The Growth Potential of Startups over the Business Cycle[*]

Petr Sedláček

University of Bonn [†]

Vincent Sterk

University College London [‡]

January 6, 2014

Abstract

This paper shows that job creation of cohorts of U.S. firms is strongly influenced by aggregate conditions at the time of their entry. Using data from the Business Dynamics Statistics (BDS) we follow cohorts of young firms and document that their employment levels are very persistent and largely driven by the intensive margin (average firm size) rather than the extensive margin (number of firms). To differentiate changes in the composition of startup cohorts from post-entry choices and to evaluate aggregate effects, we estimate a general equilibrium firm dynamics model using BDS data. We find that even for older firms, the aggregate state at birth drives the vast majority of variations in employment across cohorts of the same age. The key force behind this result are fluctuations in the composition of startup cohorts with respect to firms' potential to grow large. At the aggregate level, factors determined at the startup phase account for the large low-frequency fluctuations observed in the employment rate.

Keywords: Firm Dynamics, Heterogeneous Agents, Maximum Likelihood, DSGE
JEL Codes: E32, D22, L11, M13

1   Introduction

Following the financial crisis of 2008, the United States and many other countries experienced an unusually deep and prolonged economic downturn, raising concerns about a long-lasting drag on aggregate employment and output. These concerns are fueled by the observation of a particularly large drop in job creation by young firms in the U.S. in recent years. According to data from the Business Dynamics Statistics (BDS), employment by firms up to the age of five fell by 4.2 million between 2006 and 2010, accounting for more than half of the decline in aggregate employment.[1]

While it is broadly recognized that young firms play a pivotal role in aggregate job creation, this paper investigates the importance of aggregate economic conditions at birth for determining firms' ability to provide employment throughout their lives. One channel through which startup conditions can have persistent effects are cyclical changes in the number of newborn firms. A second possible channel, emphasized in this paper, operates through cyclical fluctuations in the *composition* of startup cohorts with respect to firms' potential to grow large. Using the BDS data, we document new stylized facts concerning job creation of newborn firms. The empirical evidence is complemented by a general equilibrium firm dynamics model, estimated using cohort-level and aggregate data.

We establish three stylized facts. First, fluctuations in employment by entrants are large and pro-cyclical. Second, cyclical deviations in the level of employment created by newborn cohorts persist as cohort age, sharply contrasting the strong mean-reversion in aggregate employment. Third, most of the variation in cohort-level employment is driven by the intensive margin (average firm size), implying a relatively modest role for the extensive margin (the number of firms).

Although the stylized facts are very suggestive, they do not precisely answer how aggregate conditions at birth influence cohort-level outcomes, nor do they allow us to evaluate the aggregate implications of fluctuations across cohorts of entrants. A first-order challenge in addressing these issues is to quantify the extent to which employment variations are driven by changes in the composition of entrants, which are likely to persist, or by post-entry decisions shaped by current business cycle conditions, which are more easily reversed. To disentangle these forces, we build a general equilibrium firm dynamics model with heterogeneous firms, endogenous entry, growth subject to adjustment costs

---

[1]By contrast, the share of firms up to five year of age in the *level* of aggregate employment is much smaller: 14 percent in 2006 and 12 percent in 2010.

and aggregate uncertainty. Before entry, firms choose a production technology affecting their returns to scale and therefore their size after birth. Using the model, we *estimate* the parameters and realizations of the underlying shocks.

Our estimates imply that at least 90 percent of variation in employment across cohorts of a given age is driven by conditions in the year of birth, even for older firms. The importance of the birth stage for long-term outcomes is largely driven by variations in the composition of cohorts with respect to firms' technologies chosen at the entry stage. Initially, such compositional variations have only a moderate effect since all firms startup relatively small. As the cohort ages, however, a fraction of the firms grows large and the effect of composition on cohort-level employment becomes increasingly pronounced.

Next, we use the model to show that decisions made at the entry phase help to understand aggregate fluctuations. We find that over our sample period, the contribution of startup conditions to aggregate employment fluctuations evolves strikingly similar to the *trend* component of the employment rate, which exhibits large variations but is often discarded in business cycle analysis. Our model thus complements more standard business cycle models in understanding the drivers of overall variations in aggregate employment. Finally, we show that general equilibrium effects associated with fluctuations in entry decisions are strong, particularly in the long run. Taking into account such effects is important for any government policy that aims to increase employment by intervening at the firm entry stage.

The model incorporates two novel features in order to suit our empirical focus, which is on cohort-level variables for the aggregate economy. First, we model heterogeneity in returns-to-scale in production. This is in contrast to standard firm dynamics models in which heterogeneity is modeled through variations in the degree in total factor productivity across firms. A motivating factor behind our choice is that we apply our model to the entire cross-section of private employers in the economy rather than confined industries, in which returns to scale are arguably more homogeneous.[2] Given the enormous differences in the type of activities performed by firms, there is arguably large variation in the degree to which various types of businesses are scalable.[3] Moreover, only a small degree of heterogeneity in returns to scale in our model is sufficient to fit the size distribution of

---

[2]Even so, Holmes and Stevens (2012) provide evidence of substantial heterogeneity in returns to scale even within narrowly defined industries, and build a model with large-scale standardized plants and small-scale specialty plants.

[3]Basu and Fernald (1997) provide evidence in favor of heterogeneity in returns to scale across sectors.

firms in the data, conditional on age.[4]

The second novel feature of our framework is the modeling of the firm entry phase. In our model, potential entrants choose the technology of their businesses (i.e. their scalability) after paying an entry cost. However, a coordination friction prevents all startups from being successful and some aspiring startups are forced to exit immediately. The probabilities of successfully starting up a business of a certain type endogenously adjust such that potential entrants are indifferent between technology types, akin to the directed search literature (see e.g. Moen, 1997). As a result, all aggregate shocks affect the composition of startups through their differential impact on the respective firm values influencing the incentives to startup a particular type of business. Moreover, the model predicts entry in all technology types enhancing the model's ability to fit the data which includes many old small firms.[5] This feature of our model is also consistent with empirical evidence that many starting entrepreneurs have low growth expectations, see Campbell and De Nardi (2009) and Hurst and Pugsley (2011).

An important element of the analysis is the estimation of our model. This is a complex problem given the presence of a large aggregate state, which includes distributions of firm-specific variables, giving rise to rich endogenous propagation of shocks. We overcome this hurdle using a novel computational strategy that allows us to solve the model fast using standard techniques.[6] We estimate the parameters and the realizations of the aggregate shocks using a combination of aggregate and cohort-level data. The structure of our model, combined with the use of cohort-level data, enables us to sharply identify changes in the composition of startups, a key feature of our framework.

The empirical findings presented in this paper complement those by Haltiwanger, Jarmin, and Miranda (2013) and Fort, Haltiwanger, Jarmin, and Miranda (2012), who document the importance of young firms for average aggregate job creation. They are also

---

[4]Nonetheless, our general framework would allow TFP to be added. In typical firm dynamics models introducing more than one dimension of heterogeneity is more challenging, since solutions in these models are described by cutoff rules.

[5]In 2007, the fraction of firms with 10 or less employees among firms between 21 and 25 years of age was about two thirds.

[6]Our model can be solved in one step using first-order perturbation, as opposed to most models with a large number of agents, which are typically solved using the computationally demanding algorithm of Krusell and Smith (1998). An exception is Campbell (1998), who pioneered the use of perturbation methods to solve heterogeneous-agents models, replacing functional equations with quadrature approximations. Our setup avoids the need for such approximations because the economic state is large but finite-dimensional, preserving exact aggregation.

related to Bartelsman, Haltiwanger, and Scarpetta (2009) who in a cross-country study investigate *average* post-entry behavior of young firms and to Lee and Mukoyama (2012) who document productivity and size patterns of entering and exiting firms in the manufacturing sector. Unlike the above studies, our paper investigates how job creation of firm cohorts relates to the business cycle at the time of their birth using data which covers the entire U.S. economy. Earlier studies using BDS data include Moscarini and Postel-Vinay (2012) and Fort, Haltiwanger, Jarmin, and Miranda (2012) who study the cyclicality of large versus small firms but do not focus on entrants.

Our structural model builds on a rich literature studying the dynamics of firms and variations in firm size. Lucas (1978) explains observed variations in firm size using a model with managers who differ in their "span of control", that is the ability in running large firms. Firm dynamics are introduced in Jovanovic (1982) who develops a model in which new firms grow faster and are more likely to fail compared to older firms as they learn about their efficient scales of operation. A workhorse firm dynamics model is presented in Hopenhayn and Rogerson (1993) who analyze the welfare effects of firing taxes in a general equilibrium (without aggregate uncertainty). Abbring and Campbell (2004) estimate a firm dynamics model using sales data on newly opened bars in Texas and find that pre-entry scale decisions are an important source of variation in sales across firms. Campbell (1998) incorporates vintage technology shocks in a firm dynamics model. More recently, Lee and Mukoyama (2012) and Clementi and Palazzo (2010) use the Hopenhayn-Rogerson framework to study how entry and exit decisions propagate shocks and apply their models to the manufacturing industry. In contrast to these papers, we focus on aggregate outcomes and incorporate multiple sources of aggregate uncertainty and estimate the realizations of the continuous-support shocks, as well as parameters pertaining to their laws of motion.

The organization of the remainder of this paper is as follows. Section 2 describes the data and presents empirical stylized facts. The model and its parametrization are described in Sections 3 and 4, respectively. Section 5 presents the model results. Concluding remarks are made in Section 6.

2 Empirical evidence

This section documents broad patterns in job creation by young U.S. firms over time, without imposing any model structure. Our units of analysis are cohorts, that is, aggregates over firms born in the same year. We use data from the Business Dynamic Statistics

(BDS), described in detail in Subsection 2.1. The BDS data enable us to observe job creation in the year of birth as well as during the five years after. Subsection 2.2 analyzes cyclical fluctuations in job creation by entrants and the extent to which these fluctuations persist as a cohort ages. The BDS also reports the number of firms in each cohort, which allows for a breakdown of cohort-level employment into the extensive margin (the number of firms) and the intensive margin (average firm size). We investigate the relative importance of the two margins in Subsection 2.3. Finally, subsection 2.4 takes stock of the empirical findings and discusses several possible explanations.

2.1 Data and definitions

The BDS database covers a very large fraction of US private employment (98 percent), which is an important advantage over alternative data sources, especially given our objective to study implications for aggregate outcomes. We use the available annual information on the number of firms and their associated job flows broken down into age categories, for the period 1979 until 2011.[7] The available age breakdown in the BDS allows one to follow cohorts of new firms for up to five years after they enter the economy. The BDS groups older firms into age categories spanning five years, preventing further following of an individual cohort. Nevertheless, the five year cutoff strikes a reasonable compromise between the length of a cohort and the number of cohorts available. Even though the main text refers to firms, Appendix A shows that our results remain to hold also for establishments.[8]

We introduce the following notation. Let $M_{a,t}$ be the number of firms in a cohort of age a in year t. Following the BDS notation, startups enter with age $a = 0$.[9] Similarly, let $N_{a,t}$ be the employment level of a cohort of firms of age a in year t. The employment level of a given cohort is measured as the cumulative net job creation since birth, i.e. $N_{a,t} = \sum_{i=0}^{a} NJC_{i,t-a+i}$, where $NJC_{a,t}$ is the net number of jobs created in firms of age a in year t.[10]

---

[7]The BDS data start in 1977 but we drop the initial two years following Moscarini and Postel-Vinay (2012), who cast doubt on the quality of the initial two years of data

[8]An establishment is defined as a single physical location where business is conducted or where services or industrial operations are performed. A firm is a business organization consisting of one or more establishments that were specified under common ownership or control. Therefore, the firm and the establishment are the same for single-establishment firms, but existing firms can create new establishments.

[9]A new firm is defined as a firm having a positive employment entry in March of year t, while not having an employment entry in March of $t-1$.

[10]Alternatively, one could use the employment stock data presented in the BDS. These employment

2.2 Job creation by newborn firm cohorts

Startups are widely recognized to be important drivers of long-run job aggregate creation, at least since Haltiwanger, Jarmin, and Miranda (2013). On average, startups created about 2.8 million jobs over our sample period, which is nearly 170 percent of aggregate *net* job creation. The strong decline in job creation by young firms during the recent downturn suggests that startups may play an important role in the business cycle as well. This subsection establishes that fluctuations in job creation by entrants are large and co-move positively with the aggregate business cycle. Also, we document that compared to aggregate employment, cohort-level employment is very persistent.

To visualize these facts, Figure 1 displays the employment levels for cohorts of startups born between 1979 and 2011 together with their employment levels five years later. The figure also plots aggregate employment, linearly de-trended. Several patterns stand out. First, note that fluctuations in cohort-level employment are large. Compared to the aggregate employment rate, the volatility of employment by entrants is 4.35 as large. Also, the cohort-level volatility does not appear to diminish with age. Second, job creation by entrants and aggregate employment move together and drop during recession years, indicated by shaded areas. Finally, there is strong positive co-movement between the employment levels of cohorts at birth and five years later, indicating substantial persistence in cohort-level employment.

2.2.1 Startups and the business cycle

To investigate the cyclical properties of cohort-level job creation at birth more closely, we correlate employment in startups with several measures of the business cycle. We use linearly de-trended employment rate and real GDP data as our main business cycle indicators.[11] Nevertheless, we also report results using the HP filter as well as the level of the

numbers do not equal the sum of net job creation, which is because the net job creation data is cleaned from observed entrants that are not believed to be true startups, while the employment data is not. BDS documentation states that: "...it may be determined that an establishment's entry/exit as shown by the data is not credible. These establishments are excluded from the change calculations in a given year" (http://www.census.gov/ces/dataproducts/bds). Thus, the net job creation data are superior, at least for our purpose.

