Lost Generations of Firms and Aggregate Labor Market Dynamics

Petr Sedláček*

[a] *University of Oxford*

Abstract

Can the unprecedented lack of startups during the U.S. Great Recession have persistently negative effects? While fewer firms hiring workers can mechanically reduce employment for many years, this may be offset by feedback effects on lower wages, slacker labor markets and higher profits. An estimated model of firm dynamics and frictional labor markets suggests that such feedback effects are too weak to offset the direct impact of fewer startups. Had firm entry remained constant during the Great Recession, output would have recovered 4-6 years earlier and unemployment would have been 0.5 percentage points lower even 10 years after the crisis.

Keywords: firm age, firm dynamics, heterogeneous firms, unemployment

JEL Classification: E24, E32, J64

*Address: University of Oxford, Department of Economics, Manor Road, Oxford OX1 3UQ, UK. Tel.: +44 1865 2719 51. Email: petr.sedlacek@economics.ox.ac.uk

1. Introduction

The number of startups in the U.S. hit an all time low at the end of the Great Recession in 2010 (31 percent below its pre-crisis level). At the same time, the unemployment rate peaked at 10 percent and remained close to this level even two years after the official end of the downturn. This paper asks to what extent these two phenomena are interlinked and what the lost generation of firms in the Great Recession implies for the U.S. economy in the medium- to long-run.

As a first step, this paper highlights that changes in firm entry impact the economy not only directly, but also indirectly in later years as affected cohorts of startups age. In particular, using Business Dynamics Statistics (BDS) data it is shown that the pro-cyclical nature of firm entry creates a ripple effect resulting in pro-cyclical movements of the share of young firms (not older than five years). Moreover, it is shown that young firms account for 40 percent of aggregate employment fluctuations (even though they employ only 16 percent of all workers). Cyclical changes in the firm age distribution therefore help shape aggregate fluctuations. These findings complement the results in Haltiwanger, Jarmin, and Miranda (2013) regarding young firms' disproportionate contributions to aggregate job creation in the long-run.[1]

The above findings raise concerns about the medium- to long-run impact of the recently lost generation of firms. A simple simulation of an exogenous drop in the number of startups (of the magnitude observed in the Great Recession) together with fixed survival and growth rates of incumbent firms suggests that the impact may be severe: even 10 years after the shock subsides, the unemployment rate remains more than 1 percentage point above its initial level.

However, the simple simulation exercise abstracts from potential general equilibrium feedback effects that may dampen the unemployment rate response. Therefore, in the next

[1] A related paper is Pugsley and Sahin (forthcoming) who analyze the effect of the secular decline in the share of startups on the aggregate economy.

step this paper uses a general equilibrium model of firm dynamics and a frictional labor market to demonstrate that while such effects are indeed important in the short-run, periods of subdued entry do negatively impact the economy in the long-run. The reason is that in the short run incumbent firms take advantage of the lack of job creation by startups and they almost fully compensate for the drop in employment. However, in future years the missing entrants generate fewer older firms (which on average account for the bulk of aggregate employment). This creates a persistent dent in the employment potential of the economy raising the "natural" rate of unemployment.

The general equilibrium effects dampening the short-run impact of a drop in firm entry operate mainly through the frictional labor market. In particular, because new (young) firms account for a large chunk of overall hiring, an exogenous drop in startups leads to a fall in aggregate vacancies. The slack labor market makes it easier for incumbent firms to hire. Moreover, employees become more reluctant to leave their current jobs because their outside options worsen. This resembles the "insulation effect" of recessions pointed out by Caballero and Hammour (1994). These factors are also reflected in the drop of wages further promoting new hiring. Finally, lower wages raise profits and thus induce an endogenous increase in the number of startups following the initial (exogenous) drop.

In order to quantify the impact of the lost generation of firms in the Great Recession, the structural model is estimated using BDS data. This enables us to determine to what extent fluctuations in firm entry were driven by forces specific to startups (i.e. not directly related to incumbents) and to what extent they were an endogenous reaction to a shock common to all firms. The results suggest that about 40 percent of the Great Recession drop in the number of startups can be attributed to factors specific to new firms, while the remainder is an endogenous response to a particularly strong recession.

Moreover, the estimated time-paths of the structural shocks and the model variables allow us to conduct counterfactual scenarios. The results reveal that had the number of startups remained at its pre-crisis level during and in the aftermath of the crisis, the immediate impact on the aggregate economy would have been relatively small. However, the larger share of young firms would have helped to speed up the recovery in later years. Specifically, output would have reverted back to its trend 4-6 years earlier and the unemployment rate would have been 0.5 percentage points lower even 10 years after the end of the crisis.

By documenting the cyclicality of the firm-age distribution and its effect on aggregate employment, this paper extends the results in Haltiwanger, Jarmin, and Miranda (2013) and Fort, Haltiwanger, Jarmin, and Miranda (2013) regarding the job creation prowess of young firms in the long-run and the cyclicality of young and small businesses. The structural model relates to several recent studies extending versions of the Mortensen and Pissarides (1994) model to include multi-worker firms and firm dynamics.[2] The focus on firm entry is related to Samaniego (2008), Lee and Mukoyama (2013) and Clementi and Palazzo (2013) who study entry and exit patterns within an extended version of the Hopenhayn and Rogerson (1993) model and to Siemer (2014) who investigates the impact of a drop in the number of startups after a financial shock within a partial equilibrium framework. In contrast to the above studies, the presented model focuses on firm entry in a search and matching framework highlighting the importance of general equilibrium feedback effects operating via the frictional labor market which go beyond an endogenous response of wages.

Finally, this paper builds on Sedláček and Sterk (2017) from whom it borrows two features. First, it uses their modelling of firm entry as an endogenous choice of a particular technology type rather than the result of a random draw from an exogenous distribution (as in e.g. Hopenhayn and Rogerson, 1993). Second, this paper uses their solution method which allows for the estimation of this general equilibrium heterogeneous firm model. In contrast to Sedláček and Sterk (2017), who focus on changes in the composition of startups over the business cycle, this paper investigates the aggregate impact of cyclical movements in the number of startups.

[2] See e.g. Elsby and Michaels (2013), Gertler and Trigari (2009), Kaas and Kircher (2015) and Schaal (2017).

The next section provides empirical evidence on young firms' contributions to aggregate employment variation and uses a simple exercise to show that a missing generation of firms has a potentially severe negative impact on the aggregate labor market. Section 3 builds a structural model with firm dynamics and a frictional labor market and Section 4 describes its calibration and quantitative properties. Section 5 presents the main results and the last section offers some concluding remarks.

2. Lost generations of firms and aggregate employment in the data

The Great Recession in the US was accompanied by a large drop in the number of startups. According to the Business Dynamics Statistics of the U.S. Census Bureau, an annual administrative dataset of almost all US private employers dating back to 1979, the number of startups hit an all time low in 2010, 31 percent below the pre-crisis level.[3] This section first documents how the cyclicality of firm entry and subsequent changes in the number of young firms (those younger than six years) contribute to aggregate employment variation.[4] Next, it shows that young firms contributed strongly to the observed drop in aggregate employment during and in the aftermath of the Great Recession. Finally, a simple counterfactual exercise suggests that the observed lack of firm entry may have a substantial and persistent negative impact on the aggregate labor market.

2.1. Contribution of young firms to aggregate employment variation on average

It is well known that plant entry is pro-cyclical in the manufacturing sector (see e.g. Campbell, 1999; Lee and Mukoyama, 2013). This pattern also holds for the aggregate economy. Depending on the detrending method, the correlation of the number of startups and

[3] The data represent a snapshot taken in March of each year. Availability starts in 1976, but the analysis drops the initial three years following Moscarini and Postel-Vinay (2012), who cast doubt on the data quality for the years prior to 1979.

[4] A firm is a business organization consisting of one or more establishments that were specified under common ownership or control. An establishment is defined as a single physical location where business is conducted or where services or industrial operations are performed. Results based on establishment data are very similar.

real GDP varies between 0.30 (linear trend) and 0.66 (HP-filter).[5]

Importantly, fluctuations in firm entry affect aggregate employment not only directly, but also indirectly in later years as cohorts of startups age. For instance, only about 2% of the variation in the number of young firms is driven by changes in their survival rates, while the rest is accounted for by fluctuations in (past) firm entry. This ripple effect of firm entry is then reflected in the pro-cyclical nature of the share of young firms in the economy. In particular, the correlation coefficient between the HP-filtered share of young firms and real GDP is 0.73.

This business cycle variation in the firm-age distribution may be inconsequential for aggregate labor market fluctuations as long as young and old firms contribute in similar magnitudes to overall employment variation. The variance of aggregate employment can be decomposed into contributions of young and old firms as follows

$$var\left(\frac{\widehat{E}_t}{\overline{E}_t}\right) = cov\left(\frac{\widehat{E}^y_t}{\overline{E}_t}, \frac{\widehat{E}_t}{\overline{E}_t}\right) + cov\left(\frac{\widehat{E}^o_t}{\overline{E}_t}, \frac{\widehat{E}_t}{\overline{E}_t}\right) + cov\left(\epsilon_t, \frac{\widehat{E}_t}{\overline{E}_t}\right), \tag{1}$$

where a bar stands for an HP-filter trend and a hat indicates deviations from this trend.[6] E_t denotes aggregate employment in period t, E^y_t and E^o_t denote employment by young and old firms, respectively, and ϵ is a residual term coming from the detrending method.

