.

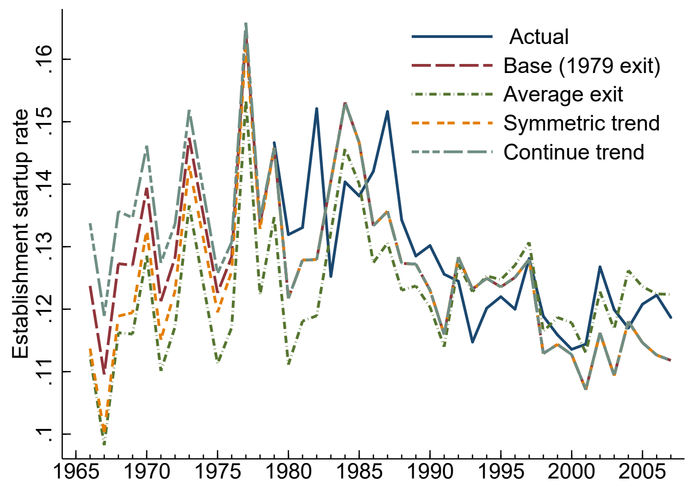


Figure C.15: Imputed startup rate under alternative extrapolation of exit rate by size groups

As a further refinement, we add an aggregate state variable to the exit regression

$$x_t^{sj} = \bar{x}^{sj} + \lambda^{sj} t + \beta^{sj} Z_t + \varepsilon_t^{sj}.$$

Here as a business cycle indicator, we use annual real GDP growth. Figure C.16 plots the same set of alternatives where the exit rate prediction also includes any predicted business cycle fluctuations using state variable Z_t. The results are very similar and also feature the hump shaped pattern.

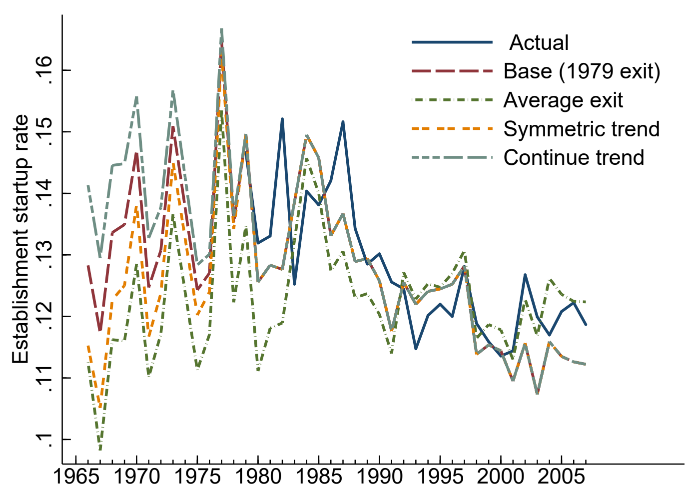


Figure C.16: Imputed startup rate under alternative extrapolation of exit rate with cyclical adjustment

References

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio, and S. J. Terry (2020): "Immigration, Innovation, and Growth," Discussion paper, National Bureau of Economic Research.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust inference with multiway clustering," *Journal of Business & Economic Statistics*, 29(2), 238–249.

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

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

——— (2014b): "The Secular Decline in Business Dynamism in the U.S.," mimeo.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016): "Where has all the skewness gone? The decline in high-growth (young) firms in the US," *European Economic Review*, 86, 4–23.

——— (2020): "Changing business dynamism and productivity: Shocks versus responsiveness," *American Economic Review*, 110(12), 3952–90.

Dent, R. C., F. Karahan, B. Pugsley, and A. Şahin (2016): "The role of startups in structural transformation," *American Economic Review*, 106(5), 219–23.

Driscoll, J. C., and A. C. Kraay (1998): "Consistent covariance matrix estimation with spatially dependent panel data," *Review of economics and statistics*, 80(4), 549–560.

Duarte, M., and D. Restuccia (2010): "The role of the structural transformation in aggregate productivity," *The Quarterly Journal of Economics*, 125(1), 129–173.

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

Foote, C. L. (2007): "Space and Time in Macroeconomic Panel Data: Young Workers and State-Level Unemployment Revisited," Working Paper WP 07-10, Federal Reserve Bank of Boston.

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

Gutiérrez, G., C. Jones, and T. Philippon (2019): "Entry costs and the macroeconomy," Discussion paper, IMF Working Paper 19-233.

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

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

Jiang, H., and F. Sohail (2017): "Skill biased entrepreneurial decline," Discussion paper, Working paper, September 2017. 5.

Kerr, S. P., and W. R. Kerr (2016): "Immigrant entrepreneurship," in *Measuring entrepreneurial businesses: Current knowledge and challenges*, pp. 187–249. University of Chicago Press.

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

Lucas, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, pp. 508–523.

Pugsley, B. W., and A. Şahin (2019): "Grown-up Business Cycles," *The Review of Financial Studies*, 32(3), 1102–1147.

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

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

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

Thompson, S. B. (2011): "Simple formulas for standard errors that cluster by both firm and time," *Journal of financial Economics*, 99(1), 1–10.