[11]The employment rate is defined as 1 minus the unemployment rate taken from the BLS. Both the employment rate and real GDP are averages over March-to-March periods, consistent with the BDS timing. All variables are logged prior to de-trending. Because our analysis deals with time series of different lengths (e.g. information on five year old firms starts only in 1984), we always de-trend the given data over the

Figure 1: Total employment of firm cohorts at age 0 and 5 by year of birth, and aggregate employment by year
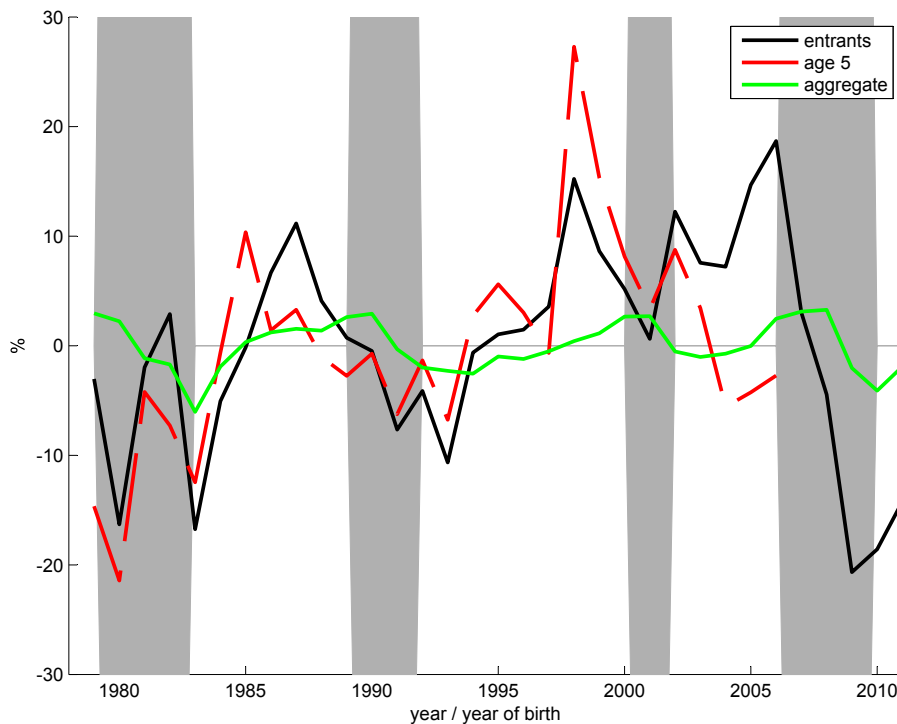
Notes: cohort-deviations plotted in percentage deviations from sample mean. Aggregate employment plotted in percentage deviation from linear trend. Shaded areas are the NBER recessions.
Source: BDS.

Table 1: Correlations of entrant job creation with business cycle indicators

	linear trend		HP filter		levels	
	e-rate	GDP	e-rate	GDP	e-rate	GDP gr.
JC entrants	0.59	0.68	0.31	0.39	0.59	0.47

Notes: The table reports correlation coefficients between the variables in the columns and job creation by firm and establishment entrants. "e-rate" stands for employment rate and is defined as 1 minus the unemployment rate. "GDP gr." is the growth rate of real GDP. Except for the case of levels, all variables are logged prior to detrending.

employment rate and the growth rate in real GDP.[12] Table 1 reports the resulting correlations and confirms that job creation by newborn cohorts is pro-cyclical.[13] This relationship holds robustly over different business cycle indicators and de-trending methods.

2.2.2 Persistence of cohort-level job creation

To quantify the persistence of cohort-level job creation, we correlate cohort-level employment in year t with that in year $t + a$ for the same cohort. Figure 2 plots the correlation coefficients for ages $a = 1$ to $a = 5$ both for the individual cohorts, as well as the corresponding correlation for aggregate employment.[14] While cohort-level employment at birth and 5 years into existence are highly correlated, with correlation coefficient of 0.68, its aggregate counterpart displays no persistence after two to three years. Thus, deviations in job creation by individual startup cohorts persist as the cohorts age, to a degree that far exceeds the persistence inherent to the aggregate business cycle.

2.3 Intensive versus extensive margin of job creation

Having established that recession-born cohorts create fewer jobs and that this persists into later years of their existence, we now investigate whether the observed variation of cohort-level employment is driven primarily by a persistently lower number of firms within the cohort, or by a persistently lower growth potential of the firms that do enter in recessions. To this end, we first decompose the observed variation of employment of individual cohorts into the contributions of the extensive margin (the number of firms) and the intensive margin (average firm size). Moreover, we break both margins down by age to separate out the contributions of the year of birth and subsequent years.

---

longest possible sample with the earliest starting point of 1979.

[12]Throughout the paper, the smoothing parameter in the HP filter is set to 100. However, using the value of 6.23 suggested by Ravn and Uhlig (2002) delivers similar results.

[13]All correlation coefficients above 0.30 are statistically significant.

[14]While cohort-level employment does not display a trend, aggregate employment does and therefore we choose to linearly detrend both time series. Using alternative detrending methods gives similar results. For the aggregate we simply correlate period t employment with employment in years $t, t + 1, ..., t + 5$.

Figure 2: Autocorrelations
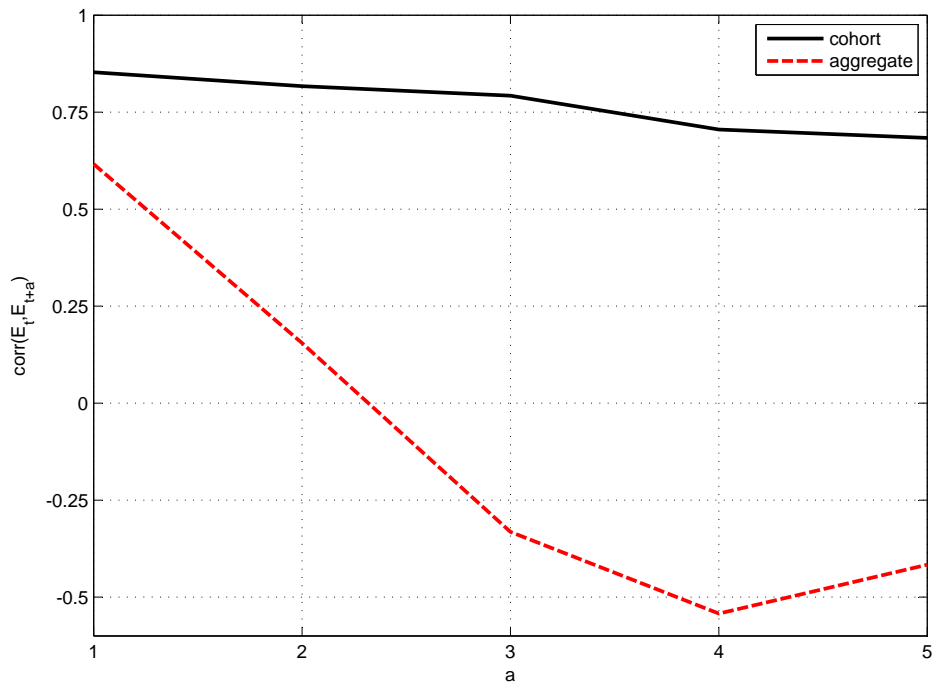
Notes: Correlation coefficients of employment in year $t = 0$ and in year $t + age$, with $age = 1, 2, 3, 4, 5$, at both the level of a cohort born in period $t = 0$ and at the aggregate level.
Source: BDS.

2.3.1 Variance decomposition

To quantify the relative importance of the two margins, we decompose the natural logarithm of cohort-level employment as:

$$\ln N_{a,t} = \ln S_{0,t-a} + \ln M_{0,t-a} + \sum_{j=1}^{a} \ln \gamma_{j,t-a+j} + \sum_{j=1}^{a} \ln \delta_{j,t-a+j},$$

where $S_{a,t}$ is average firm size within the cohort, $\gamma_{j,t} \equiv \frac{S_{j,t}}{S_{j-1,t-1}}$ denotes average size growth and $\delta_{j,t} \equiv \frac{M_{j,t}}{M_{j-1,t-1}}$ denotes average firm survival rate. Based on the above expression, the variance of employment can be decomposed as:

$$Var\left(\hat{N}_{a,t}\right) = Cov(\hat{N}_{a,t}, \hat{S}_{0,t-a}) + Cov(\hat{N}_{a,t}, \hat{M}_{0,t-a}) + \sum_{j=1}^{a} Cov(\hat{N}_{a,t}, \hat{\gamma}_{j,t-a+j})$$

$$+ \sum_{j=1}^{a} Cov(\hat{N}_{a,t}, \hat{\delta}_{j,t-a+j}),$$

where a hat indicates deviations from a linear trend of a logged variable.[15]

Figure 3 plots the contributions of average firm size and the number of firms to employment levels of cohorts at age five, expressed as a percentage of the total variance. Average firm size accounts for the majority of the variation in cohort-level employment of five year old firms: about 69 percent. Within the part accounted for by average size, more than a third is due to entrant size. Moreover, entrant size and firm growth in the year after birth jointly account for than 40 percent of the overall variation in employment. Thus, average firm size at the early stage of a cohort's existence emerges as a key determinant of a cohort's success in providing jobs later in life.

2.3.2 Potential importance of composition

The variance composition reveals that entrant size plays a prominent role in accounting for cohort-level employment in later years. But what generates variations in entrant size in the first place? One possibility is that entrant size is driven primarily by fluctuations in the type of firms that enter. Alternatively, average size variations may reflect post-entry decisions, for a given mix of firm types (or a combination of the two). The structural model, laid out in the next section, will be used to *estimate* the contribution of pre- and

---

[15]This decomposition is convenient because it exactly sums to 1 for linearly detrended data, see Fujita and Ramey (2009) for a more detailed description.

Figure 3: Variance decomposition
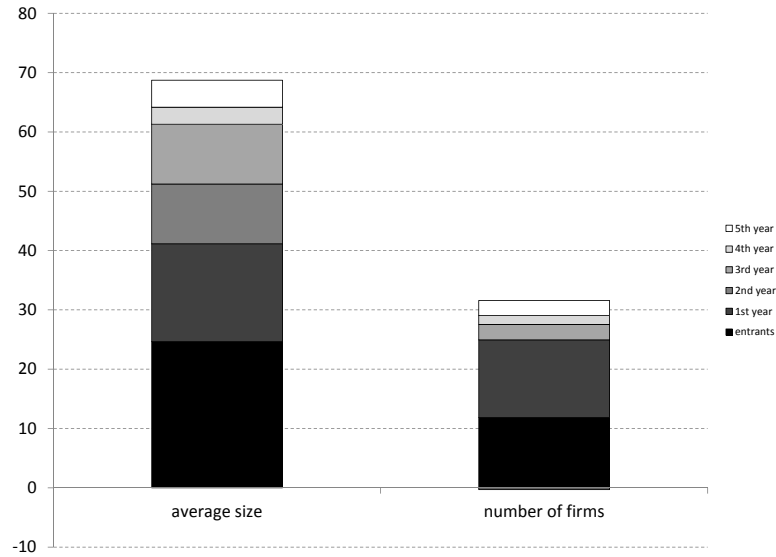
Notes: The figure plots contributions of average firm size and the number of firms to the variation in cohort-level employment of five year old firms expressed as percent of the total variation.

Source: BDS and authors' calculations.

post-entry decisions. Nevertheless, we can get a glimpse of their relative importance by conducting a simple exercise which exploits that BDS data can be broken down into size categories.[16]

We compute two counterfactual time series for average entrant size, plotted in Figure 4. The first counterfactual is calculated by fixing the distribution of the number of firms over the size brackets and letting only average firm size within each size bracket vary. Conversely, the second counterfactual fixes average firm size and lets only the distribution of the number of firms over the size brackets vary. The two counterfactuals reveal that shifts *between* size brackets drive the majority of fluctuations in average size. With size-variation within brackets shut off, the counterfactual time series of average entrant size is very close to the actual time series, suggesting that composition effects are very important. But clearly, we need to take this exercise with a grain of salt since shifts between size brackets are mere proxies for true fluctuations in underlying firm heterogeneity.[17]

---

[16]The smallest size category, firms with 1-4 employees, accounts for about a third of employment. The remaining categories individually account for less than 20 percent.

[17]In particular, the figure could be overstating the effect of composition since fluctuations in firms' post-entry decisions may cause them to be reclassified into a different size bracket. Nonetheless, using wider

Figure 4: Entrant size: potential importance of composition
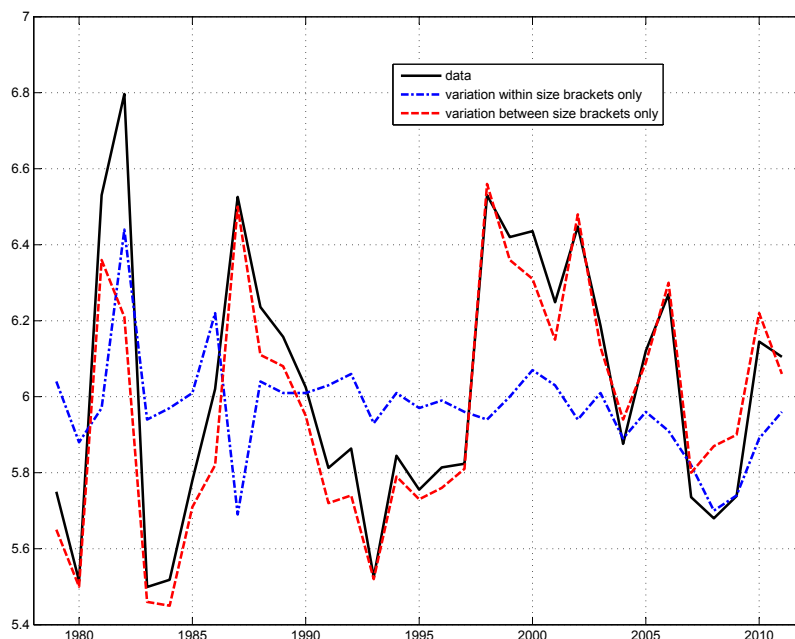
Notes: The figure plots entrant size ("actual") and two counterfactuals. "Variations within size brackets only" assumes that the firm shares within the BDS size categories are fixed and only average size within these categories varies. "Variations between size brackets only" assumes that the composition of firms varies but the average sizes within the BDS size groups are fixed.
Source: BDS and authors' calculations.

2.4 Summary and possible explanations

The findings in the previous subsections can be condensed into three stylized facts:

Fact 1. *Entrant job creation is volatile and pro-cyclical.*

Fact 2. *Cohort-level employment is largely determined in the year of birth.*

Fact 3. *The intensive margin (average size) is the main driver of variations in cohort-level employment.*

These stylized facts are difficult to reconcile with the view that cohort-level employment at a given point in time is primarily driven by the *current* state of the business cycle.

size brackets in the counterfactual, which should alleviate this issue, yields similar results.