Expressing the relative contributions in percentage terms of total variance, young firms account for 38% of all fluctuations in aggregate employment on average. This number is rather striking when compared to the employment share of young firms which amounts to only 16%.[7] This result, related to business cycle frequencies, complements recent findings

[5] Throughout the paper the smoothing coefficient in the HP-filter is set to 1600 (100) for quarterly (annual) data.

[6] The components combine the effect of employment growth rates in young and old firms together with the changes in their shares: $\frac{\widehat{E}^y_t}{\overline{E}_t} = \frac{\widehat{E}^j_t}{\overline{E}^j_t}\frac{\overline{E}^j_t}{\overline{E}_t}$ for $j = y, o$ and $\epsilon_t = \frac{\overline{E}^y_t + \overline{E}^o_t}{\overline{E}_t}$.

[7] This is consistent with Fort, Haltiwanger, Jarmin, and Miranda (2013) who find that young/small businesses are more volatile than old/large ones. It is also not inconsistent with Moscarini and Postel-Vinay (2012) who document that the differential growth rate between small and large businesses is pro-cyclical. As Fort, Haltiwanger, Jarmin, and Miranda (2013) point out, small-old businesses, of which there are many in the BDS data (e.g. 60 percent of firms aged 25 and older has less than 10 employees), are much less

that young firms on average account for a disproportionately large fraction of overall net job creation (Haltiwanger, Jarmin, and Miranda, 2013). Combined, these facts suggest that a lost generation of young firms may not only lead to a sluggish recovery in the short run, but possibly also to a persistent drag on aggregate employment.

Further decomposing employment changes of young firms reveals that about half of the variation is accounted for by fluctuations in the number of young firms and the rest is due to changes in their average size.[8] Recalling that variation in firm survival rates accounts for very little of the changes in the number of young firms suggests that firm entry is responsible for about $15-20\%$ of aggregate employment fluctuations, even though the employment share of startups is only 3%. Moreover, this constitutes a lower bound because part of the contribution of old firms to aggregate employment fluctuations could also be traced back to past changes in firm entry.

2.2. Lost generation of firms during the Great Recession and the potential long-run impact

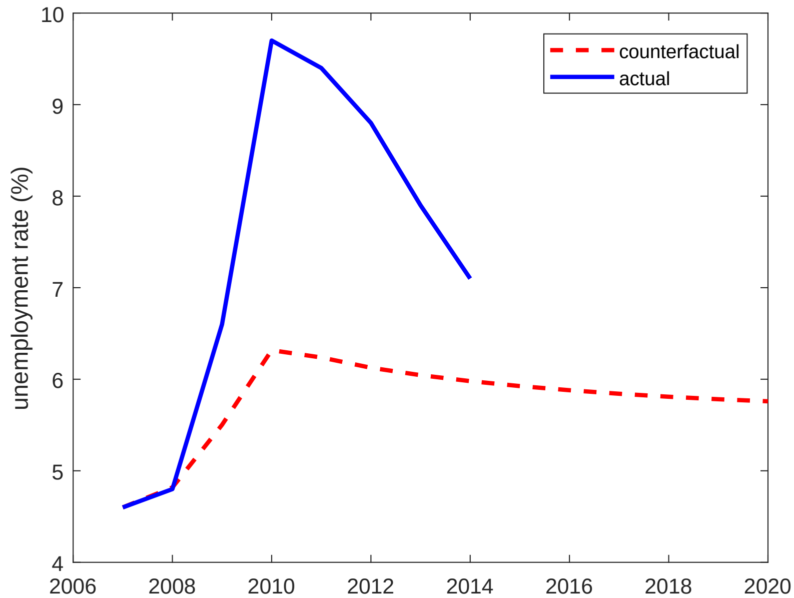
This subsection focuses specifically on the latest downturn for which the common view is that startups (and young firms) were hit particularly hard (see e.g. Fort, Haltiwanger, Jarmin, and Miranda, 2013; Sahin, Kitao, Cororaton, and Laiu, 2011). In particular, between 2007 and 2009 44% of the cumulative drop in aggregate employment was driven by young firms. In the aftermath of the crisis (between 2010 and 2012), when aggregate employment was still 5.1 percent below its pre-crisis level, this contribution was 70 percent. Moreover, three quarters of young firms' cumulative employment drop during and in the aftermath of the Great Recession was driven by a decline in their numbers, rather than a fall in their average size.

What is the potential medium- to long-run impact of the lost generation of firms coming from the Great Recession? A simple simulation can answer how the unemployment rate

cyclically sensitive than small-young firms.

[8] Conducting the decomposition on startups shows that 58% of employment variation is driven by changes in the number of startups. For young firms the contribution is 42%.

Figure 1: Unemployment rate; actual and counterfactual



Notes: The figure plots the actual unemployment rate together with a counterfactual one based on a fixed firm life-cycle as observed in the BDS and the observed drop in startups during the Great Recession.

would have responded to the observed firm entry drop but holding all else unchanged. Similar to Gourio, Messer, and Siemer (2014), Figure 1 shows a counterfactual unemployment rate based on an exogenous drop in the number of startups of the magnitude observed in 2008-2009, but where survival and growth rates of all firms are fixed to their sample averages.

Given the relatively small (3 percent) employment share of startups, the short-term impact of the drop in firm entry is mild (pushing unemployment up to 6.3% in 2010). Therefore, the lack of startups alone cannot explain the observed unemployment rate increase during the crisis. More strikingly, however, the negative effect is extremely persistent. Even ten years after entry reverts back the unemployment rate remains more than 1 percentage point above its pre-crisis level.

While this simple exercise gives an indication of the potential importance of lost generations of firms for aggregate labor market dynamics, the assumption of unchanged behavior of incumbent firms is unrealistic. The rest of the paper is therefore devoted to analyzing the effects of lost generations of firms within a structural model of firm dynamics in which businesses optimally hire and fire workers on a frictional labor market.

3. Labor market model with firm dynamics

This section builds a structural general equilibrium model of a frictional labor market with endogenous firm dynamics. This framework is particularly suitable for the question at hand because it allows for the possibility that a drop in firm entry feeds back into the employment behavior of incumbent firms through several channels. In particular, changes in the number of entrants will not only be reflected in bargained wages, but also in the chances of incumbent firms hiring new workers and via their effect on workers' outside options also in the probabilities of the employed separating from their current jobs.

Among other things, the model will be able to quantify to what extent the drop in firm entry was driven by a shock common to all firms or a shock specific to startups. A key prerequisite of such a decomposition is a model which can deliver realistic fluctuations in firm entry. While many existing models of firm dynamics have trouble replicating entry patterns (see e.g. Samaniego, 2008), the setup of firm entry in this paper, explained in detail in Subsection 3.4, does not suffer from such a drawback.

The following paragraphs describe the model economy. To facilitate the exposition of the model, aggregate variables are denoted by upper case letters, firm- or worker-specific variables are denoted by lower case letters and next period values are indicated by a prime.

3.1. Matching in the labor market

In each period, unemployed workers search for jobs and firms that wish to hire employees post vacancies. The total number of unemployed (U) and the total number of posted vacancies (V) engage in (random) matching on the aggregate labor market. The number of new hires is determined by an aggregate matching function

$$M = mU^{\mu}V^{1-\mu}, \tag{2}$$

where I follow the majority of the literature and assume a Cobb-Douglas functional form with m being the (constant) matching efficiency and μ being the elasticity of matches with respect to the number of unemployed. Given the above, the probability with which an unemployed worker finds a firm is given by $F = M/U$ and the probability with which a firm hires a worker is given by $Q = M/V$.

Already at this point it is possible to highlight one of the general equilibrium effects of a fall in the number of startups. In particular, the total number of vacancies is a sum over vacancies posted by individual firms. Therefore, a drop in firm entry will, ceteris paribus, decrease the total number of vacancies. This in turn increases the probability of hiring a worker (Q) making it easier for existing firms to compensate for the lack of job creation by startups.

3.2. Household behavior

The economy is populated by a representative household consisting of a continuum of risk-neutral and ex-ante homogeneous workers. Workers can find themselves either unemployed (and searching for jobs) or employed by one of the heterogeneous firms. At the beginning of each period, employed workers obtain an iid draw of (worker-specific) productivity z from a distribution $H(z)$. Particularly bad draws will result in employment relationships that are not profitable and such workers will be fired. Denote the cutoff value of worker-specific productivity below which employment relationships are severed by $\widetilde{z}$.

Household members pool their income from work and non-work activities and spend it on the consumption good. Formally, the household maximizes the present value of life-time utility (i.e. the present discounted value of consumption), subject to a budget constraint

$$C = bU + W + P, \tag{3}$$

where C is aggregate consumption, b is the value of home production, W is aggregate wage income and P are aggregate profits (the latter two are defined at the end of the section).

3.3. Firm behavior

This subsection describes the behavior of incumbent firms of which there is an endogenous mass. After the realization of aggregate shocks, but prior to observing worker-specific shocks, firms bargain with their workers over wages (w) that will be paid out in the current period. Thereafter, worker-specific productivity shocks are realized and firms decide to fire a fraction of their workforce for which the idiosyncratic productivity shocks were particularly bad, i.e. those with productivity draws below $\widetilde{z}$. Firms pay out the bargained wages to the remaining employees and produce output.