Instead, the observed patterns lend support to a view in which firm characteristics *at the entry stage* are important in determining a cohort's potency to create jobs, both initially and later in its life. It is perhaps not surprising that recession-born cohorts create fewer jobs because fewer firms start up during downturns. However, our results point to another channel, namely that recession cohorts consist of smaller firms on average. Moreover, this second channel seems to be the dominant force in determining the variation of cohort-level employment in later years of their existence. While a subsequent economic recovery may contribute to average size growth at the cohort level, the upward effect does not seem to be strong enough to offset the lower average size that recession cohorts are born with.

One can think of several plausible explanations for why the composition of entrants may fluctuate over the business cycle. In this subsection we discuss two candidate explanations for our stylized facts.

2.4.1 Sectoral reallocation

One possibility is that during recessions job creation within newborn cohorts declines because of a reallocation of activity between sectors. Appendix B, however, documents that our stylized facts continue to hold, with a few exceptions, also within the broad sectors defined in the BDS.[18]

2.4.2 Very small entrants

Another possibility is that our findings are driven by fluctuations in entry of very small firms. Several studies emphasize the role of entrepreneurship as a way to escape unemployment ("necessity entrepreneurs"), e.g. Hurst and Pugsley (2011) and Poschke (2012). Businesses created out of a "necessity" motive are likely to remain very small. Given that unemployment is high during recessions, one may expect necessity entrepreneurship to have a negative effect on job creation of firms born in recessions.

To investigate the importance of small firms we construct a counterfactual time series for the total employment of cohorts at age five, but set the employment levels of firms with less than 10 employees equal to the sample average. Similarly, we construct a second counterfactual time series by setting the employment of firms with 10 employees or more

---

[18]The exceptions are the mining sector and transportation, communication, and public utilities in which entrant job creation is counter- and a-cyclical, respectively. Also, the extensive margin dominates employment variations in retail trade and construction.

Figure 5: Counterfactual: necessity entrepreneurs
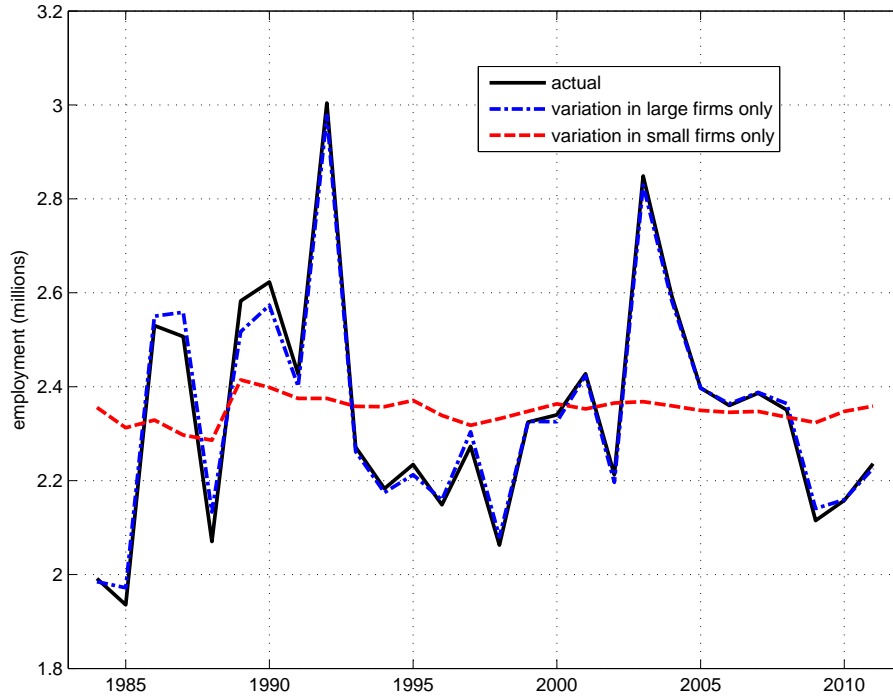
Notes: The figure plots employment of five year old firms for the data ("actual") and for two counterfactuals. "Variation in small firms only" assumes that employment is fixed to its sample average for firms with 10 or more employees. "Variation in large firms only" assumes that employment is fixed to its sample average for firms with less than 10 employees.

equal to the sample average. Figure 5 shows that the vast majority of fluctuations in employment of five year old firms is in fact driven by firms with 10 or more employees. This observation does not of course refute the existence of necessity entrepreneurship, but it appears unlikely that cyclical variations in this entrepreneurship motive are driving our stylized facts.

3 The model

The empirical evidence presented in the previous section suggests that fluctuations in the composition of startup cohorts are important for cohort-level employment in later years. However, the data does not allow us to quantify the importance of composition fluctuations, since we do not directly observe the distribution of cohorts with respect to

firm types. For the same reason, the empirical facts can provide only limited information on the aggregate implications of decisions made at the entry stage. Moreover, this may be especially relevant at longer horizons, beyond the age of five, as the cohorts mature.

To address the limitations of our reduced-form analysis, this section proposes a general equilibrium model of firm dynamics in which startups can choose the technology type affecting the scalability of their aspired businesses. The model further features endogenous firm entry, labor adjustment costs and several sources of aggregate uncertainty. To quantify the contribution of composition effects and post-entry employment choices for the evolution of cohort-level employment in later years, we estimate the aggregate shocks using information about average size from the BDS. While post-entry shocks have a transitory impact on average firm size, the impact of changes in the composition of startups increases as the cohort ages. Therefore, information about average firm size at birth and at the age of five enables us to sharply identify fluctuations in the composition of cohorts of entrants.

The estimation procedure also delivers predicted values for the full economic state of the model in each sample period, which includes the distribution of firms with respect to age and types. The knowledge of this distribution enables us to quantify what fraction of the fluctuations in the aggregate employment rate is accounted for by decisions made at the entry stage. We thus actively use the large aggregate state of our model in our quantitative analysis.

Our model is designed for an application to cohort-level and aggregate data, rather than for analyzing individual firms. We therefore abstract from firm-specific technology shocks and associated endogenous exit after entry. Instead, we model an age-dependent exit rate calibrated using BDS data. In Appendix D we introduce stochastic time-variation in exit rates and show that our main results or not substantially affected.

The model economy is populated by an infinitely-lived representative household and a continuum of heterogeneous firms with finite lives. All agents have rational expectations. Firms and households trade on a goods and a labor market, both of which are perfectly competitive. Firm dynamics models with more detailed descriptions of the labor market include Acemoglu and Hawkins (2010), Elsby and Michaels (2013), Kaas and Kircher (2011), Moscarini and Postel-Vinay (2010), Schaal (2010) and Sedláček (2013) who extend the Diamond-Mortensen-Pissarides model to include multi-worker firms.

3.1 Existing firms

There is an endogenous mass of heterogeneous firms which produce a homogeneous good. Before describing the entry decision, which is key to our analysis, we lay out how incumbent firms behave. While all firms in the economy use labor as the only factor of production, the production technology itself differs across firms. In particular, there is a finite number of technology types, indexed by $i = 1, 2, ..., I$. Existing firms grow only gradually towards their optimal sizes due to costs related to adjusting their employment levels. Finally, firms face an exogenous, age-dependent, probability of shutting down, ρ_a. By symmetry, all firms of the same type and age make the same decisions and we therefore index them only by technology type and age.

Technology types differ in the degrees of returns to scale. In particular, a firm with technology type i is characterized by the following production function

$$y(n_t; i) = A_t n_t^{\alpha_i},$$

where A_t is an exogenous and stochastic aggregate TFP variable with mean one, n_t is the firm's level of employment, and α_i is a technology-specific returns to scale parameter.[19] In our quantitative simulations we parameterize $\alpha_i \in (0,1)$ for each technology type i, i.e. returns to scale are decreasing. As a result, there exists a type-specific "optimal size" beyond which further growth is undesirable.

Firms choose their employment level in order to maximize their discounted stream of profits

$$V_{i,a}(n_{i,a-1,t-1}, \mathcal{F}_t) = \max_{n_{i,a,t}} \left[\begin{array}{c} y_{i,a,t} - W_t n_{i,a,t} - \frac{\varsigma Q_t}{2}(n_{i,a,t} - n_{i,a-1,t-1})^2 \\ + (1 - \rho_a) \mathbb{E}_t \Lambda_{t,t+1} V_{i,a+1}(n_{i,a,t}, \mathcal{F}_{t+1}) \end{array} \right], \qquad (1)$$

where $V_{i,a}(n_{i,a-1,t-1}, \mathcal{F}_t)$ is the asset value of a firm of type i and age a, $\mathbb{E}_t$ is the conditional expectations operator and $\mathcal{F}_t$ is the aggregate state to be described later. W_t is the economy-wide wage rate, $\Lambda_{t,t+1}$ is the firm's stochastic discount factor between period t and $t+1$ and $\frac{\varsigma Q_t}{2}(n_t - n_{t-1})^2$ is the adjustment cost of changing employment with Q_t being a stochastic shock with mean one.[20] Since firms in our model are typically on an upward growth path, we label Q_t an "expansion cost shock". Given that firm expansion is a form of investment in our model, Q_t resembles the investment-specific technology shock that

---

[19]For computational feasibility we assume firm-level employment is a continuous variable.

[20]It is assumed that firms that shut down do not pay adjustment costs and that new-born firms have an initial employment level of zero.

features prominently in the DSGE literature and is sometimes thought of as a stand-in for time-varying financial frictions.

The first-order necessary condition for the firm's optimal choice of labor can be written as

$$W_t + \zeta Q_t(n_{i,a,t} - n_{i,a-1,t-1}) = \alpha_i \frac{y_{i,a,t}}{n_{i,a,t}} + \beta \Lambda_{t,t+1}(1 - \rho_a)\zeta \mathbb{E}_t Q_{t+1}(n_{i,a+1,t+1} - n_{i,a,t}), \quad (2)$$

which equates the marginal costs of firm expansion to the marginal benefits. Marginal costs consist of the wage and the marginal adjustment cost. Marginal benefits equal the sum of the marginal product of labor and the expected discounted marginal reduction in adjustment costs to be paid next period.

3.2   Entry decisions

Let us now turn to the decision of potential firms wishing to start up a business. Starting up a firm requires the sacrifice of a cost $\chi > 0$. After paying this cost, a potential entrant chooses one business idea, associated with a technology type, from a given measure of opportunities. While the total measure of business opportunities, denoted by $\Psi > 0$, is assumed to be constant, its composition with respect to technology types varies stochastically over time. This happens according to an exogenous "composition" shock X_t, the interpretation of which is discussed below.[21] However, not all startup attempts are successful and not all business ideas are realized because of a *coordination* friction. The underlying idea is that without perfect coordination, some potential entrants select the same business opportunity, forcing all but one to exit directly. At the same time, some business opportunities are not selected. Our modeling approach thus follows search and matching models, and in particular directed search models, which are routinely applied to the labor market. Therefore, starting up a firm with a given technology type happens only with a certain (endogenous) probability. A free entry condition then ensures that potential firm entrants are indifferent between technology types.

Specifically, the number of startups in a given technology type is determined by a "matching" function. Let $e_{i,t}$ denote the measure of aspiring entrants selecting a business opportunity of type i and let $\psi_{i,t}$ be the mass of (exogenously given) business opportunities of technology type i, such that $\Psi = \sum_i \psi_{i,t}$. The total number of successful startups of

---

[21]The precise details of how the composition changes with X_t is described in Subsection 3.4.

type i, $m_{i,0,t}$ is assumed to be determined by a Cobb-Douglas matching function

$$m_{i,0,t} = e_{i,t}^{\phi} \psi_{i,t}^{1-\phi}, \text{ for } i = 1, 2, .., I, \tag{3}$$

where ϕ is a parameter common to all technology types in order to avoid any a priori heterogeneity in entry sensitivities. The probability of successfully starting up a firm of type i, conditional on paying the entry cost, is then given by $P_{i,t} = m_{i,0,t}/e_{i,t}$. Free entry implies the following condition for each technology type

$$\chi = P_{i,t} V_{i,0,t}(0, \mathcal{F}_t), \text{ for } i = 1, 2, .., I, \tag{4}$$

where $V_{i,0,t}(0, \mathcal{F}_t)$ is the value of a newborn firm of type i. The above equation makes clear that if the value of a new firm of type i increases, the probability of a successful startup, $P_{i,t}$, adjusts downward to restore equilibrium.[22] Hence, the value of newborn firms within each technology type directly pins down the number of startups of each type.

Our modeling of the entry stage is technically convenient but we believe it is also has conceptual appeal for several reasons.[23] First, agents have a *choice* of what type of firm to start up, rather than being exogenously confronted with a particular technology. Second, agents who aim to start up more ambitious firm types - associated with larger firm values - face tougher competition in starting up their business. Third, technologies of successful entrants are no longer strictly superior to those of failed attempts. This implication is attractive in the light of the empirical evidence that many entrepreneurs have low growth ambitions and are not very skilled (see e.g. Hurst and Pugsley, 2011; Campbell and De Nardi, 2009), while at the same time a substantial fraction of highly skilled and experienced entrepreneurs fail to get a new business ambition off the ground.

Moreover, note that the composition of startups with respect to scalability is endogenous in our model. All shocks affecting the value of firms impact the composition of startups through the free entry condition (4). Taking logs and total differentiation of the free entry condition implies that $d \ln m_{i,0,t} - d \ln e_{i,t} = d \ln P_{i,t} = -d \ln V_{i,0,t}$. Similarly, the matching function implies that $d \ln P_{i,t} = (\phi - 1)(d \ln e_{i,t} - d \ln \psi_{i,t})$. It follows that elasticity of the number of actual entrants, $m_{i,0,t}$, with respect to the associated firm value

---

[22]Note that there is no entry when $V_{i,0,t}(0, \mathcal{F}_t) < \chi$.

[23]A technical advantage of our approach that we do not exploit in this paper is that it naturally accommodates the presence of multiple dimensions of heterogeneity among firms. In firm dynamics models along the lines of Hopenhayn and Rogerson (1993), entry decisions are typically characterized by a cut-off rule, which relatively difficult to implement in the presence of several dimensions of heterogeneity.

is given by $\phi/(1-\phi)$.[24] The parameter ϕ thus commands the degree to which the number of entrants changes in response to the changes in firm values. This property is exploited in our calibration, described in Section 4, which uses the volatility of the number of entrants to pin down ϕ.

Fluctuations in firm values are in turn driven by fluctuations in firm profits. One important reason why profits can respond differently across types is that profit margins depend on the degree of returns to scale. Firms with higher returns to scale optimally choose larger steady-state output quantities but lower profit margins. Aggregate shocks tend to affect profit margins across firms by similar amounts, which means that the *percentage* change in profit margins is higher for larger firms. Hence, their values tend to respond more to aggregate shocks. A very similar effect is described in Gourio (2007), who compares firms that differ in their share of labor versus capital in production.

Moreover, entrant composition is also directly affected by the exogenous composition shock X_t. This shock can be thought of as a primitive technology shock, resembling the vintage-technology models of Campbell (1998) and Gilchrist and Williams (2000). Over our estimation sample, which runs from 1979 to 2011, innovations in information technology may have generated opportunities to create particularly large businesses like Google and Microsoft. Also, ICT innovations may have affected the scalability of existing business formats. An alternative interpretation of the shock is that it reflects time-variation in frictions affecting startups of firms, for example related to obtaining finance as in Bassetto, Cagetti, and De Nardi (2012). Here, we avoid taking a strong a priori stance on the source of the composition shock.

3.3 Households

There is a representative household which consists of a continuum of members, some of which supply labor on a perfectly competitive market. The household maximizes the expected present value of life-time utility, subject to its budget constraint:

$$\max_{\{C_t, N_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma} - 1}{1 - \sigma} - \nu Z_t N_t \right) \quad \text{s.t.}$$

$$C_t = W_t N_t + \Pi_t, \tag{5}$$

---

[24]Similarly, one can show that, keeping firm values constant, the elasticity of $m_{i,0,t}$ with respect to the measure of opportunities $\psi_{i,t}$ is 1.

where C_t is the total amount of goods purchased by the household, N_t denotes total employment within the household, $\sigma > 0$ is the coefficient of risk aversion, $\nu > 0$ is a parameter capturing the disutility of labor, Z_t is a stochastic preference shock, Π_t denotes firm profits and $\beta \in (0, 1)$ is the household's subjective discount factor. Following the indivisible labor models of Hansen (1985) and Rogerson (1985), we assume that utility is linear with respect to labor supply. Profits and the wage are taken as given by the household. The optimal employment choice takes on the familiar form:

$$W_t = Z_t \frac{\nu}{C_t^\sigma}. \tag{6}$$

The first-order condition makes clear that Z_t drives a wedge between the marginal product of labor and the households intratemporal marginal rate of substitution. Hence it has been labeled a "labor wedge" in the literature and is typically thought of as a shock that may also capture time-varying labor market frictions.

3.4 Shock processes

Before implementing the model quantitatively, we fill in the final details of the shock processes. First, we specify how precisely the composition shock affects the distribution of business opportunities. Assume, without loss of generality, that technology types reflect the degree of returns to scale in increasing order, such that $i = 1$ is associated with the lowest and $i = I$ with the highest degree of returns to scale. Finally, let ι be the median technology type.[25] The measure of business opportunities of type i in period t is given by

$$\psi_{i,t} = X_t \overline{\psi}_i \qquad \text{if} \quad i < \iota, \tag{7}$$

$$\psi_{i,t} = \overline{\psi}_i \frac{\Psi - \overline{\psi}_\iota - X_t \sum_{j=1}^{\iota-1} \overline{\psi}_j}{\sum_{j=\iota+1}^{I} \overline{\psi}_j} \qquad \text{if} \quad i > \iota, \tag{8}$$

where a bar indicates steady state values and $\Psi = \sum_i \psi_{i,t}$. The composition shock thus shifts mass from the upper half of returns to scale technologies to the lower half, in proportion to the respective steady state levels.

Next, we specify the stochastic processes for all four shock variables. We assume each to follow an AR(1) process:

$$J_t = 1 - \rho_J + \rho_J J_{t-1} + \epsilon_t^J, \tag{9}$$

---

[25]In the quantitative simulations, the number of technology types is odd. However, allowing for an even number of technology types is straightforward.

where $J = A, Q, X, Z$, where ϵ_t^J are i.i.d. innovations distributed normally with mean zero and standard deviations σ_J and ρ_J being the respective persistence parameter.

3.5   Equilibrium

For computational reasons to be explained in Section 4 we impose a maximum firm age K, that is, we set the exit probability at age K, ρ_K, equal to one. Using that all firms of the same age and technology type take the same decisions, the aggregate resource constraint, the labor market clearing condition and the law of motion for the measure of firms by age and technology type can be written, respectively, as:

$$\sum_{i=1}^{I} \sum_{a=0}^{K} m_{i,j,t} \left(y_{i,a,t} - \frac{\zeta Q_t}{2} (n_{i,a,t} - n_{i,a-1,t-1})^2 \right) - \sum_{i=1}^{I} e_{i,t} \chi = C_t, \tag{10}$$

$$\sum_{i=1}^{I} \sum_{a=0}^{K} m_{i,a,t} n_{i,a,t} = N_t \tag{11}$$

$$m_{i,a,t} = (1 - \rho_{a-1}) m_{i,a-1,t-1} \text{ for } a = 1, 2, ..K \text{ and } i = 1, 2, .., I \tag{12}$$

Let $\mathcal{F}_t = \left[A_t, Q_t, X_t, Z_t, \{m_{i,a-1,t-1}, n_{i,a-1,t-1}\}_{i=1,..I,\ a=0,..,K-1} \right]$ be the aggregate state, consisting of the measure of firms of each age-technology combination up to age $K-1$, the employment levels of these firms in the previous period, as well as the values of the stochastic aggregate shocks. We are now ready to define a recursive equilibrium.

Definition (recursive equilibrium).
A recursive competitive equilibrium is defined by laws of motion for:
- the representative household's labor supply, $N(\mathcal{F}_t)$, and consumption $C(\mathcal{F}_t)$,
- the wage $W(\mathcal{F}_t)$,
- firm value functions $V_{i,a}(n_{i,a-1,t-1}, \mathcal{F}_t)$ and employment choices $n_{i,a}(n_{i,a-1,t-1}, \mathcal{F}_t)$, for $i = 1, 2, .., I$ and $a = 0, 1, .., K$,
- the measure of potential entrants $e_i(\mathcal{F}_t)$ and startup probabilities $P_i(\mathcal{F}_t)$ for $i = 1, 2, .., I$,
- the measure of operating firms of type i and age a, $m_{i,a}(\mathcal{F}_t)$, that solve the household's problem, solve the firm's problem, satisfy the free entry condition for each technology type $i = 1, 2, .., I$, satisfy the aggregate resource constraint, clear the labor market, and obey the laws of motion for the elements of the aggregate state $\mathcal{F}_t$.

The system of model equations we use to solve for the equilibrium consists of equations (1)-(9) plus the law of motion for the aggregate shocks.

4 Quantitative Implementation

4.1 Parametrization strategy

We parameterize the model using a combination of the Simulated Method of Moments (SMM) and Maximum Likelihood (ML) estimation. The dynamic model is solved using a first-order perturbation method around the stationary equilibrium (i.e. around the steady state growth path of firms). We impose a maximum firm age of 50 years, which makes the aggregate state finite. This enables us to track the aggregate state *entirely*, given the approximated policy functions, instead of being forced to revert to iterative methods in the spirit of Krusell and Smith (1998) which rely on an approximation of the aggregate state. The solution method is explained in detail in Appendix C.

Even though the aggregate state consists of over 900 state variables and the aggregate shocks have continuous support, our computational strategy enables us to solve the model in several minutes on a desktop.[26] This relatively fast computation of the equilibrium makes it possible to estimate parameters, which requires the model to be solved many times.

4.2 Parameters calibrated to match long-run targets

Following the frequency of the BDS data we set the model period to one year. We divide the parameters calibrated to match long-run targets into three groups. First, parameters pertaining to the household, second firm-level parameters that are common to all firm types, and third parameters that are specific to technology types. All model parameters are summarized in Table 2.

4.2.1 Household parameters

Household preferences are chosen in line with conventional values in the macro literature. The household's discount factor, β, is set to 0.96, corresponding to an annual real interest rate of four percent. The households' coefficient of relative risk aversion, σ, is set to one which implies log utility with respect to consumption. The preference parameter ν is backed out from the household's first order condition for a given wage and total

---

[26]The large number of state variables results from the fact that the aggregate state consists of the measure of active firms and its distribution across age/type categories, together with the associated employment levels from the previous period.

Table 2: Calibrated parameters

	parameter	value	target/estimate
β	discount factor	0.96	annual interest rate 4%
σ	relative risk aversion coefficient	1	log-utility
υ	disutility of leisure	0.06	unit labor supply
ζ	adjustment cost, age 1-50	0.004	size of 1 year old firms
ξ_0	exit rate coefficient	0.050	exit rates by age, BDS data
ξ_1	exit rate coefficient	0.170	exit rates by age, BDS data
χ	entry cost	0.930	entry costs = 0.073 GDP
Ψ	measure of business opportunities	0.090	$M = 1$, normalization
ϕ	elasticity in entry function	0.300	std(entry)/std(y)
ρ_A	TFP shock, persistence	0.896	
σ_A	TFP shock, standard deviation	0.011	
ρ_Q	expansion cost shock, persistence	0.593	
σ_Q	expansion cost shock, standard deviation	0.158	
ρ_X	composition shock, persistence	0.415	
σ_X	composition shock, standard deviation	$9e-6$	
ρ_Z	labor wedge, persistence	0.751	
σ_Z	labor wedge, standard deviation	0.012	

α_i	returns to scale									average size in BDS size classes
	0.890	0.932	0.946	0.956	0.963	0.968	0.972	0.976	0.988	
$P_i = \left(\frac{\psi_i}{x_i}\right)^{1-\phi}$	probability of starting up a type i firm									firm shares in BDS size classes
	0.625	0.357	0.218	0.123	0.070	0.040	0.022	0.013	0.002	

Notes: The table the calibrated parameters and their respective targets or sources. Since the magnitude of the measure of business opportunities of type i firms is hard to grasp, we rather report the probabilities of successfully starting up a business of type i (P_i) conditional on paying the startup cost.

Table 3: Exit rates by age

	Firm age					
	1	2	3	4	5	> 5
data	0.22	0.15	0.12	0.11	0.10	0.06
model	0.22	0.14	0.11	0.09	0.08	0.06

Notes: The table reports firm exit rates in the BDS database and in the model according to age.

consumption. We normalize the wage such that the model matches a profit rate of 3% taken from Hornstein, Krusell, and Violante (2005).

4.2.2   Parameters common to all firm types

Parameters that are common across all technology types are the adjustment cost level, ζ, the exogenous firm exit rate, ρ_a, the entry cost, χ, and the mass of potential entrants, Ψ. We set ζ to match the average size of entrants, which is 6.1 in the BDS data. As will become clear below, the model also matches the average size of $16 - 20$ year old firms which then pins down the average firm growth rate. To capture age-dependency of exit rates observed in the data, we introduce the following parametric relation between age and the exit probability

$$\rho_a = \xi_0 + \frac{\xi_1}{a}, \ \varphi_0, \varphi_1 > 0, \ a < K,$$

where a is the firm's age and K is a maximum age. The top row of Table 3 contains average exit rates for firms aged one to five, as well as the average exit rate for older firms. We target these numbers, setting $\xi_0 = 0.05$ and $\xi_1 = 0.17$. The bottom row of Table 3 shows that the implied exit rates match their data equivalents very closely.

The last two parameters in this category pertain to firm entry. We set the entry cost χ such that total entry costs are equal to 0.73% of GDP which is the average value for the US economy in the years 2004 to 2010 as documented by the "Doing Business" database of the World Bank. Finally, the measure of business opportunities Ψ is set such that the total mass of firms in the economy (M) is normalized to 1 in the steady state.

4.2.3   Firm-type parameters

Model parameters that describe firm technology types are the returns to scale parameters and the steady-state measure of business opportunities in each technology type $(\overline{\psi}_i)$. The

Table 4: Average firm size and firm shares in BDS size categories

	Firm size								
	$1-4$	$5-9$	$10-19$	$20-49$	$50-99$	$100-249$	$250-499$	$500-999$	>1000
average size	2	6	13	30	68	149	335	658	$3,315$
firm shares	47.6%	23.8%	14.6%	9.1%	2.7%	1.4%	0.4%	0.2%	0.2%

Notes: The table reports average firm sizes and firm shares within a given size class.

presence of heterogeneity in technology types implies a cohort-level size distribution of firms, which we can confront with the BDS data. We set the total number of technology types equal to the number of size groups available in the BDS database, where we group the three largest size categories into one, giving us $I = 9$ technology types.

We exclude production functions with increasing returns to scale, that is, $\alpha(i) < 1$ for $i = 1, .., I$. To pin down the returns to scale parameters, we target average firm size in the 9 size categories reported in the BDS data for firms aged between 16 and 20 years (averaged over the period $2000 - 2010$). The implied values for the returns to scale parameters are shown in the bottom part of Table 4. They range between 0.890 and 0.988, which is within the range of estimates of Basu and Fernald (1997).

To pin down the measure of business opportunities in each technology type (ψ_i), we match the distribution of the *number* of firms between 16 and 20 years old, over the nine size categories reported in the BDS data, again averaged over the period $2000 - 2010$. The firm shares and average firm sizes in the BDS size brackets are reported in Table 4.

4.3 Estimated parameters

The remaining parameters (the elasticity parameter in the entry matching function and parameters related to the aggregate shocks) are estimated using either the Simulated Method of Moments (SMM) or Maximum Likelihood (ML) estimation.

We choose SMM to estimate the elasticity of entry with respect to firm value, ϕ, because it is closely related to the volatility of the number of entrants, a key second moment of our model. Specifically, we target the relative volatility of (the log of) the number of entrants with respect to (the log of) real GDP.