While it is assumed that all firms operate the same decreasing returns to scale production technology, they differ in the efficiency with which they operate it. In particular, there is a finite number of (permanent) technology types, indexed by $i = 1, 2, ..., I$, which differ in the level of total factor productivity (ϵ). The production function, which uses labor as its only input, is assumed to take on the following form

$$y = A\epsilon\widehat{z}n^{\alpha}, \tag{4}$$

where A is the level of aggregate productivity, $\widehat{z} = \int_{\widetilde{z}} z \frac{h(z)}{1-H(\widetilde{z})} dz$ is average worker-specific productivity of employees who remain in the firm, n is the number of workers in production and α is the parameter of decreasing returns to scale.

After observing all shocks, firms post vacancies on a frictional labor market to attract new workers for production in the next period. The costs of hiring are assumed to take on the following form

$$\frac{\kappa}{\gamma}x^{\gamma}n, \tag{5}$$

where $x = v/n$ is the vacancy rate (vacancies over employment) and $\gamma > 1$ and $\kappa > 0$ are parameters. This functional form is borrowed from Merz and Yashiv (2007) and states that the costs of hiring are proportional to the size of the firm and that they are convex in the hiring rate (see e.g. Kaas and Kircher, 2015; Gertler and Trigari, 2009, for models using such a functional form).

At the end of each period, firms face an exogenous but age-dependent probability of shutting down δ_a. The assumption of an exogenous exit probability is made not only for greater tractability, but is also justified by the BDS data which indicate that job destruction by exiting firms accounts for only one percent of the variation in aggregate employment.[9]

Notice that under the above assumptions all firms of the same type and age make the same decisions. This property increases the tractability of the model and simplifies the solution method. In what follows I will therefore index individual firms by their type (i) and age (a).

Formally, an incumbent firm maximizes expected firm value ($\Pi_{i,a}$) by choosing employment available at the beginning of the next period ($\widetilde{n}'_{i,a+1}$), the vacancy rate ($x_{i,a}$) and the cutoff for worker-specific productivity ($\widetilde{z}_{i,a}$) below which workers get fired, subject to the law of motion for firm-specific employment:

$$\Pi_{i,a} = \max_{\widetilde{n}'_{i,a+1}, x_{i,a}, \widetilde{z}_{i,a}} \left[y_{i,a} - w_{i,a} n_{i,a} - \frac{\kappa}{\gamma} x_{i,a}^{\gamma} n_{i,a} + \beta(1-\delta_a)\mathbb{E}\Pi'_{i,a+1} \right] \quad \text{s.t.} \tag{6}$$

$$\widetilde{n}'_{i,a+1} = (1 - H(\widetilde{z}_{i,a}))(1 + Q x_{i,a})\widetilde{n}_{i,a}, \tag{7}$$

$$y_{i,a} = A\epsilon_i \widehat{z}_{i,a} n_{i,a}^{\alpha}, \tag{8}$$

$$n_{i,a} = (1 - H(\widetilde{z}_{i,a}))\widetilde{n}_{i,a}, \tag{9}$$

where $H(\widetilde{z}_{i,a})$ is the endogenous separation rate defined as the fraction of workers who obtain a productivity draw below $\widetilde{z}$. The resulting first-order conditions can be combined into a "optimal hiring" condition and a condition implicitly defining the worker-specific

[9]Nevertheless, as a robustness check the online appendix shows that the results change little when considering realistic variation in firm survival rates.

productivity cutoff value. The optimal hiring condition is given by

$$\frac{\kappa x_{i,a}^{\gamma-1}}{Q} = \beta(1-\delta_a)\mathbb{E}\mathcal{J}'_{i,a+1}, \tag{10}$$

where $\mathcal{J}_{i,a} = \frac{\partial \Pi_{i,a}}{\partial \tilde{n}_{i,a}}$ is the beginning-of-period marginal value of a job for the firm. The optimal hiring condition therefore takes on a familiar form where the effective marginal costs of posting a vacancy are equal to the expected marginal benefits.

The condition defining the worker-specific productivity cutoff also balances the benefits and costs of changing the separation cutoff ($\widetilde{z}_{i,a}$). Intuitively, raising the separation cutoff will increase average worker-specific productivity of the remaining employees, but at the same time it will reduce the size of the workforce in the firm. This is formalized in the following condition which implicitly defines the separation cutoff and which anticipates that wages depend on the number of workers in the firm:

$$\frac{\partial y_{i,a}}{\partial \widetilde{z}_{i,a}} - \frac{\partial w_{i,a}}{\partial \widetilde{z}_{i,a}} n_{i,a} = -\frac{\partial n_{i,a}}{\partial \widetilde{z}_{i,a}} \frac{\mathcal{J}_{i,a}}{1 - H(\widetilde{z}_{i,a})}. \tag{11}$$

3.4. Firm entry and subsequent survival

Starting up a firm requires the sacrifice of a (potentially time-varying) cost $X > 0$ which captures initial costs of doing market research, formulating a business plan etc. This entry cost may also be interpreted as a reduced-form way of capturing the effects of (time-varying) financial frictions (see the online appendix). Upon paying this cost, a potential entrant chooses one business opportunity from a (fixed) finite measure of possibilities given by ψ_i. Each business opportunity allows for at most one successful startup. All startups begin with an employment value n_0, which is independent of the business cycle and the chosen technology type.

It is assumed that potential entrants cannot coordinate on which business opportunities to select. That is, not all individual opportunities are seized whereas others are pursued by several aspiring startups. This results in the number of startups within type i ($\omega_{i,0}$) being strictly smaller than both the number of business opportunities and the number of startup attempts (e_i). It follows that an attempted startup of a business type i is successful only with probability $\frac{\omega_{i,0,t}}{e_i}$. Unsuccessful startups exit before production takes place. This way of modeling firm entry is similar in spirit to models of innovation and research and development (see e.g. Klette and Kortum, 2004; Saint-Paul, 2002).

The coordination friction among aspiring startups is concisely summarized by an entry matching function, borrowed from the search and matching literature. This function relates the number of startups within each type to the respective number of startup attempts and business opportunities. It is assumed to be increasing in both arguments and to display constant returns to scale. In particular, $\omega_{i,0} = e_i^{\phi} \psi_i^{1-\phi}$, where $\phi \in (0, 1)$ is the elasticity with respect to the number of startup attempts.[10]

Free entry implies that in equilibrium the costs of starting up a firm of a specific technology type equal the expected benefits

$$X = \frac{\omega_{i,0}}{e_i} \Pi_{i,0}, \text{ for } i = 1, 2, .., I, \tag{12}$$

The free entry conditions (12) imply that entry happens in all technology types. This is because technology types associated with high firm values will attract many potential new firms which in turn lowers the probability of successfully starting up. This, in turn, encourages entry of firms into technology types with lower firm values. In equilibrium, aspiring entrants are indifferent between all of the business opportunities, akin to models of directed search.

Furthermore, while firm values may be relatively insensitive to aggregate shocks of reasonable magnitudes, the entry elasticity parameter ϕ enables the model to nevertheless match the firm entry patterns in the data even with aggregate productivity being the only source of exogenous variation. Without this feature, fluctuations in startups would be counterfactually

[10]See Saint-Paul (2002) for a similar specification in the context of firms' research and development.

small due to the relative stability of firm values (see e.g. Samaniego, 2008).

Finally, firm exit is governed by an exogenous, but age-dependent probability δ_a. The evolution of the mass of firms of type i and age a is thus given by

$$\omega'_{i,a+1} = (1-\delta_a)\omega_{i,a} \quad \text{for i=1,2,...,I and a} \geq 0. \tag{13}$$

3.5. Wage setting

The presence of a frictional labor market results in employment relationships being characterized by positive surplus values over which workers and firms bargain. The assumption of decreasing returns to scale in firms' production functions implies that these surplus values will depend on the number of workers employed in a given firm.

In this paper, wage setting is assumed to be conducted under the bargaining solution of Stole and Zwiebel (1996) which generalizes the Nash bargaining solution to a setting with decreasing returns to scale.[11] Under this bargaining setup, the resulting wage is the same as under Nash bargaining, but where the bargaining happens over the marginal surplus. The resulting bargained wage takes on the following form

$$w_{i,a} = \eta\left(\frac{\alpha y_{i,a}/n_{i,a}}{1-\eta(1-\alpha)} + \frac{\kappa(\gamma-1)}{\gamma}x_{i,a}^{\gamma} + \Phi\right) + (1-\eta)b, \tag{14}$$

where $\Phi = \kappa\frac{V}{U}\sum_i\sum_a \frac{\omega_{i,a}v_{i,a}}{V}x_{i,a}^{\gamma-1}$. The bargained wage has the familiar interpretation of being a weighted average of the marginal product, savings on hiring costs and the flow income in unemployment. Moreover, for linear hiring costs and constant returns to scale ($\gamma = 1$ and $\alpha = 1$) the above expression collapses to the standard Nash wage.