The parameters of the aggregate shocks, on the other hand, come with no obvious associated moments and therefore a likelihood based approach seems more appropriate. In particular, we estimate the persistence parameters ρ_J and volatility parameters σ_J,

with $J = A, Q, X, Z$. We use four data series for this purpose: aggregate real GDP, the aggregate employment rate, computed as one minus the unemployment rate, the average size of entrants and the average size of five year old firms taken from the BDS data.[27]

Our estimation strategy has the following steps. First we guess the value for ϕ and then estimate the remaining parameters using ML and simulate the model. Next, we evaluate the second moment of interest over the sample period and go back to the first step, updating the the initial guess of ϕ. We repeat this procedure until we are reasonably close to our target. An important by-product of the estimation is that we obtain estimated time-series for the aggregate shocks and the endogenous state, which we use in counterfactual exercises. More details on the solution and estimation procedure are provided in Appendix C.

4.3.1 Shock identification

Because the results will depend crucially on the estimated importance of the various shocks, we now discuss their identification on more detail. While real GDP and the employment rate are informative about the aggregate TFP and labor wedge shocks, the average size time-series provide identification of the the composition and expansion cost shocks. To help understand the identification of composition effects implied by the model structure, Figure 6 depicts the impulse response functions to the four shocks in the model of the average size among entrants and cohorts of age five, two key variables used in the estimation. The response of these two variables to TFP and labor wedge shocks is qualitatively similar. However, their response to the expansion cost and the composition shock is very different. The main identification comes from the fact that while the effect of the composition shock gets stronger as the cohort ages (and firms with high returns to scale account for a larger share of employment), the effect of the expansion cost shock diminishes as the cohort ages (bottom panels of Figure 6). In particular, both shocks trigger a substantial decline in the size of entrants on impact but there is hardly any contemporaneous response of five year old firms. Five years after the shock, the size of five year old firms declines strongly. These declines apply to the cohort born in the initial period of the shock. Note however, that the effect of the expansion cost shock on this cohort has become smaller once it has reached five. This happens because the expansion cost shock mainly affect post-entry decisions, which can be gradually reversed as the shock

---

[27]To be consistent with the timing in the BDS, we construct annual time series for GDP and employment over March-to-March time intervals. All time series are in logs and linearly detrended.

Figure 6: Shock identification: impulse response functions
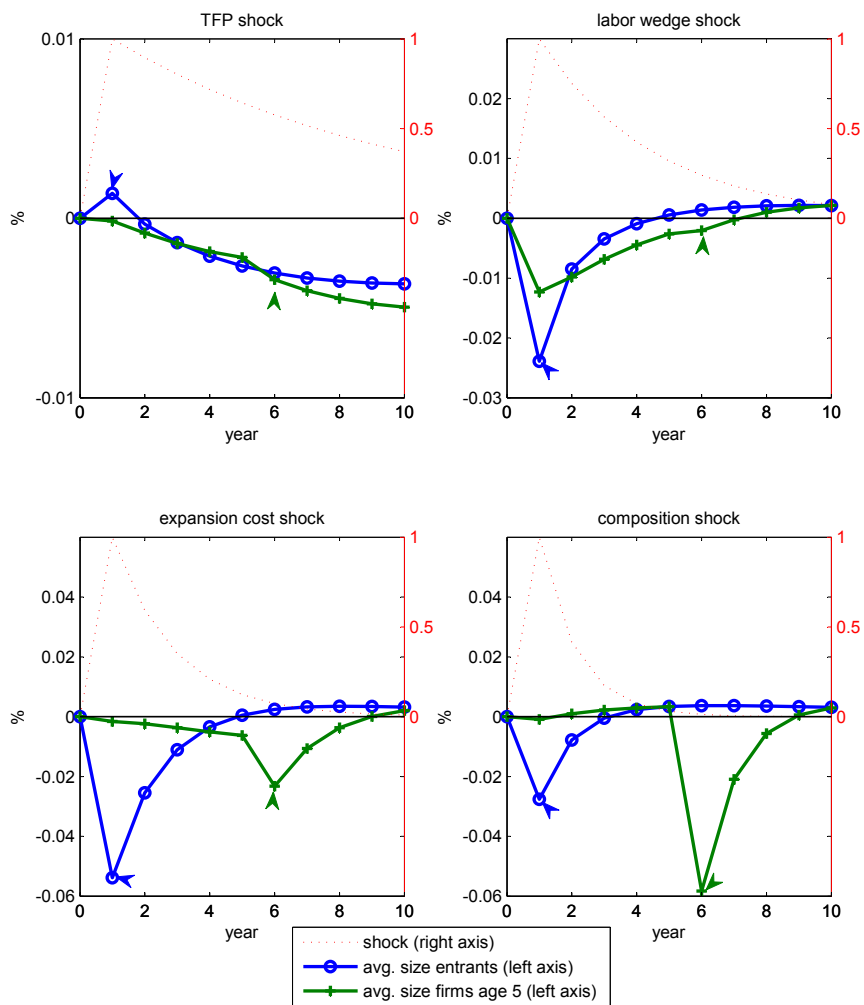
Notes: Responses to shocks of size one standard deviation. Dashed lines denote responses of the exogenous shock variables and have been re-scaled to increase by one on impact. Arrows denote the cohort born in the initial period of the shock

Figure 7: Steady state: firm size by age and type
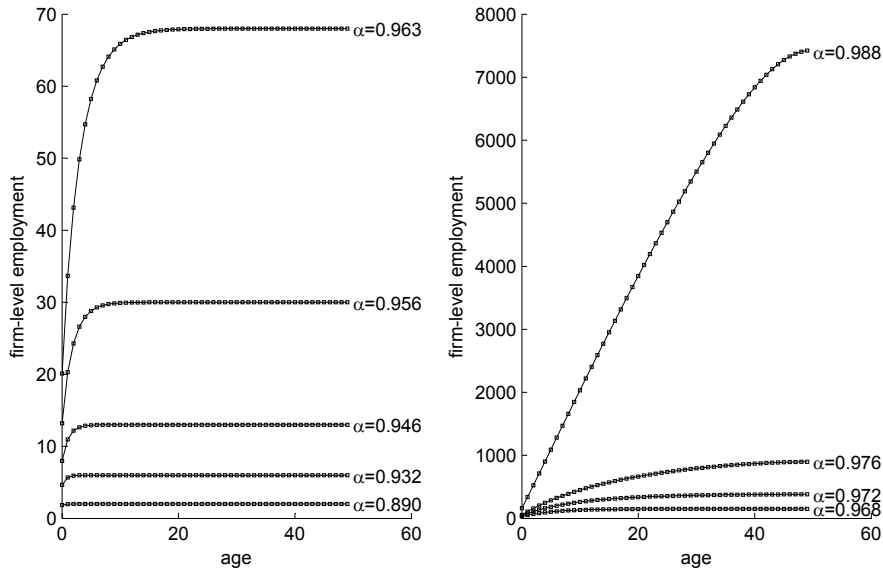
dies out. On the contrary, the effect of the composition shock becomes larger as the cohort ages and the firms with high returns to scale come to dominate the cohort.

4.4 Properties of the steady-state equilibrium

Before further evaluating the dynamics predicted by the model, we analyze the deterministic steady state equilibrium. Figure 7 plots the steady-state employment patterns of firms by age and technology type. Firms of the lowest returns-to-scale type ($\alpha = 0.890$) are born small and stay small during their entire life, starting off with an employment level of 1.8 which grows to only 2 later on in the firms' lives. On the other extreme, the most scalable firms ($\alpha = 0.988$) have nearly constant returns to scale and grow from 247 employees in the year of startup to a maximum of 7800 employees.

As firms with high returns to scale grow more in the years after birth, they account for an increasingly large share of the cohort's total employment. The shares of firm types in total cohort-level employment are also displayed by age in Figure 8. While the most scalable firms account for only about 7 percent of the cohort's total employment in the year of birth, they provide more than half of the cohort's employment by the age of fifty. Firms with low returns to scale, on the other hand, are relatively important during the early years of a cohort's life, with the firms of the three lowest returns-to-scale types creating more than 50 percent of the cohort's jobs in the year of birth. By the age of fifty,

28

Figure 8: Steady state: contributions to cohort-level employment by age and type
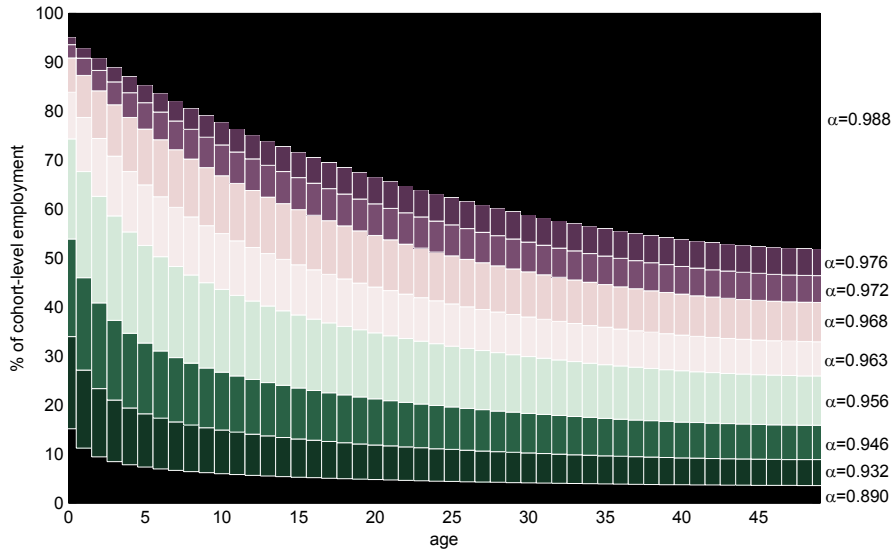
however, their share in the cohort's employment level has declined to about 16 percent.

4.5 Model performance

We now evaluate the model's dynamic performance along several dimensions not directly exploited in the estimation. First, let us assess the model's success in capturing the empirical stylized facts presented in Section 2. As in the data, the model predicts a high positive correlation between job creation of entrants and aggregate employment (correlation coefficient of 0.85).[28] Also, the correlation between cohort-level employment at birth and five years later is 0.71, very close to the empirical value of 0.68. Moreover, a variance decomposition of cohort-level employment of five year old firms within the model reveals that 65% of the variation is driven by the intensive margin, which is again very close to the 69% found in the data.

Next, we evaluate the predictions of the model for firms older than five years. Although a limitation of the BDS data is that we cannot track individual cohorts above age 5, we do observe averages over firms aged 5 to 10, as well as from 11 to 15. Each year, a new cohort enters the group whereas another leaves the to a higher age category. One could therefore expect that cohort-level differences are reflected in the (percentage) *change* in

---

[28]Similarly high correlations are found for job creation of entrants and aggregate output, also when using alternative detrending methods.

employment and average size within these groups over time. Hence we compute time series for these statistics using our model estimates and compare these to the equivalents in the BDS data. In the BDS data, we observe a correlation between employment growth and the growth rate of average size of firms aged 5 to 11 of 0.85. In our model, this correlation is only slightly higher (0.90). For firms aged 11-15, the correlation is 0.88 in the data and 0.73 in the model. Hence, also for older firms the average size margin appears to be a main driver behind fluctuations in cohort-level employment, both in the model and the data.

We also directly correlate the model-predicted time series with their BDS equivalent, we find that for the percentage change in employment of firms aged 5-11, the correlation 0.61. For the percentage change in average size we find an even higher correlation (0.67). For firms aged 11 to 15, these correlations are still significantly positive: 0.39 and 0.34, respectively. We think these results are reassuring for several reasons. First, we do not use any direct information on firms between 5 and 11 in our calibration or estimation procedure. Second, in the estimation we only use cohort-level information about average size in the estimation, not total cohort-level employment. Third, there is uncertainty around our estimation of parameters and shock realizations. Finally, one could think of additional, unmodelled factors driving year-to-year differences in the averages over age groups, for example related to time variation in exit rates. Nonetheless we remain reluctant to draw strong conclusions for very old cohorts, for which few data points are available.

As discussed in the previous section the elasticity parameter, ϕ, in the entry matching function is closely related to the volatility of the number of entrants. The estimated value delivers a relative volatility of the number of entrants and output of 2, which is close to the empirical value of 2.5. Moreover, the model also correctly predicts that firm entry is pro-cyclical as in the data.[29]

Finally, we compare the model's predictions on real wages to the data.[30] The correlation between the real wage in the model and the data is encouragingly high: 0.6. Also, the volatility of the real wage is only moderately higher in the data than in the model (1.33 in the data versus 1.08 in the model). The pro-cyclicality of wages, however, is too strong in the model relative to the data, a common finding in Neoclassical models of the business cycle.[31]

---

[29]The correlation coefficient between firm entry and real GDP is 0.67 in the data and 0.92 in the model.

[30]As a data equivalent we use real hourly compensation in the nonfinancial corporations sector, as reported by the Bureau of Labor Statistics. We linearly de-trend the series.

[31]The correlation between the real wage and the employment rate is 0.15 in the data and 0.76 in the

5 Model results

This section presents the main results obtained from the model. First, we analyze cohort-level fluctuations implied by the model. Our primary goal is to quantify the importance of the year of birth in determining a cohort's success in providing jobs and to understand the underlying drivers. Next, we investigate whether birth-determined factors can help to understand fluctuations in aggregate employment.