Finally, as has been pointed out in other models with frictional labor markets and heterogeneous firms, heterogeneity per se does not necessarily lead to greater amplification of shocks (see e.g. Kaas and Kircher, 2015; Hawkins, 2011). Therefore, the model will be al-

[11]While the Stole-Zwiebel bargaining solution has not yet been verified empirically, this paper adopts it for its conceptual proximity to models with constant returns to scale and Nash bargaining.

lowed to display a certain degree of wage rigidity, which is known to be one of the solutions to the labor market volatility puzzle. In particular, wages in individual firms will be a weighted average of the Stole-Zwiebel wage derived in (14) and its steady state counterpart. The weight (ζ) given to the steady state wage then governs the degree wage rigidity in the model.[12]

3.6. Aggregate shocks and market clearing

There are two aggregate shocks present in the model. The first is a shock to aggregate productivity (A) and the second is a shock to the entry cost (X). Time variation in entry costs is meant to capture driving forces particularly affecting startups, e.g. changes in the tightness of credit conditions, house prices or uncertainty.[13] Both are assumed to follow an AR(1) process in logs

$$\ln A' = \rho_A \ln A + \eta^A, \quad \eta^A \sim N(0, \sigma_A^2), \tag{15}$$

$$\ln X' = (1-\rho_X)\overline{X} + \rho_X \ln X + \eta^X, \quad \eta^X \sim N(0, \sigma_X^2), \tag{16}$$

where $e^{\overline{X}}$ is the steady state entry cost, ρ_j is the autocorrelation coefficient and η_t^j is the shock innovation assumed to be distributed identically and independently according to a normal distribution with zero mean and standard deviation σ_j, with $j = A, X$.

Given all the above the aggregate resource constraint can be written as

$$Y + bU = C + \sum_i \sum_a \omega_{i,a} \frac{\kappa}{\gamma} x_{i,a}^{\gamma} n_{i,a} + X \sum_i e_i, \tag{17}$$

where aggregate firm output $Y = \sum_i \sum_a \omega_{i,a} y_{i,a}$ together with home production are spent on consumption, the costs of posting vacancies and the cost of starting up firms. The elements of the household's budget constraint left to define are total wage income $W = \sum_i \sum_a \omega_{i,a} w_{i,a} n_{i,a}$

[12]In the quantitative exercises the resulting wages lie within the bargaining sets of workers and firms. In addition, the online appendix provides robustness exercises with respect to the wage rigidity parameter ζ.

[13]For examples of studies linking these factors to startup activity see Drautzburg (2014), Siemer (2014) and Schmalz, Sraer, and Thesmar (2016). See the online appendix for one explicit micro-foundation of financial frictions.

and aggregate profits $P = \sum_i \sum_a \omega_{i,a}(y_{i,a} - w_{i,a}n_{i,a} - \frac{\kappa}{\gamma}x_{i,a}^{\gamma}n_{i,a}) - X\sum_i e_i$. Finally, the number of unemployed is given by $U = L - \sum_i \sum_a \omega_{i,a}n_{i,a}$ with L being the size of the labor force.

Let $\mathcal{S} = \{A, X, \omega_{i,a}, \widetilde{n}_{i,a}\}_{i=1,2,\ldots,I,a\in\mathbb{N}}$ be the aggregate state consisting of both aggregate shocks, but also of the entire firm distribution (the mass and beginning-of-period employment levels of all firms). The entire firm distribution is a state variable because individual firms need to know and be able to predict the evolution of the matching probabilities on the frictional labor market in order to make their optimal decisions. The matching probabilities, however, depend on aggregate vacancies and unemployment which are in turn determined by the sum of employment levels and vacancies at individual firms.

Finally, the model's equilibrium is defined by

- individual firms' policy rules for available employment ($n_{i,a+1}$), the vacancy rate ($x_{i,a}$), workers' productivity cutoffs ($\widetilde{z}_{i,a}$), wages ($w_{i,a}$), masses of startup attempts (e_i), firm values ($\Pi_{i,a}$)

- the representative household's policy rule for aggregate consumption (C), the aggregate matching probability for firms (Q),

- the distribution of firms across types and ages ($\omega_{i,a}$) and the exogenous shocks (A and X)

which (for all firm types i and ages $a > 0$) solve the firms' optimization problems (6) and satisfy the law of motion for employment (7), the optimal hiring decision (10), the optimal separation rule (11), the bargained wage rule (14), the free entry conditions (12), the aggregate resource constraint (17), the matching function (2), the law of motion of firm masses across types and ages (13) and the laws of motion of the exogenous shocks (15 and 16).

4. Quantitative implementation

This section describes the parametrization procedure. In addition, the end of this section is devoted to inspecting the model's predictions along cross-sectional and business cycle dimensions and comparing them to the observed patterns in the data. The online appendix then provides robustness exercises to various model parameters.

4.1. Parametrization

All model parameters, except for those pertaining to the two aggregate shocks, are calibrated using aggregate and BDS data for the sample period 1979-2012. The parameters, and time-paths, of the two structural shocks are then estimated using Maximum Likelihood on data for the unemployment rate and the number of startups.[14] To facilitate the exposition, the parametrization strategy as a whole is discussed first. Next, the exposition moves on to parameters found in standard search and matching models and finally it describes parameters pertaining to firm dynamics. Table 1 summarizes all the parameter values, their targets and the respective model predictions. In the benchmark specification of the model, the number of firm types is set to $I = 5$ and the maximum firm age is set to $K = 100$.

The parametrization strategy is based on the interpretation of the entry cost shock as "residual" variation in startups not driven by fluctuations in firm values. In other words, it is meant to capture forces particular to startups and not directly affecting incumbent firms. Therefore, the entrant elasticity ϕ is chosen such that the model generates a relative volatility of startups with respect to output of 2.5 as in the BDS data without exogenous variation in entry costs.[15] The estimation procedure will then assign all other variation in startups to the entry cost shock.

[14]Both time-series are linearly detrended in order not to impose artificial autocorrelations into the estimation. Nevertheless, the results are similar when estimated on HP-filtered data. While the unemployment rate is available at a quarterly frequency, the BDS data offers only annual information. This, however, can be conveniently handled by the Kalman filter used to construct the likelihood function. More details on the estimation procedure are in the online appendix.

[15]In a similar fashion, the calibration of parameters related to the volatility of separations and wages (discussed below) is also based on a model without entry cost shock variation.

The model period is assumed to be one quarter and the discount factor (β) is therefore set to 0.99 implying an annual interest rate of 4 percent. The size of the labor force (L) is set such that the resulting steady state unemployment rate is equal to 6.3 percent. The level of match efficiency (m) is set to target a job filling probability of 0.71 as in den Haan, Ramey, and Watson (2000). The matching elasticity (μ) takes on the value of 0.6, which is in the middle of values found in the literature (see Petrongolo and Pissarides, 2001, for an overview). The value of home production (b) is set to match the overall separation rate of 4.6 percent, taken from the Current Population Survey. The intuition for this target is that for a given parametrization of worker-specific productivity, a higher value of b reduces the surplus of the employment relationship implying a greater chance of separation (see Elsby and Michaels (2013) for a similar calibration strategy). The resulting value gives rise to an average replacement rate of 66 percent. Following most of the existing literature, it is assumed that workers and firms have equal bargaining power, i.e. $\eta = 0.5$. The distribution of worker-specific productivity shocks is assumed to be logistic with mean 1 and scaling parameter σ_H which is set such that the volatility of separations relative to output volatility is 3.4 as in the data. Given all the model parameters, the costs of posting vacancies are determined via the optimal hiring condition. The resulting value implies overall costs of less than 1 percent of output. Last, as anticipated already in Subsection 3.5, the model requires an additional (exogenous) source of wage rigidity in order to be able to match the volatility of labor market variables. As explained earlier, wages are a weighted average of the Stole-Zwiebel bargained wage and the respective steady state counterpart. The weight (ζ) is chosen such that the resulting wage elasticity with respect to labor productivity is 0.5 as in Hagedorn and Manovskii (2008).

Firm dynamics parameters which are common to all firms include the returns-to-scale parameter is set to 0.85, within the range of estimates in Basu and Fernald (1995), Basu (1996) and Basu and Kimball (1997). Vacancy posting costs are assumed to be quadratic in the vacancy rate, i.e. $\gamma = 2$, as in Gertler and Trigari (2009). In order to capture the

Table 1: Calibrated parameters

	parameter	value	target/source	model
β	discount factor	0.99	4% annual interest rate	
L	labor force	785.8	$u = 6.3\%$, BLS	6.3%
m	matching efficiency	0.605	$Q = 71\%$, den Haan, Ramey, and Watson (2000)	71%
μ	match elasticity	0.600	Petrongolo and Pissarides (2001)	
b	value of non-work activities	0.477	$\rho_T = 4.6\%$, CPS	4.4%
η	worker bargaining power	0.500	symmetric bargaining	
κ	vacancy posting cost level	2.45	optimal hiring condition	
ζ	wage rigidity parameter	0.85	$\epsilon_{w,z} = 0.5$, Hagedorn and Manovskii (2008)	0.46
μ_H	idio-syncartic productivity shock, mean	1.000	normalization	
σ_H	idio-syncratic productivity shock, shape parameter	0.240	$\sigma(ln(\rho_T))/\sigma(\ln(Y)) = 3.3$, CPS	3.1
α	returns-to-scale parameter	0.850	Schaal (2017)	
γ	convexity of hiring costs	2.000	Gertler and Trigari (2009)	
$\overline{\delta}_0$	firm exit level parameter	0.005	exit rate age profile, BDS	
$\overline{\delta}_1$	firm exit age parameter	0.330	exit rate age profile, BDS	
n_0	initial employment stock	6.000	average entrant size of 6.2, BDS	6.4
$\overline{X}$	firm entry cost level	2.351	success probability 34%, BDS approximation	34%
ϕ	entry elasticity	0.735	$\sigma(\ln(\sum_i \omega_{i,0}))/\sigma(\ln(Y)) = 2.5$, BDS	2.5
σ_Z	standard deviation, aggregate productivity	0.009 (0.001)	estimated	
ρ_Z	AR coefficient, aggregate productivity	0.925 (0.021)	estimated	
σ_Z	standard deviation, startup cost	0.015 (0.002)	estimated	
ρ_Z	AR coefficient, startup cost	0.876 (0.050)	estimated	
ϵ_i	firm productivity	1	1.47 1.82 2.23	2.95
$100\frac{\omega_{i,0}}{\sum_i \omega_{i,0}}$	fraction of startup types (%)	92.7	5.9 0.9 0.4	0.1