5.1 Cohort-level fluctuations and the importance of conditions at birth

At any age after birth, a cohort's employment level is to some extent determined by the economic state in the year of birth. The remainder is due to shocks that realized after birth. It is, however, difficult to disentangle the relative importance of the two contributors by using reduced-form empirical techniques, if only because the aggregate state may include unobservable variables. Within our estimated model, however, we can quantify the contribution of the economic state at birth precisely. To do so, first define cohort-level employment as $N_{a,t} \equiv \sum_{i=1}^{I} n_{i,a,t}$. We can then decompose cohort-level employment as:

$$N_{a,t} = \mathbb{E}_{t-a} \left[N_{a,t} \right] + \widehat{N}_{a,t}.$$

Here, $\mathbb{E}_{t-a} \left[N_{a,t} \right]$ is the expectation of $N_{a,t}$ conditional on information available in the year of birth, which is a function of only the aggregate state in the year of birth. It follows that $\widehat{N}_{a,t}$ is the prediction error, which depends only the shocks realized in the years after birth, which are orthogonal to the state in the year of birth. Using this orthogonality we can decompose the unconditional variance of $N_{a,t}$ as:

$$Var \left(N_{a,t} \right) = Var \left(\mathbb{E}_{t-a} \left[N_{a,t} \right] \right) + Var \left(\widehat{N}_{a,t} \right)$$

The top left panel of Figure 9 plots the results of the variance decompositions for cohorts up to twenty years after birth. The importance of the aggregate state at birth is overwhelming, contributing over 90 percent to employment variance, regardless of age. A very similar pattern is found for cohort-level average size (middle left panel). However, for the employment of an *individual* firm of a certain type, the state at birth quickly loses importance in the years following birth (bottom left panel). The main reason is that the composition of newborn firms does not play any direct role in the evolution of an individual firm. The persistence that remains is driven by persistent inherent to the shock process and the endogenous part of the aggregate state.
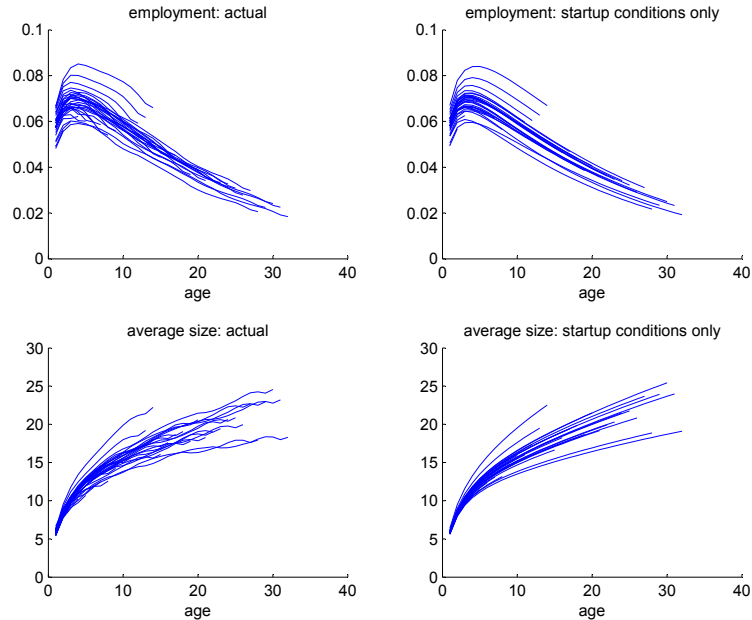
model.

Figure 9: Model: variance decompositions.
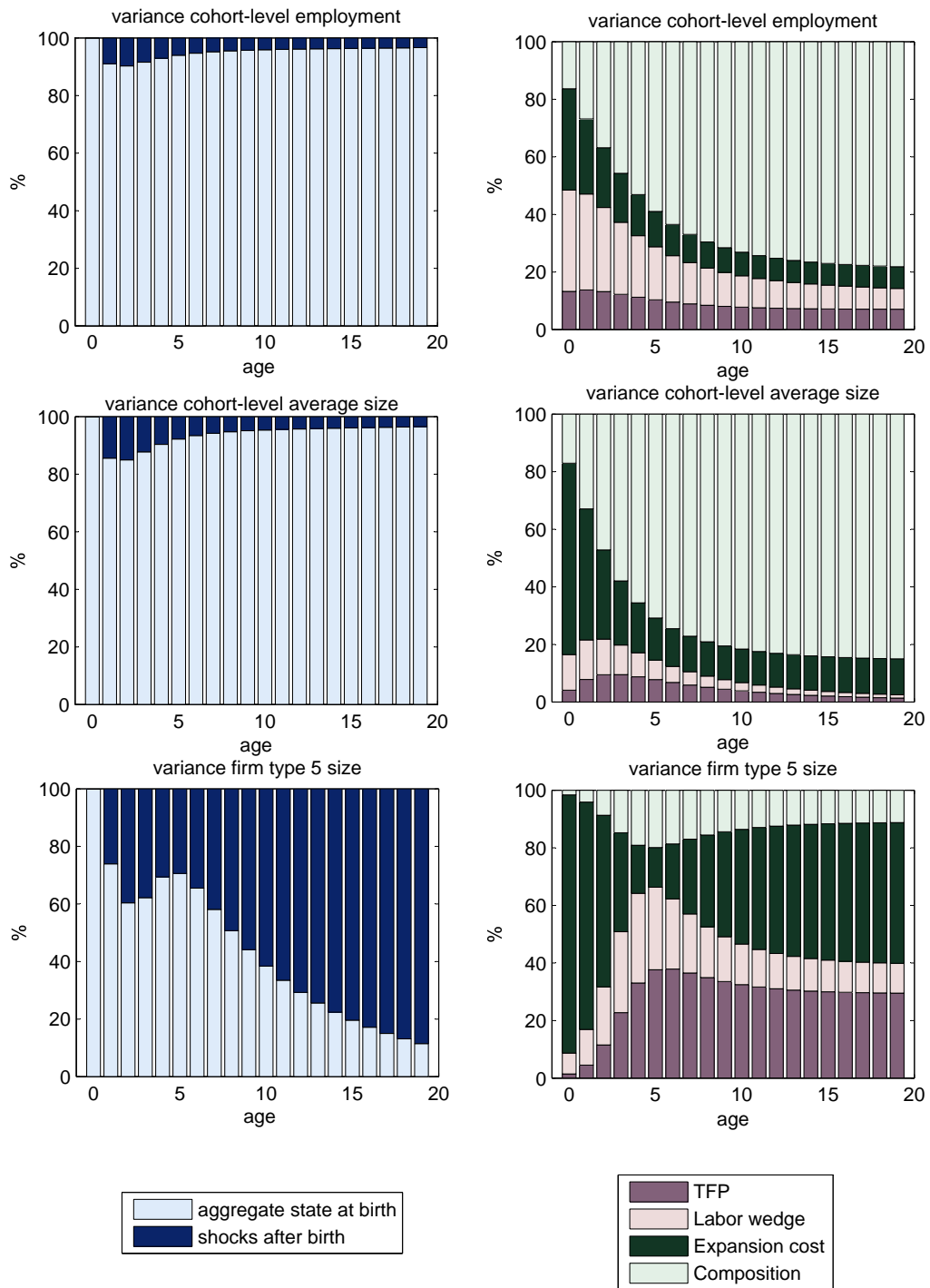
Left panels: aggregate state at birth versus shocks after birth. Right panels: decomposition with respect to shocks.

Additional insight into the drivers of cohort-level persistence is obtained by quantifying the contributions of the four aggregate shocks (right panels of Figure 9). For cohort-level employment, the composition shock is of minor importance in the year of birth. As the cohort ages the composition shock gains prominence, accounting for nearly 80 percent of cohort-level employment twenty years after birth. For average size, we observe a similar pattern, with TFP and labor wedge shocks being nearly irrelevant for variation in cohort-level variables of firms at age 20. The expansion cost shock, however, remains relatively important. It is clear from the variance decomposition for an individual firm of a given type that the composition shock does not account for much of the fluctuations. Interestingly, the contribution of composition shocks do not fall to zero, which is because of general equilibrium responses triggered by the shock.

Figure 10: Model: estimated cohort-level outcomes over the period 1979-2011

Notes: Employment is expressed as a fraction of steady-state aggregate employment. Left panel: model estimations. Right panels: estimations in which employment within each age/type category is set to the steady-state level ("startup conditions only").

We can also visualize the importance of the state at birth specifically for our sample period 1979-2011. The two left panels of Figure 10 plot the estimated levels of employment (top panel) and average size (bottom panel) over our data sample (recall that average size at age zero and five match the true data). Moving beyond age five, average size across

cohorts diverges further while employment differences across cohorts persist.

More interestingly, the two right panels of Figure 10 plot the same cohorts, but now re-setting the employment choice within each age/type bracket to their values in the deterministic steady state. This exercise therefore isolates fluctuations that are purely due to factors determined at birth, i.e. the number of firms in the cohort and their composition over types. The figure shows that the resulting cohort-level outcomes are quite similar to those obtained from the full simulation, only smoother. The figure thus makes clear also that over the specific sample period pre-birth effects account for the majority of cohort-level outcomes.

5.2 Aggregate implications

This subsection discusses the aggregate implications of changes in startup conditions. First we isolate, in an accounting sense, the contribution of startup decisions to fluctuations in the employment rate. Second, we quantify to what extent changes in startup conditions can matter for the aggregate economy in general equilibrium.

5.2.1 The contribution of startup conditions to aggregate fluctuations

Within our model, one can inspect the effect of startup conditions not only on cohort-level variables, but also on the aggregate economy. To this end, we again fix average size in each age/type bracket to its steady state value, but let the distribution of firms (their number and composition) vary as predicted by the estimated model, as in the right panels of Figure 10. Aggregating over all firms in any given year, we obtain a time series representing fluctuations in the *aggregate* employment rate accounted for by changes in startup conditions alone. Figure 11 plots this time series together with the actual aggregate employment rate. The figure also plots two time series for the trend component of the actual employment rate, constructed using the HP filter with smoothing coefficient 6.23 and 100.

Figure 11 shows that the magnitude of aggregate employment fluctuations due to startup conditions alone is large, with a volatility of nearly two thirds relative to the actual employment rate series. Secondly and remarkably, the employment rate implied by changes in startup conditions closely resembles the empirical trend components in aggregate employment. The correlation coefficient between the counterfactual employment rate and the HP-filter trend is 0.61 for smoothing coefficient 6.23 and 0.56 for smoothing coef-

Figure 11: Employment rate: data, HP-trends and estimated contribution of startup decisions
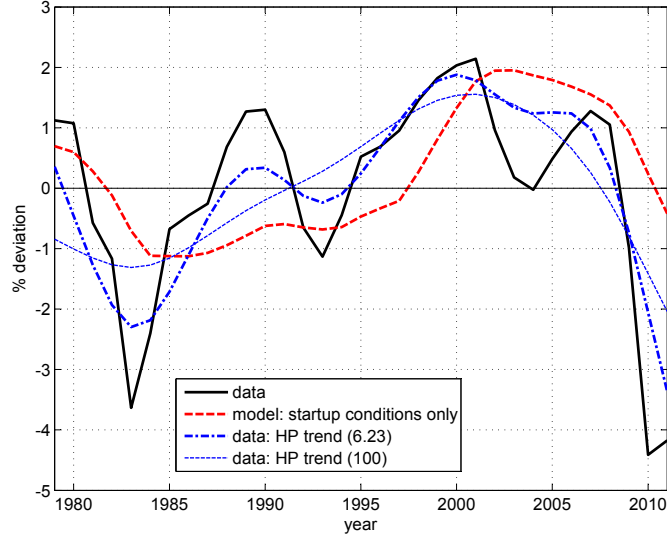
Notes: The figure plots the employment rate data used in the estimation, a counterfactual employment rate based on fixing the age/type size to their respective steady state values and letting only the mass of each age/type vary and the HP-filter trend in the employment rate.

ficient 100.[32] Thus, startup decisions appear important to understand the low-frequency movements of *aggregate* employment, often ignored in business cycle analysis.

Decomposing the counterfactual further into the contribution of the number of firms and the contribution of the composition of firms reveals a roughly equal importance. The volatility of the counterfactual employment rate resulting from changes in the composition of firms only is 45% that of the counterfactual which allows for changes in both the composition and the number of firms.

5.2.2 Counterfactual scenarios and equilibrium effects

The importance of startup conditions in accounting for low-frequency fluctuations in employment raises the question whether there is a role for policy to either stabilize fluctuations or to raise employment permanently. The extent to which such policies are desirable

---

[32]The estimation uses linearly detrended employment rate data. However, the linear trend is very modest and therefore comparing the counterfactual with the HP-filter trend of the data used for estimation delivers very similar results.

or even possible depends crucially on the deeper disturbances and frictions driving our estimated shocks. In this paper, we deliberately avoid imposing structure on the precise nature of these forces, limiting the scope for detailed policy evaluation. However, a second important determinant of the success of policy interventions is the degree to which general equilibrium effects work against the desired policy outcome. Our model framework incorporates clearing in all markets and is therefore suitable to investigate such effects.

We analyze the aggregate effects of two counterfactual scenarios over our sample period. In the first counterfactual we study full stabilization of the number of entrants and their composition. Panel 1 of Table 5 displays the effect on various aggregate variables, averaging over 5 year periods between 1979 and 2011.[33] The effects of stabilizing startup conditions are sizeable. Changes in aggregate employment and output can be as large as one third to one half of a percentage point on average in the five year windows.[34] However, the table also makes clear that stabilizing startup conditions does not fully eliminate the low-frequency components of aggregate employment. In equilibrium, post-entry employment decisions adjust in response to the policy, offsetting part of its effects. Finally, effects on labor productivity and wages appear to be relatively slow moving: in the counterfactual these variables are higher during the first half of the sample, but lower during in the post-1995 period.

In the second counterfactual, we study a once-and-for-all change in the distribution of business opportunities, increasing the fraction startups with high returns to scale. We implement this counterfactual as a semi-permanent shock to the composition of startup-opportunities, with a magnitude equal to twice the estimated standard deviation.[35] Panel 2 of Table 5 shows that the change gradually pushes up average firm size. This is because composition effects gain in strength as cohorts age. A similar, but quantitatively weaker, increase is seen for aggregate output and average labor productivity. By contrast, aggregate employment actually falls in the long-run. Intuitively, the shift towards startups with higher returns to scale increases average labor productivity and output. In the longer run, the increase in income drives up consumption reducing households' labor supply and hence aggregate employment. Thus, our simulations point to the presence of strong equilibrium effects, especially in the medium and long run.

---

[33]To make full use of our sample the first and last columns average over six years.

[34]Further splitting up the period into annual intervals gives a maximum effect of 0.7 percent

[35]The maximum realization over our sample is 3.8 standard deviations.

Table 5: Counterfactual scenarios: aggregate implications

	1979-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2011
1. fixed entry and composition						
average firm size	0.63	−0.21	0.36	1.32	1.71	0.85
employment	0.34	−0.02	−0.02	−0.31	0.11	0.19
output	0.46	0.09	0.04	−0.35	0.01	0.13
labor productivity	0.12	0.11	0.06	−0.04	−0.10	−0.06
real wage	0.07	0.15	0.07	−0.04	−0.15	−0.09
2. larger startups						
average firm size	0.80	1.99	3.25	4.41	5.34	5.70
employment	0.18	0.12	0.06	0.01	−0.03	−0.05
output	0.20	0.25	0.27	0.28	0.29	0.29
labor productivity	0.03	0.12	0.21	0.27	0.32	0.34
real wage	0.00	0.10	0.18	0.23	0.27	0.27

Notes: percentage deviations from the benchmark model. All variables denote aggregates.