Notes: The table shows the calibrated parameters (based on a model without entry cost shocks), their respective targets or sources and the last column shows the value of the statistic as predicted by the calibrated model. ρ_T is the total separation rate, Y is aggregate output, $\overline{w}$ denotes the average wage, $\epsilon_{w,z}$ is the elasticity of wages with respect to productivity and u is the unemployment rate.

age-dependent character of firm exit, it is assumed that the exit probabilities are given by $\delta_a = \overline{\delta}_0 + \overline{\delta}_1/a$ for $a < K$ and $\delta_K = 1$. The coefficients $\overline{\delta}_0$ and $\overline{\delta}_1$ are then chosen to match the observed exit rates in the BDS data, conditional on firm age. All firms are assumed to start with an employment level of n_0 which is set to match average firm size of startups. Finally, the mass of entrants is normalized to 1. Given all the above choices, the level of the entry cost does not affect any variables in the model, except for the number of startup attempts, or equivalently the startup probabilities. Interpreting the startup probability as the survival rate during the first period, the entry cost is set such that the model matches a probability of success of 34 percent implied by the calibrated firm exit function.

Finally, turning to firm-specific parameters, the productivity levels (ϵ_i) and associated masses of business opportunities (ψ_i) are pinned down by requiring the model to have a realistic firm size distribution. Specifically, the productivity parameters are set such that the model matches the employment shares within five size brackets found in the BDS data (with the least productive firm type normalized to 1): $1-49$, $50-249$, $250-999$, $1,000-9,999$ and $> 10,000$. Similarly, the firm mass parameters ψ_i are chosen to match the firm shares within the above five size brackets.

4.2. Cross-sectional and business cycle properties

Before moving on the the main results of the paper, this subsection documents that the presented structural model is consistent with the data along several cross-sectional and business cycle dimensions.

4.2.1. Firm size and age distributions

The previous subsection described that the firm-specific parameters of the model are set to match the firm size distribution observed in the BDS data. Table 2 shows these targets, together with firm and employment shares of new, young and old firms. The table documents that the structural model does well not only in capturing the firm size distribution but also the empirical firm age distribution which was not targeted by the calibration.

Table 2: Firm age and size distributions

shares (in %)	new	young	old	1–49	50–249	250–999	1,000–9,999	> 10,000
				BDS data				
firm	10.6	31.7	57.7	95.04	4.15	0.57	0.22	0.02
employment	3.0	13.2	83.8	37.8	21.1	12.5	17.4	11.2
				model				
firm	11.3	31.0	57.7	94.66	4.44	0.60	0.27	0.03
employment	2.8	17.4	79.8	33.4	19.6	11.4	21.4	14.2

Notes: The table reports employment and firm shares within firm age (first 3 columns) and size (last five columns) bins. "New" firms are less than one year old, "young" firms are between 1 and 5 years of age, and "old" firms are defined as more than five years of age. All values are in percent.

4.2.2. Net job creation in young and old firms

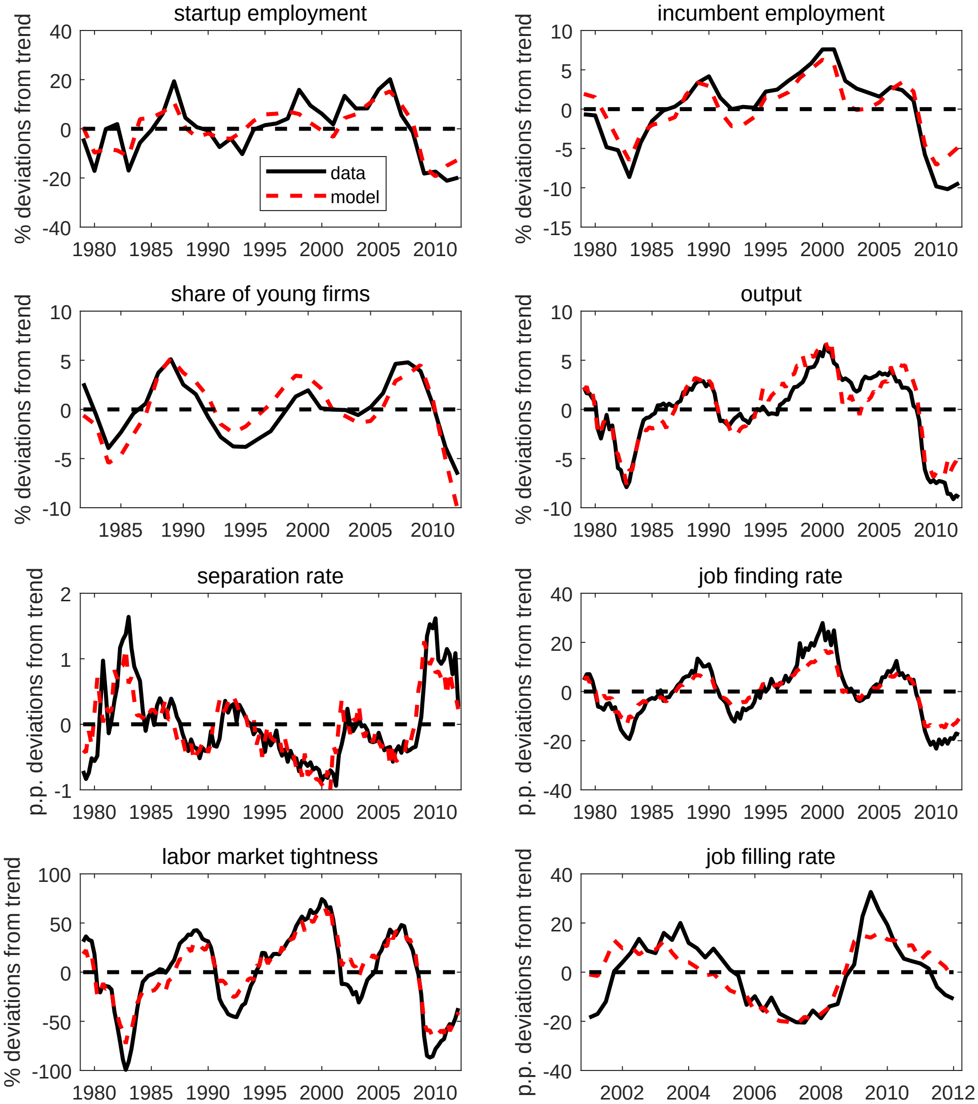
Because the model will be used to evaluate the labor market implications of a lost generation of (young) firms, it is important that the model replicates net job creation rates conditional on age in the cross-section.[16] As in the data (see e.g. Haltiwanger, Jarmin, and Miranda, 2013), net job creation in the model declines quickly with age from about 20 percent for one year old firms to about 5 percent for five year old businesses.

4.2.3. Time-paths of variables not in estimation

Finally, Figure 2 shows that the model performs well in terms of capturing the dynamics of variables not directly used in the estimation. In particular, the top row shows that the model does well in capturing the employment dynamics of new and incumbent firms. At the aggregate level, labor market tightness is slightly less volatile than in the data and this spills over to the job finding and filling rates. This is due to a lower volatility of vacancies predicted by the model. The reason is that a large fraction of vacancies is posted by young

[16]Following Haltiwanger, Jarmin, and Miranda (2013), the net job creation rate is defined as $g_{i,a} = (n_{i,a,t} - n_{i,a-1,t-1})/(0.5(n_{i,a,t} + n_{i,a-1,t-1}))$.

Figure 2: Actual and model-predicted time-paths



Notes: actual and model-predicted time-paths of variables (based on an estimation using the unemployment rate and the number of entrants as data inputs). The employment in new and incumbent firms as well as the share of young firms is based on BDS data, the separation and job finding rates are taken from the CPS, labor market tightness is defined as vacancies (using the time series constructed by Barnichon (2010)) over unemployment and finally the job filling rate is the vacancy yield from Job Openings and Labor Turnover Survey. All data are linearly detrended.

firms which are characterized by being far away from their optimal size, and thus having relatively large marginal products of labor. This in turn, increases their surplus resulting in relatively lower sensitivity to aggregate shocks.[17] Finally, it is worth noting that the strong

[17]This contrasts the results of Elsby and Michaels (2013) who find that a labor market model with heterogeneous firms does well in replicating the volatility of labor market variables (even without exogenous

Beveridge curve relation in the data is preserved by the model (a correlation of -0.91).

5. Aggregate dynamics and lost generations of firms

The purpose of this section is to understand the quantitative effects of a lost generation of firms within a general equilibrium model in which existing firms are free to adjust their hiring behavior in response to the lack of jobs created by young firms. Towards this end, the next subsection presents a set of impulse response functions (IRFs) to a one-time shock to the entry cost which results in a 30 percent drop in the number of entrants (the magnitude observed in the Great Recession). Note that this shock immediately reverts back to zero and therefore any persistent dynamics are generated endogenously.

Then, the structural model is used to estimate the time-path of the two aggregate shocks (productivity and entry cost) during and in the aftermath of the Great Recession. This enables the model to quantify the relative contribution of a strong but "standard" recession shock and a shock that specifically affected entrants. Moreover, it is possible to use the model and the estimated shocks to conduct counterfactual scenarios and ask how the economy would have developed had the number of startups not fallen as much.

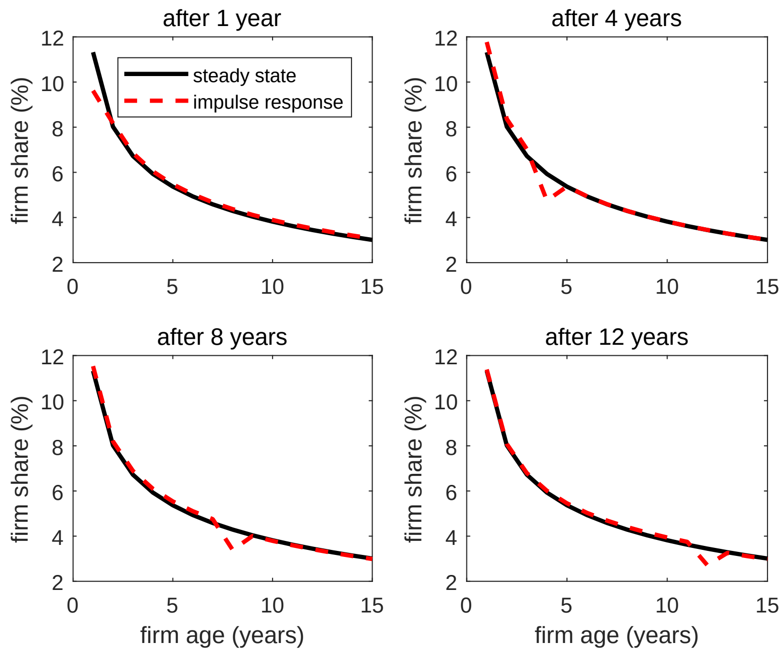
5.1. Impulse response functions

5.1.1. Firm-age distribution

Figure 3 shows the steady state firm-age distribution and the impulse response of the distribution to an exogenous increase in the startup cost. The response of the distribution is depicted 1, 4, 8 and 12 years after the shock subsides. The figure makes clear that even a one-time shock to entry may affect the economy many years after it subsides as the affected cohort of firms grows old and works its way through the firm-age distribution. Importantly, it also makes clear that the slight increase in firm entry following the shock is not strong

wage rigidity). The key difference is that in their framework there is no notion of a firm's life-cycle (i.e. they abstract from firm entry and exit) and thus all firms (including small ones) are close to their optimal sizes.

Figure 3: Impulse response of the firm-age distribution



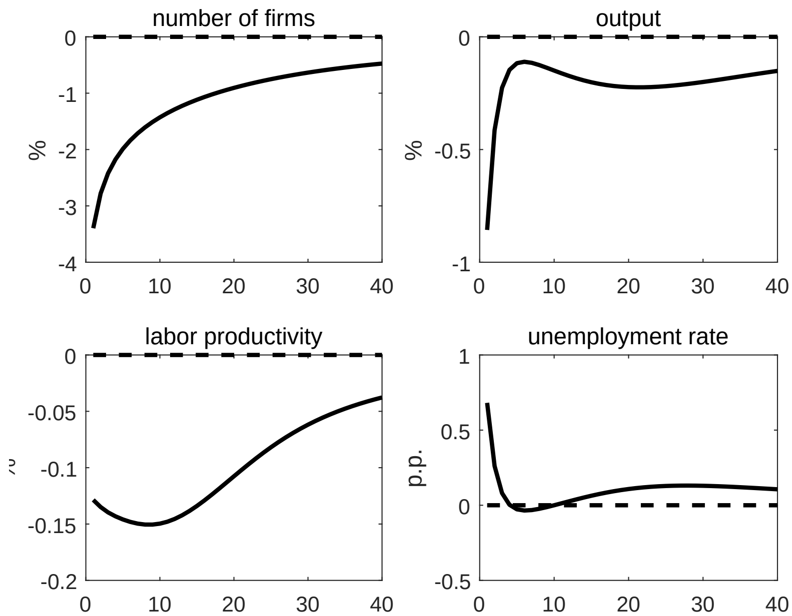
Notes: the steady state firm-age distribution (up until the age of 15 for ease of exposition) and the impulse response of the distribution 1, 4, 8 and 12 years after the shock to the entry cost subsides. The magnitude of the entry cost shock is such that the resulting drop in the number of startups resembles the one observed during the Great Recession.

enough to offset the initial drop in the number of startups. This provides a first hint that changes in firm entry may have long-lasting effects.

5.1.2. Output, productivity and unemployment

Figure 4 plots the IRFs (over a period of ten years) of the number of firms, output, labor productivity and unemployment to the one-time shock to the entry cost. The top left panel depicts the response of the number of firms which remains persistently below its steady state level owing to the fact that the affected cohort of firms dies out only slowly over time. The top right panel shows the response of output which after an initial almost 1 percent drop remains to be 0.2 percent below its steady state even ten years after the initial shock subsides. The output response stems both from a persistent decline in labor productivity (bottom left panel) and a persistent increase in the unemployment rate. The reason for the former is that the economy shifts away from young firms which are more productive than older businesses (because of the decreasing returns to scale).

Figure 4: Impulse response functions of aggregate variables



Notes: impulse response functions of the number of firms, output, labor productivity and the unemployment rate to a one-time shock to the entry cost. The magnitude of the entry cost shock is such that the resulting drop in the number of startups resembles the one observed during the Great Recession.

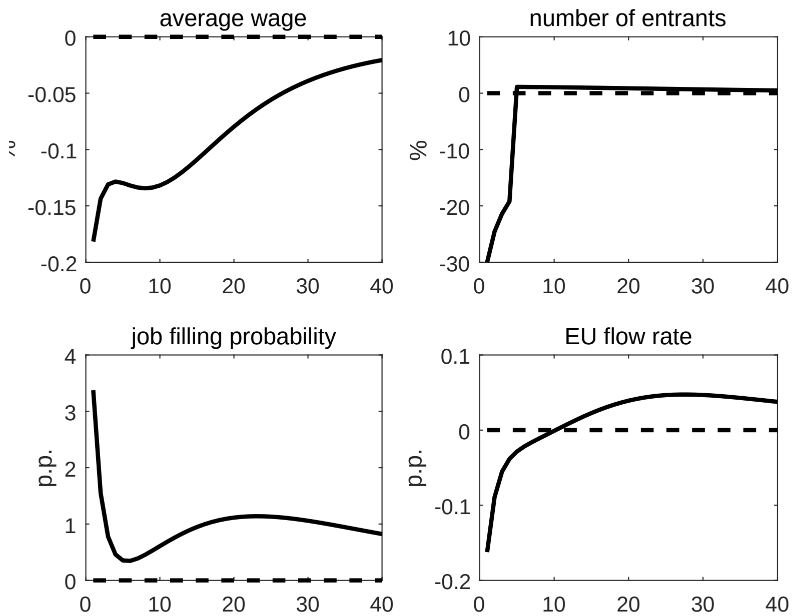
While the impact of the lost generation of firms is as persistent as in the simple counterfactual exercise in Section 2, the magnitude of the effect is substantially dampened. The reason for this stark difference is the presence of several general equilibrium effects present in the structural model. The following paragraphs inspect these general equilibrium (GE) effects in more detail.

5.1.3. General equilibrium effects

After an entry cost shock, the number of entrants falls and this reverberates through the firm-age distribution as a lost generation of firms (Figure 3). This brings with it a shift in the composition of firms away from young businesses lowering the number of posted vacancies and thus creating a slack labor market ($\theta = V/U$ falls). This effect serves to dampen the initial increase in the unemployment rate through four distinct channels depicted in Figure 5.

First, because of the slack labor market workers' outside options worsen and they settle

Figure 5: Impulse response functions: general equilibrium effects



Notes: impulse response functions of the number of entrants, average wages, the probability of hiring a worker and the separation rate to a one-time shock to the entry cost. The magnitude of the entry cost shock is such that the resulting drop in the number of startups resembles the one observed during the Great Recession.

for lower wages (top left panel).[18] This by itself increases the incentives for existing firms to hire more workers.

Second, because of the fall in wages, firm profits increase. The higher profitability increases the incentives to start up new businesses. The top right panel shows that after the initial (exogenous) drop in firm entry, the number of startups immediately jumps back and remains persistently above its steady state level. The higher number of startups in subsequent years helps offset some of the initial lack of job creation by young firms.

Third, while the slack labor market makes it harder for the unemployed to find jobs, it makes it relatively easier for incumbent firms to find workers (bottom left panel). This again increases hiring incentives as the effective costs of posting vacancies (which include the expected duration of a vacant position) fall.