6 Conclusion

This paper exploits the recent opportunity to break down aggregate employment data into cohort-level observations, in order to improve our understanding of fluctuations in macroeconomic aggregates. New stylized facts direct our attention to the birth stage of newborn firms and in particular decisions affecting their scalability. Our results indicate that the impact of these decisions not only persists as cohorts mature, but actually grows over time since highly scalable firms need time to reach their full potential. Hence, compositional differences across cohorts become increasingly pronounced with age, accounting for slow-moving but large fluctuations in aggregate employment.

We also gain insight into the sources of compositional differences in scalability across newborn cohorts. The reduced-form empirical analysis suggests that these are related to the state of the aggregate business cycle. Within our estimated structural model, much of the variation across cohorts is accounted for by a shock to the distribution of new business opportunities. One can interpret this shock as a fundamental disturbance to the economy resulting from innovations. This interpretation, however, is premature since it is not difficult to think of endogenous frictions affecting the scalability of newly created

businesses, for example frictions related to their finance. Given that these frictions may be affected by policies, an important avenue for future research is to scrutinize various competing explanations for the observed variations across cohorts, possibly using more disaggregated data sources.

References

ABBRING, J., AND J. CAMPBELL (2004): "A Firm's First year," Tinbergen Institute Discussion Paper 05-046/3.

ACEMOGLU, D., AND W. HAWKINS (2010): "Wages and Employment Persistence with Multi-worker Firms," mimeo.

BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2009): "Measuring and Analyzing Crosscountry Differences in Firm Dynamics," In *Producer Dynamics: New Evidence from Micro Data*,NBER Chapters, pp. 15–76. National Bureau of Economic Research, Inc.

BASSETTO, M., M. CAGETTI, AND C. DE NARDI (2012): "Credit Crunches and Credit Allocation in a Model of Entrepreneurshipa," working paper.

BASU, S., AND J. FERNALD (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105(2), 249–283.

CAMPBELL, J. (1998): "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, 1(2), 371–408.

CAMPBELL, J., AND M. DE NARDI (2009): "A Conversation with 590 Nascent Entrepreneurs," *Annals of Finance*, 5(3), 313–340.

CLEMENTI, G. L., AND D. PALAZZO (2010): "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," mimeo.

ELSBY, M., AND R. MICHAELS (2013): "Marginal Jobs, Heterogeneous Firms and Unemployment Flows," *American Economic Journal: Macroeconomics*, 5(1), 1–48.

FORT, T., J. HALTIWANGER, R. JARMIN, AND J. MIRANDA (2012): "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," mimeo.

FUJITA, S., AND G. RAMEY (2009): "The Cyclicality of Separation and Job Finding Rates," *International Economic Review*, 50(2), 415–430.

GILCHRIST, S., AND J. WILLIAMS (2000): "Putty-Clay and Investment: A Business Cycle Analysis," *Journal of Political Economy*, 108(5), 928–960.

GOURIO, F. (2007): "Labor Leverage, Firms Heterogeneous Sensitivities to the Business Cycle, and the Cross-Section of Returns," mimeo.

HALTIWANGER, J., R. JARMIN, AND J. MIRANDA (2013): "Who Creates Jobs? Small vs. Large vs. Young," *The Review of Economics and Statistics*, 45(2), 347–361.

HANSEN, G. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16(3), 309–327.

HOLMES, T., AND J. STEVENS (2012): "An Alternative Theory of the Plant Size Distribution with Geography and Intra- and International Trade," mimeo.

HOPENHAYN, H., AND R. ROGERSON (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101(5), 915–938.

HORNSTEIN, A., P. KRUSELL, AND G. VIOLANTE (2005): "Unemployment and Vacancy Fluctuations in the Matching Model: Inspecting the Mechanism," *Federal Reserve Bank of Richmond Economic Quarterly*, 91, 19–51.

HURST, E., AND B. PUGSLEY (2011): "What Do Small Businesses Do?," NBER working paper no. 17041.

JOVANOVIC, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649–670.

KAAS, L., AND P. KIRCHER (2011): "Efficient Firm Dynamics in a Frictional Labor Market," IZA discussion paper no. 5452.

KRUSELL, P., AND A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(51), 867–896.

LEE, Y., AND T. MUKOYAMA (2012): "Entry, Exit, and Plant-level Dynamics over the Business Cycle," mimeo.

LUCAS, R. E. (1978): "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 9(2), 508–523.

MOEN, E. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*, 105(2), 385–411.

MOSCARINI, G., AND F. POSTEL-VINAY (2010): "Stochastic Search Equilibrium," mimeo.

———— (2012): "The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment," *American Economic Review*, 102(6), 2509–2539.

POSCHKE, M. (2012): "Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice," *Journal of Economic Dynamics and Control*, forthcoming.

RAVN, M., AND H. UHLIG (2002): "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 84(2), 371–380.

ROGERSON, R. (1985): "Indivisible Labor, Lotteries, and Equilibrium," *Journal of Monetary Economics*, 21(1), 3–16.

SCHAAL, E. (2010): "Uncertainty, Productivity and Unemployment in the Great Recession," mimeo.

SEDLÁČEK, P. (2013): "Firm Age, Business Cycles and Aggregate Labor Market Dynamics," mimeo.

Appendix

A   Establishments

The main text documents our new stylized facts for firms. Using the BDS data one can also inspect establishment-level information. An establishment is defined as a single physical location where business is conducted or where services or industrial operations are performed. A firm, on the other hand, is a business organization consisting of one or more establishments that were specified under common ownership or control. Therefore, the firm and the establishment are the same for single-establishment firms, but existing firms can create new establishments. The following paragraphs show that at the establishment-level our empirical findings remain to hold.

As for firms, the variation in the number of jobs created by new establishments is robustly pro-cyclical and large. The correlation coefficient of establishment entrant job creation with the employment rate (real GDP) is 0.63 (0.69) using linear detrending. The correlations when considering HP-filetered data or data in levels remain large and positive.[36] Moreover, the volatility of jobs created by new establishments (in logs) is large amounting to 5.4 times that of the volatility of (log) real GDP.

Figure (12) shows the correlation coefficient of employment in year t with that in year $t + a$ of the same cohort. The figure shows a very high persistence for cohort-level employment, which strongly contrasts that of aggregate employment. Notice, that the correlation of employment of entrants and five year old establishments of the same cohort is even higher than that computed using firm-level data (0.77 for establishments compared to 0.68 for firms).

Finally, when decomposing employment variation of five year old establishments into the intensive (average establishment size) and extensive (number of establishments) margins, one finds that the majority is driven by the intensive margin (58%). As in the firm-level case, it also holds true that entrants and the first year after entry accounts for the vast majority of the variation (85%).

---

[36]For HP-filtered data the correlation coefficients are 0.35 (0.38) when considering the employment rate (real GDP) as business cycle indicators. The correlation with the level of the employment rate (growth rate in real GDP) is 0.66 (0.15).

Figure 12: Autocorrelations: establishments
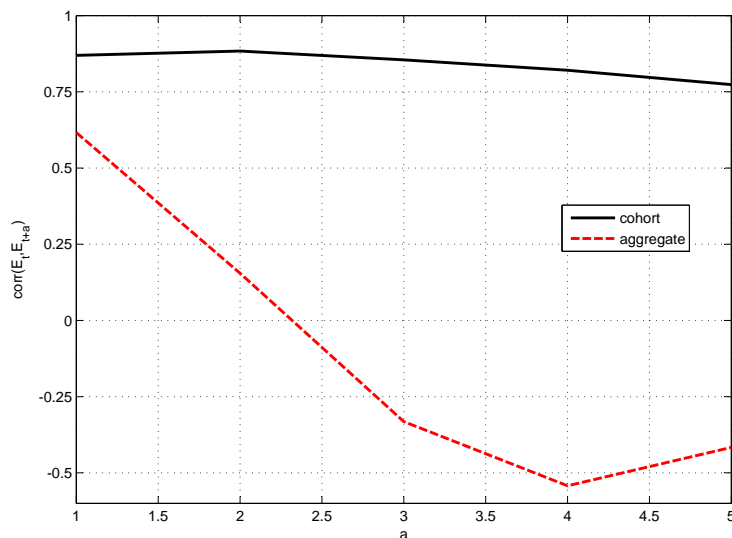
Notes: Correlation coefficients of employment at establishments in year $t = 0$ and in year $t + a$, with $a = 1, 2, 3, 4, 5$, at both the level of a cohort born in period $t = 0$ and at the aggregate level.
Source: BDS.

B Sectoral evidence

This section investigates to what extent our empirical findings may be driven by cyclical sectoral composition changes of entering firms. For this purpose we use the BDS sectoral breakdown, which includes information on nine 1-digit sectors.[37]

To gain insight into the importance of sectoral shifts for aggregates, we compute a two counterfactual time series of average entrant size. First, we construct a counterfactual entrant size under the assumption that the distribution of the number of entrants over the nine sectors remains fixed over time, setting the fractions equal to their sample averages. This series captures variation that is due within sector variations in average size only. Second, we compute a counterfactual series that captures only between-sector shifts, by setting the average entrant size within each sector equal to the sample average, but let fractions of entrants in the nine sectors to vary over time as in the data. Figure 13 displays the two counterfactual time series, as well as the actual series for average size

---

[37]The data is broken down into the following sectors: (i) Agriculture, Forestry, and Fishing, (ii) Mining, (iii) Construction, (iv) Manufacturing, (v) Transportation, Communication, and Public Utilities (vi) Wholesale Trade, (vii) Retail Trade, (viii) Finance, Insurance, and Real Estate, (iv) Services.

Figure 13: Average entrant size: data and counterfactuals
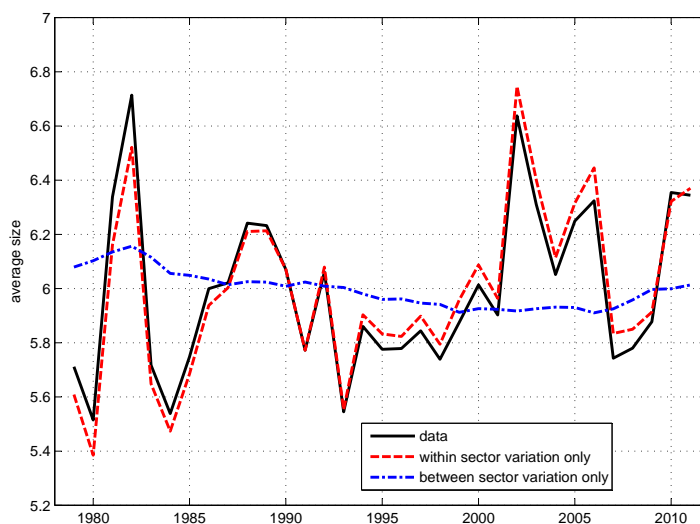
Notes: "within sector variation only" average size is constructed by fixing the sectoral shares of entrants to their sample average. "Between sector variation only" average size is constructed by fixing the average size of entrants within sectors to their respective sample averages.

Source: BDS, authors' calculations.

within newborn cohorts. It is immediately clear that within-sector variations account for almost all of the variation in average size; between-sector shifts appear to play an extremely limited role.

Next, we repeat our empirical analysis within each of the nine sectors in the BDS separately. The results are reported in Table (6) and show that our earlier findings also broadly hold within sectors. This gives further support that the economy-wide results are not driven by cyclical sectoral shifts.

In particular, all sectors are characterized by large persistence in cohort-level employment, which is in stark contrast to the persistence found in the sector as a whole. Most sectors are also characterized by strongly pro-cyclical job creation by entrants. The exceptions are mining (strongly counter-cyclical) and transportation, communications and utilities (a-cyclical).[38] Both of these sectors account for a very small fraction of firms and employment in the economy and therefore are unlikely to influence the aggregate cyclical

---

[38]Retail trade also displays positive, but statistically insignificant correlations at the firm level. At the establishment level, however, the pro-cyclicality is strong.

Table 6: Summary of stylized facts within sectors

		AGR	MIN	CON	MAN	TCU	WHO	RET	FIRE	SRV
Cyclicality										
firms	e-rate	0.46	−0.37	0.73	0.60	0.10	0.45	0.15	0.54	0.44
	GDP	0.56	−0.38	0.76	0.63	0.15	0.51	0.25	0.67	0.52
establishments	e-rate	0.51	−0.39	0.68	0.54	0.20	0.46	0.58	0.54	0.51
	GDP	0.63	−0.38	0.74	0.62	0.28	0.54	0.58	0.62	0.54
Persistence										
firms	cohort	0.42	0.79	0.75	0.66	0.81	0.76	0.69	0.36	0.56
establishments	cohort	0.65	0.84	0.68	0.64	0.89	0.80	0.80	0.85	0.72
	total	−0.21	0.11	−0.24	0.22	0.13	0.12	−0.15	−0.37	0.16
Variance decomposition										
firms	extensive	0.13	0.51	0.89	0.38	0.14	0.44	0.59	0.10	0.33
	intensive	0.87	0.49	0.11	0.62	0.86	0.56	0.41	0.90	0.67
establishments	extensive	0.37	0.55	0.90	0.41	0.06	0.48	0.35	0.30	0.42
	intensive	0.63	0.45	0.10	0.59	0.94	0.52	0.65	0.70	0.58

Notes: "Cyclicality" reports the correlation coefficients between linearly detrended log job creation of entrants and the employment rate or real GDP in the different sectors for firms and establishments. "Persistence" reports the correlation coefficients between (linearly detrended) entrant job creation and employment in 5 year old firms or establishments within the same cohort, both for the individual cohorts and for the sector as a whole. Finally, "Variance decomposition" reports the contribution of the extensive (number of firms or establishments) and intensive (average size) margin to variation in employment of 5 year old firms or establishments (based on linearly detrended data).

properties.[39] Finally, in most sectors it is the intensive margin which drives the majority of variation of employment among five year old firms or establishments. The exceptions are construction and retail trade where the the intensive margin contributes with 11 and 41%, respectively.

Our findings are also related to results of Lee and Mukoyama (2012) who document that in recessions entering plants in manufacturing are on average larger than those entering in booms. Their findings are based on the Annual Survey of Manufacturers from the U.S. Census Bureau for the period 1972-1997. Their measure of the business cycle is given by the growth rate of manufacturing output. Interestingly, we confirm their finding in the BDS. When we compute the correlation of average size of newborn firms in manufacturing and the growth rate of real GDP, we find it is significantly negative. However, for other de-trending methods and business cycle indicators and when using data on establishments this correlation drops to virtually zero in the BDS data.

C   Computation and Estimation

This part of the appendix provides details on the numerical solution procedure and the estimation. To economize on notation, we can express the model compactly as:

$$\mathbb{E}_t f\left(y_{t+1}, y_t, x_{t+1}, x_t; \Upsilon, \eta\right) = 0$$

where x_t is a vector containing the state variables (all variables in $\mathcal{F}_t$) and y_t is a vector containing the non-predetermined variables, Υ is a vector containing all parameters of the model and η is a scalar parameter pre-multiplying the covariance matrix of the shock innovations, as in Schmitt-Grohé and Uribe (2004). Importantly, the above is system of a finite number of expectational difference equations.

C.1   Solving for the steady state without aggregate uncertainty

We first solve for the equilibrium of a version of the model without aggregate uncertainty. That is, we find vectors $\overline{y}$ and $\overline{x}$ that solve $f\left(\overline{y}, \overline{y}, \overline{x}, \overline{x}; \Upsilon, 0\right) = 0$. As described in the main text, we calibrate various parameters to match long-run targets. The calibration procedure has the following steps:

---

[39]The firm (employment) share of mining and transportation, communications and utilities are 0.5(0.6)% and 4.4(4.2)%, respectively.

1. given values for the technology type parameters (α_i), the adjustment cost level (ζ) and the aggregate wage rate (w), one can calculate the growth paths of firm-level employment and firm value leading towards the average size targets of 16-20 year old firms (and the growth paths for higher ages). While the average size targets are informative about the values of α's (for given values of ζ and w), the wage rate is pinned down by the average profits target, and the adjustment cost is pinned down by targeting average entrant size (effectively targeting the growth rate between entry and 16-20 years).

2. given firm values of startups from (1.), and a value of the entry cost (recall that the calibration targets a total cost of startups as a fraction of output), one can back out the startup probabilities from the free entry condition.

3. given the startup probabilities from (2.), and the firm shares of technology types taken from the BDS (shares of 16-20 year old firms in the given size brackets), one can back out the mass of business opportunities of the given technology types.

4. given the mass of business opportunities from (3.), the probabilities of starting up from (2.), and the age-dependent death rates (taken from the BDS), one can calculate the mass of firms in each age-type cell.

5. given firm masses from (4.), employment choices from (1.), one can calculate all the aggregates (total employment, output, consumption etc.) and back out the disutility of supplying labor from the households first-order condition.

## C.2 Solving for the equilibrium with aggregate uncertainty

Next, we solve for the dynamic equilibrium using first-order perturbation around the deterministic equilibrium found in the previous step. The first-order approximated solutions, denoted by hats, have the following form:

$$
\begin{aligned}
\widehat{x}_{t+1} &= \overline{x} + \Theta \left(\widehat{x}_t - \overline{x} \right) \\
\widehat{y}_{t+1} &= \overline{y} + \Phi \left(\widehat{x}_t - \overline{x} \right)
\end{aligned}
$$

where Θ and Φ are matrices containing the coefficients obtained from the approximation. The perturbation procedure is standard and carried out in one step.

An advantage of perturbation methods is that the computational speed is relatively high and many state variables can be handled. An important prerequisite for perturbations

Figure 14: Employment levels: types 3, 6 and 9
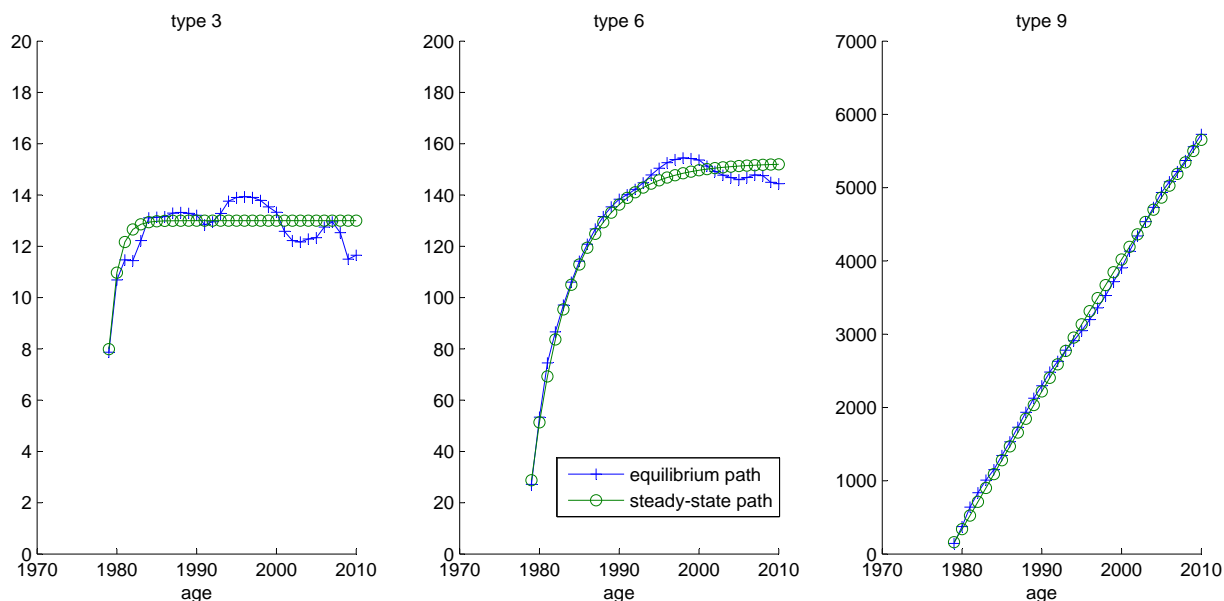
Notes: Green line represents the steady-state path. Blue line is the estimated equilibrium path for firms born in 1979.

to be accurate, however, is that deviations from the steady-state are not too large. For firm dynamics models like ours this may seem problematic because differences in employment levels across firms may be very large. Our approach, however, overcomes this problem since the steady state we perturb around contains the entire *growth path* of firms. These growth paths, captured by the constants in the above equations, are themselves non-linear functions of age and type.

Hence, the fact that most newborn firms starts off much below their eventual sizes does not involve large accuracy losses since the same is true for the steady-state sizes of newborn firms. Similarly, the fact that the equilibrium features various firm types with very different optimal sizes does not reduce accuracy since we perturb around the growth path for each individual firm type. To illustrate these points, the Figure (14) plots a simulated employment levels of firms of various types. The figure also plots the steady-state path in the absence of aggregate shocks, the centre of the first-order approximation. At each point in the simulation, the employment level of the firm is close to the steady-state path used for the approximations, even though differences across type- and age-groups are very large.

C.3 Estimation

Having solved the model for given parameter values we can compute the likelihood of the linearized model. To do so, the log-linearized model is set into state-space form and the parameters of the aggregate shock processes are estimated using Maximum Likelihood.

Using the Kalman filter enables one to conveniently characterize the likelihood function as well as obtain estimates of the underlying aggregate shocks, given the four observable time series. An important by-product of estimating the model is that we obtain model-predicted time series for all the variables in the model. This means that we obtain the entire time-varying distribution of firms (their masses, employment levels and firm values), which we can use for counterfactual analysis.

D Variation in firm exit rates

The model in the main text assumes constant, though age-dependent, firm exit rates. A concern could be that variation in exit rates is an important feature of the data responsible for a large part of variation in employment. This subsection suggests that incorporating variation in firm exit rates would result in only minor changes of our results.

D.1 How important is firm exit for employment?

To get a sense of how important can variation in firm exit rates be for employment we construct two counterfactual time series. The first tries to quantify how important is time-variation for the evolution of employment. The second goes a step further and acknowledges that different firm types might behave differently in terms of firm exit. Specifically, we construct a counterfactual employment time series according to

$$E_t^c = E_{t-1} + JC_t - (JD_t - JDd_t + JDd_t^c), \tag{13}$$

where a c in the superscript indicates a counterfactual, JC_t is gross job creation, JD_t is gross job destruction and JDd_t is gross job destruction due to firm exit in period t. The two counterfactual employment series we create differ in the way we construct JDd_t^c. First, we assume that the number of jobs lost due to firm exit is fixed at its sample average, i.e. $JDd_t^{1,c} = 1/T \sum_{t=1}^{T} JDd_t$.

Figure 15: Employment levels: data and counterfactuals
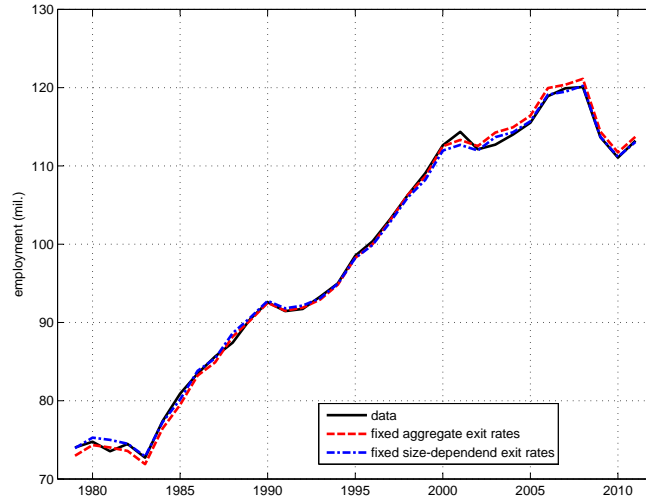
Notes: The figure plots the data and two counterfactual aggregate employment levels. "fixed aggregate exit rate" is constructed by fixing the overall firm exit rate to the sample average. "Fixed size-dependent exit rates" is constructed by fixing the firm exit rates within each size bracket in the BDS and averaging over the sample.

Second, we try to go a step further and acknowledge that firms of different types may behave differently in terms of firm exit. As a proxy for firm type, we use the BDS size brackets. Therefore, we construct the counterfactual number of lost jobs due to exit as

$$JDd_t^{2,c} = 1/T \sum_{t=1}^{T} \sum_j N_{j,t} S_{j,t} \left(\sum_j \frac{N_{j,t}}{N_t} \frac{Nd_{j,t}}{N_{j,t}} \right),$$

where j is an index representing the size brackets in the BDS data, N is the number of firms, S is the average firm size and Nd is the number of firms that shut down. In other words, we first compute the weighted average of firm exit rates, where the weights are firm shares, in each period. We multiply this by the number of firms and average firm size in each size bracket. In this way we obtain a counterfactual number of jobs lost if firms of all sizes had the same job destruction rate, but taking into account that they have different employment shares (i.e. exactly the case of our model which assumes the same exit rate over different firm types). We then average this over the sample period and create the counterfactual employment time series according to (13).

Figure 15 shows aggregate employment and the two (mean adjusted) counterfactual

Figure 16: Variance decomposition for a model with and without time-varying exit rates.
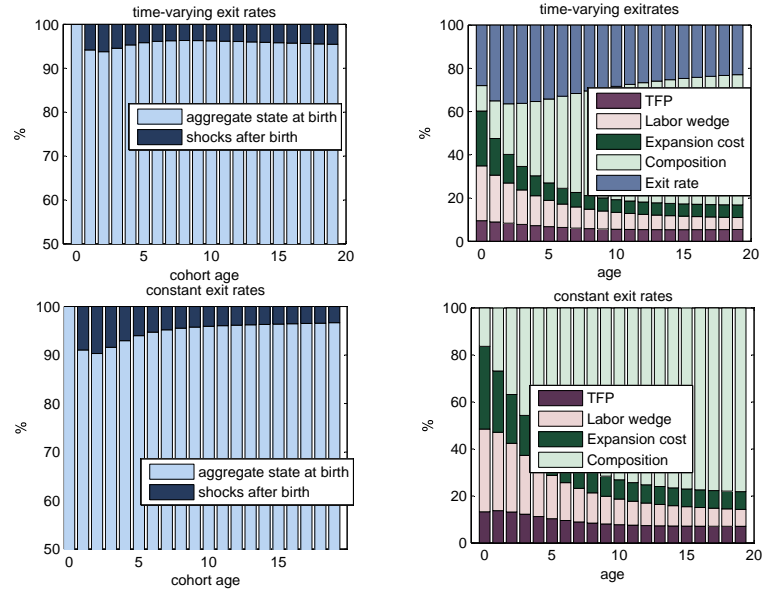
Notes: "Composition" denotes the shock to the distribution of business opportunities"

employment time series.[40] Clearly, the counterfactuals are very close to the data suggesting that not much information is lost by assuming constant firm death rates. In particular the standard deviation of the (log) employment counterfactuals are both about 0.96 that of the data.

D.2 Incorporating time variation in firm exit in the model

To investigate the importance of time variation in exit rates for our results, we introduce a stochastic aggregate shock to the exit rate to the model. We assume the shock follows an AR(1) process with normally distributed innovations. We estimate the autocorrelation coefficient and standard deviation of the innovation from BDS data on exit rates. The estimated parameters are 0.324 and 0.068, respectively. After adding the shock to the model, we re-compute the variance decomposition for cohort-level employment. Figure 16 plots the result, together with the variance decomposition for our benchmark model.

The results show that adding stochastic exit rate does not alter our conclusion on the

---

[40]Due to the averaging of firm exit rates the mean of the number of jobs lost due to firm death changes. We are, however, concerned about differences in the cyclical pattern.

importance on the state at birth. For younger firms, the initial state actually becomes more important relative to the benchmark. This is true even though exit rates shocks introduce substantial additional variation in cohort-level employment. Behind these results are equilibrium effects: a positive shock to firm exit encourages firm entry, replacing the firms that exit. This equilibrium effect on entry, however, this is fundamentally part of the "birth state" and persists as the cohort ages.