[18]The drop in the average wage also reflects the shift away from younger more productive firms (because of the decreasing returns to scale).

Fourth, it is easier to retain workers because their outside option (of finding a new job) worsens. This channel resembles the "insulation effect" of lower job creation put forward by Caballero and Hammour (1994). This insulation effect is, however, quantitatively relatively weak. The majority of the resulting separation rate decline (bottom right panel) is because of a composition shift away from young firms which have on average higher rates of shutting down.[19] All the above channels serve to dampen the increase in the unemployment rate following the lack of job creation by startups (and subsequently young firms).

5.1.4. Decomposing the unemployment rate response

To quantify the relative strength of the above channels, it is possible to construct a counterfactual unemployment rate response where the general equilibrium effects are "shut down". The solid black lines in Figure 6 indicate, respectively, the benchmark IRF ("flexible") and the IRF without the GE effects ("fixed").

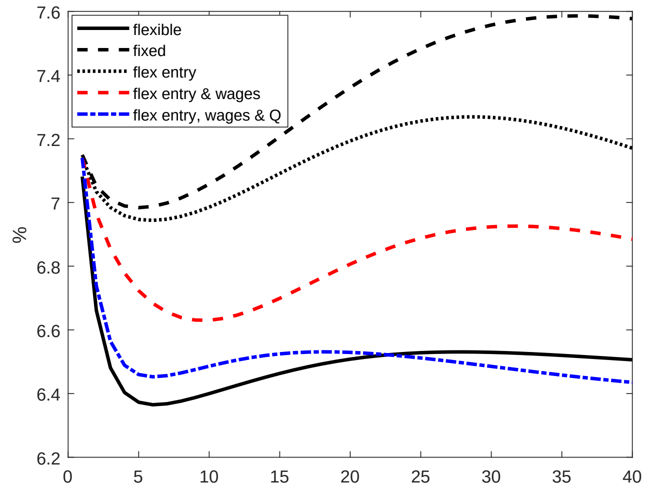
The "fixed" counterfactual is constructed by holding firm entry, wages, the probability of hiring a worker and the separation threshold at their respective steady states and by dropping the equilibrium conditions determining these endogenous variables (12), (14), (11) and (11) from the system of equations defining the model. In the same way, it is possible to construct counterfactuals where only a subset of the GE channels is shut down.

Without the GE effects unemployment is almost 1.5 percentage points above its steady state even ten years after the shock subsides. This is reassuringly similar to the unemployment response in the simple counterfactual (which also does not allow for GE effects) presented in Section 2.

The remaining lines characterize the relative strengths of the four general equilibrium channels. Specifically, "flex entry" allows for firm entry to vary (i.e. to overshoot following

[19]The subsequent increase in the separation rate is also driven by composition changes as the share of young firms rises above its steady state after about 3 years. Recall that firm entry overshoots from the second period onwards.

Figure 6: Unemployment rate response: benchmark and counterfactuals



Notes: impulse response of the unemployment rate in the benchmark specification ("flexible") and four counterfactuals. "fixed" refers to the case when the job filling probability, wages and separation rates are fixed and when entry reverts back to steady state after the initial shock. "flex entry" is like the "fixed" case, but entry responds endogenously. "flex entry & wages" is like the "flex entry" case, but in addition firms' wages respond endogenously. Finally, "flex entry, wages & Q" is like the "flex entry & wages" case, but in addition the job filling probability responds endogenously. The magnitude of the entry cost shock is such that the resulting drop in the number of startups resembles the one observed during the Great Recession.

the exogenous drop), "flex entry & wages" lets not only firm entry but also wages to respond endogenously and finally "flex entry, wages & Q" assumes only a fixed separation rate while the remaining three channels are left to endogenously adjust.

Figure 6 therefore quantifies to what extent each channel contributes to the considerably milder response of unemployment compared to the simple counterfactual which does not allow for GE effects. The closer the given counterfactual response is to the "flexible" response, the stronger the given general equilibrium channel. Therefore, the figure suggests that while the endogenous response of firm entry and wages (red dashed line) is an important channel that mutes the increase in the unemployment rate, the GE effects operating through the frictional labor market are equally strong. The dominant channel in the labor market is the changing probability of hiring workers.

5.2. *The lost generation of firms in the Great Recession*

This subsection uses the estimated time-paths of the model variables to inspect the relative contribution of the two structural shocks during the Great Recession and to conduct counterfactual scenarios determining how the U.S. economy would have evolved had it not been for the strong decline in firm entry in the past years.

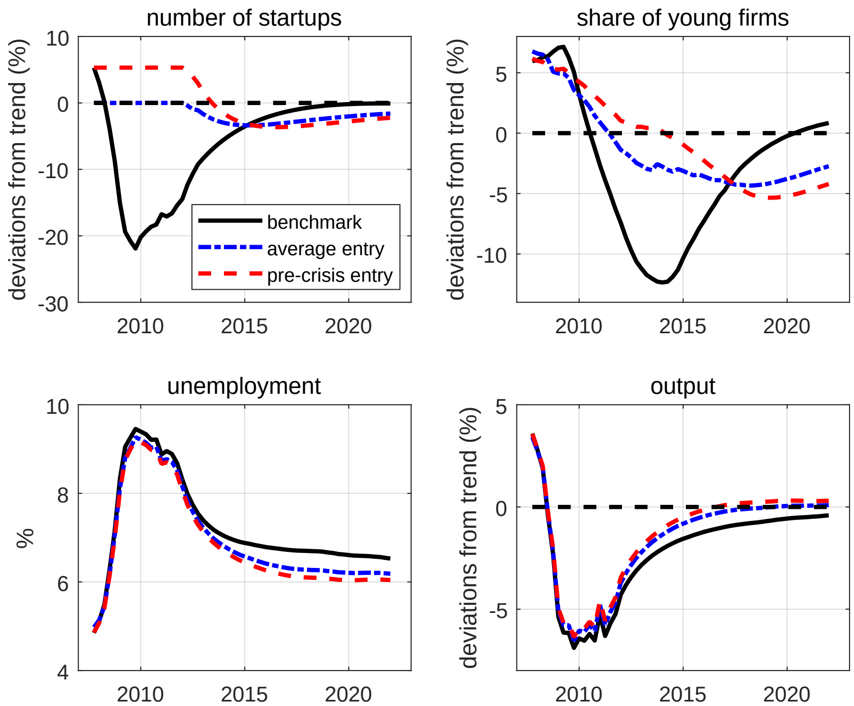
First, we begin by discussing to what extent the drop in firm entry was caused by forces specific to startups and to what extent it was a reaction to a "standard" but large recession. Towards this end, I simulate the benchmark economy using the estimated aggregate productivity shock alone, i.e. without exogenous variation in the entry cost. In this case, firm entry would have still been about 10 percent below its trend in 2010. This means that roughly 60 percent of the subdued entry during the Great Recession was driven by a "standard" response to a particularly strong recession while the rest can be attributed to factors specific to startups.

Next, we consider two counterfactual scenarios: one in which the entry cost shocks are such that firm entry during and in the aftermath of the Great Recession remains at its trend level ("average entry") and one in which it remains at its pre-crisis level ("pre-crisis entry"). After the end of the sample period, the economies are left to converge to the steady state for 10 years (i.e. both shocks are fixed to their unconditional means until 2022Q1). Figure 7 plots the actual and counterfactual time-paths for the number of startups, the share of young firms, unemployment and output.

Both counterfactual scenarios have similar quantitative effects. The immediate consequences of a drop in firm entry on the aggregate economy are small as the general equilibrium effects allow incumbent firms to compensate for the loss of job creation by startups. Specifically, the unemployment rate peak would have been only about 0.2 percentage points lower.

However, in later years the negative effect of the lost generation of firms strengthens. The reason is that the missing entrants generate fewer older firms in the future (apparent

Figure 7: Time-paths of variables under alternative scenarios



Notes: actual and couterfactual time-paths of the number of startups, the share of young firms, output and the unemployment rate using estimated shocks to aggregate productivity and the entry cost. "Average entry" ("pre-crisis entr") refers to the case when the entry costs are such that the model generates firm entry equal to its trend (pre-crisis) level during and in the aftermath of the Great Recession. The data used in the estimation (unemployment rate and the number of entrants) runs between 1977Q1 and 2012Q1. Thereafter the model is left to converge to its steady state (i.e. both aggregate shocks are set to zero).

from Figure 3) which account for the bulk of aggregate employment. Even though firm entry recovers somewhat faster in the benchmark economy compared to the counterfactuals (as explained in Figure 5) it is not enough to compensate for the depressed employment in mature firms. The reason is that the employment gains of the additional startups kick in only in later years as these new firms grow older and larger.

Therefore, the missing generation effect creates a persistent dent in the employment potential of the economy essentially raising the "natural" rate of unemployment. Specifically, had the number of startups remained constant during and in the aftermath of the Great Recession then the unemployment rate would have been 0.5 percentage points lower even 10 years after the crisis. Similarly, output would have reverted back to trend 4-6 years earlier compared to the benchmark economy.

5.3. Discussion

This paper asks whether periods of subdued firm entry can have long-lasting negative effects on the aggregate labor market. The results suggest that they can but, as long as incumbent firms play a similar role as new businesses (in this case job creation), the short-run aggregate impact of a lost generation of firms is limited. This subsection briefly discusses the robustness of this result.

5.3.1. Composition of startups

Sedláček and Sterk (2017) investigate the potential of startups to grow large over the business cycle in the BDS data. They find that recessionary periods give rise to firms that are smaller and that remain small even several years into the future. Using an estimated heterogeneous firm model they conclude that changes in the composition of startups (with respect to their potential to grow large) are extremely important for cohort-level employment variation, but they also help shape the medium- to long-run fluctuations in aggregate employment. This suggests that while changes in the number of new firms alone are unlikely to have persistent effects, a greater concern may lie in the selection effects among startups over the business cycle.

5.3.2. Mismatch

In the current model all workers are ex-ante identical and therefore young and old firms play similar roles in job creation. If, however, new (young) firms disproportionately employ a particular type of workers, then the general equilibrium effects mentioned in this paper may be weakened because older firms would not be tempted to hire from the larger unemployment pool. Even though such increased mismatch could strengthen the impact a lost generation of firms has on the aggregate labor market, current evidence suggests only a relatively modest role for mismatch in explaining unemployment dynamics (see e.g. Sahin, Song, Topa, and

Violante, 2014).[20]

5.3.3. Postponing of entry

It is possible that firms that do not enter during recessions are simply waiting until business conditions improve and they will enter in the subsequent recovery phase. In this sense startups are not lost, they are merely postponed. Interestingly, the presented model does indeed predict that an exogenous drop in firm entry is followed by an overshooting in the following periods. The magnitude of this effect is, however, rather weak. Also in the data there seems to be little evidence in support of a strong overshooting of startups following periods of subdued entry. Even in 2012, three years after the official end of the recession, the number of startups was still 28 percent below its pre-crisis level (and 16 percent below the sample average).

5.3.4. Labor force participation

The Great Recession was a time when the participation rate fell strongly (arguably contributing to some of the unemployment rate decline). It is beyond the scope of this paper to investigate the quantitative impact this additional margin of adjustment may have in the medium-run. However, given recent evidence that the participation margin can play an important role for unemployment rate fluctuations (see Elsby, Hobijn, and Sahin, 2015), incorporating this feature seems an interesting avenue for future research.

5.3.5. Wage rigidity

The baseline model is characterized with rigid wages. While this improves the model's business cycle properties, a concern may be that it also drives the dynamics of firm entry, the key channel of the model. However, as shown in the online appendix, wage rigidity has essentially no impact on the evolution of firm entry and in turn the number of firms. The reason for this is that the number of firms is determined by the free entry condition which, in

[20] Sedláček (2016), however, shows that the contribution of mismatch to unemployment rate fluctuations is substantially larger when job seekers from outside unemployment are taken into account.

turn, depends on firm values (i.e. the discounted net present value of future profits). While different degrees of wage rigidity affect the short-run dynamics of profits, they do little to firm values.

6. Conclusion

This paper documented that the firm-age distribution exhibits cyclical changes (shifts away from young firms during recessions) which help shape aggregate employment dynamics. A simple exercise would suggest that the lost generation of firms observed during the Great Recession may have substantial negative effects on the aggregate labor market for many years to come.

However, this conclusion is based on unchanged behavior of incumbent firms. A general equilibrium model of firm dynamics and a frictional labor market shows that accounting for feedback effects of firm entry into the employment behavior of incumbent firms is important for the magnitude of the aggregate response to the drop in firm entry. However, despite the strong general equilibrium effects, periods of subdued firm entry remain to negatively impact the economy in the medium- to long-run.

What is beyond the scope of this paper are statements about efficiency and therefore policy recommendations. Given the presence of ex-ante heterogeneity (and the imposed wage rigidity), it is unlikely that the simple Hosios condition would make the competitive equilibrium efficient (see e.g. Hawkins, 2014). Therefore an analysis of the efficiency conditions, together with the additional channels pointed out in the discussion seem to be interesting directions for future research.

Acknowledgements

This paper is a substantially revised part of my Ph.D. dissertation written at the University of Amsterdam. I thank my advisor Wouter den Haan for his support and guidance. I am thankful to my editor, Ricardo Reis, and two annonymous referees for excellent comments which improved the paper considerably. I am also grateful for comments from Eric Bartelsman, Christian Bayer, Michael Elsby, Bart Hobijn, Philip Jung, Leo Kaas, Philip Kircher, Keith Kuester, Matija Lozej, Petr Mach, Fabien Postel-Vinay, Richard Rogerson, Nicolas Roys, Vincent Sterk, Christian Stoltenberg, Emily and Robert Swift, Antonella Trigari, Sweder van Wijnbergen and seminar participants at CERGE, the DNB, the ECB, the HNB, the Universities of Amsterdam, Autonoma de Barcelona, Bath, Bonn, Carlos III, Edinburgh, Louvain and Pompeu Fabra.

References

Barnichon, R. (2010): "Building a Composite Help-Wanted Index," *Economics Letters*, 109(3), 175–178.

Basu, S. (1996): "Procyclical Productivity: Increasing Returns to Scale or Cyclical Utilization?," *Quarterly Journal of Economics*, 111(3), 719–751.

Basu, S., and J. Fernald (1995): "Are Apparent Productive Spillovers a Figment of Specification Error?," *Journal of Monetary Economics*, 36(1), 165–188.

Basu, S., and M. Kimball (1997): "Cyclical Productivity with Unobserved Input Variation," NBER Working Paper No. 5915.

Caballero, R., and M. Hammour (1994): "The Cleansing Effect of Recessions," *American Economic Review*, 84(5), 1350–1368.

Campbell, J. (1999): "Entry, Exit, Embodied Technology, and Business Cycles," *Review of Economic Dynamics*, 1(2), 371–408.

Clementi, G. L., and D. Palazzo (2013): "Entry, Exit, Firm Dynamics and Aggregate Fluctuations," mimeo.

den Haan, W. J., G. Ramey, and J. Watson (2000): "Job Destruction and Propagation of Shocks," *American Economic Review*, 90(3), 482–498.

Drautzburg, T. (2014): "Entrepreneurial Tail Risk: Implications for Employment Dynamics," mimeo.

Elsby, M., B. Hobijn, and A. Sahin (2015): "On the Importance of the Participation Margin for Labor Market Fluctuations," *Journal of Monetary Economics*, 72, 64–82.

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 (2013): "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," NBER Working Paper no. 19134.

Gertler, M., and A. Trigari (2009): "Unemployment Fluctuations with Staggered Nash Wage Bargaining," *Journal of Political Economy*, 117(1), 38–86.

Gourio, F., T. Messer, and M. Siemer (2014): "What is the Economic Impact of the Slowdown in New Business Formation?," Chicago FED Letter no. 326.

Hagedorn, M., and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98(4), 1692–1706.

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.

Hawkins, W. (2011): "Do Large-Firm Bargaining Models Amplify and Propagate Aggregate Productivity Shocks?," mimeo.

——— (2014): "Bargaining with Commitment between Workers and Large Firms," *Review of Economic Dynamics*, 18(2), 350–364.

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

Kaas, L., and P. Kircher (2015): "Efficient Firm Dynamics in a Frictional Labor Market," *American Economic Review*, 105(10), 3030–3060.

Klette, T. J., and S. Kortum (2004): "Innovating Firms and Aggregate Innovatino," *Journal of Political Economy*, 112.

Lee, Y., and T. Mukoyama (2013): "Entry, Exit, and Plant-level Dynamics over the Business Cycle," mimeo.

Merz, M., and E. Yashiv (2007): "Labor and teh Market Value of the Firm," *American Economic Review*, 97(4), 1419–1431.

Mortensen, D., and C. Pissarides (1994): "Job Creation and Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61(3), 397–415.

Moscarini, G., and F. Postel-Vinay (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.

Petrongolo, B., and C. A. Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

Pugsley, B., and A. Sahin (forthcoming): "Grown-Up Business Cycles," *Review of Financial Studies*.

Sahin, A., S. Kitao, A. Cororaton, and S. Laiu (2011): "Why Small Businesses Were Hit Harder by the Recent Recession," Federal Reserve Bank of New York, Current Issues in Economics and Finance.

Sahin, A., J. Song, G. Topa, and G. Violante (2014): "Mismatch Unemployment," *American Economic Review*, 104(11), 3529–3564.

Saint-Paul, G. (2002): "Employment Protection, International Specialization, and Innovation," *European Economic Review*, 46(2), 375–395.

Samaniego, R. (2008): "Entry, Exit, and Business Cycles in a General Equilibrium Model," *Review of Economic Dynamics*, 11(3).

Schaal, E. (2017): "Uncertainty and Unemployment," .

Schmalz, M., D. Sraer, and D. Thesmar (2016): "Housing Collateral and Entrepreneurship," *Journal of Fincnace*, 72(1), 99–132.

Sedláček, P. (2016): "The Aggregate Matching Function and Job Search Behavior from Employment and Out of the Labor Force," *Review of Economic Dynamics*, 21, 16–28.

Sedláček, P., and V. Sterk (2017): "The Growth Potential of Startups Over the Business Cycle," *American Economic Review*, 107(10), 3182–3210.

Siemer, M. (2014): "Firm Entry and Employment Dynamics in the Great Recession," mimeo.

Stole, L., and J. Zwiebel (1996): "Intra-firm Bargaining Under Non-binding Contracts," *Review of Economic Studies*, 63(3), 375–410.