Firm Dynamics and Random Search over the Business Cycle

Richard Audoly*
UCL and IFS

This version: November 8, 2019
First version: January 2017

Job Market Paper
Latest version available here.

Abstract

I develop a tractable model of worker reallocation over the business cycle that emphasizes the interplay between firm heterogeneity and on-the-job search. I use this framework to study the role of search frictions in determining aggregate labor productivity following a large economic contraction. In the model, low productivity firms slow down worker reallocation after a recession because workers stay longer at these firms before finding a job at a more productive firm. Quantitatively, the model implies that worker reallocation has sizable and persistent negative effects on aggregate labor productivity. This channel is consistent with evidence from firm-level UK data suggesting that the allocation of workers to firms has downgraded in the aftermath of the Great Recession.

1 Introduction

The UK Great Recession is a stark example of an economic downturn with long lasting negative effects on the efficiency of production. Labor productivity growth, one measure of

*University College London and Institute for Fiscal Studies. Email: richard.audoly@ifs.org.uk. I am thankful to my supervisors at UCL, Fabien Postel-Vinay and Jan Eeckhout, for their constant advice and support. For many helpful comments and suggestions, I also wish to thank Giuseppe Moscarini, as well as seminar participants at SaM 2017 (Barcelona), the Young Economists Conference 2018 (Belgrade), EEA 2018 (Cologne), the SaM Workshop 2019 (Bristol), SaM 2019 (Oslo), and the 3rd Dale T. Mortensen conference (Aarhus). Any errors are my own.

this efficiency, has remained slower than its pre-recession level since 2008, a pattern known in Britain as the "labour productivity puzzle". Compared to its pre-recession trend, labor productivity was 3% lower in 2012, four years after the onset of the Great Recession. This figure was 4.5% in 2016. How can such long lasting effects be reconciled with the fact that other key aggregates – firm entry, investment – were back at their pre-recession level by 2012?

In this paper, I show that search frictions in the labor market can contribute to durably dampen labor productivity in the aftermath of a recession. The key idea is that search frictions slow down the reallocation of workers through a congestion effect. As unemployment rises, workers employed at low productivity firms wait for longer before finding a job at a more productive firm. This mechanism is supported empirically by the sharp drop in job-to-job transitions observed in the UK following the Great Recession. In an experiment designed to match the sudden rise in unemployment observed in Britain after 2008, I show that my framework can account for sixty percent of the drop in the worker allocation component of aggregate labor productivity measured in the data seven years after the onset of the recession.

A central contribution of this paper is to develop a tractable model of worker reallocation combining the three following features: aggregate shocks, search frictions, and firm dynamics. Aggregate shocks are a pre-requisite to studying the evolution of labor productivity over the business cycle. Search frictions constrain the transition of workers out of unemployment. In addition, I allow workers to search while employed, in the spirit of the random search framework with on-the-job search proposed by Burdett and Mortensen (1998). As highlighted by Barlevy (2002), on-the-job search can contribute to lower labor productivity in search models, as it gives unemployed workers the option to take bad jobs as a stepping-stone to get better ones later. Lastly, firm dynamics allows the selection of firms to adjust over the business cycle through entry and exit. It also gives a model counterpart to the micro-level concept of productivity available in the data – which is defined at the firm or establishment level – thus allowing to calibrate the model with firm level data.

To motivate the existence of a worker reallocation channel stemming from less productive firms, I build a labor productivity measure from the ground up, aggregating from British firm-level administrative data over the period 2000-2016. Importantly, this sample covers about a decade before and after the Great Recession – officially starting in 2008Q2 in the UK – the largest post-war economic contraction in Britain. I can then study separately the component of aggregate labor productivity coming from the productivity of individual firms and that arising from the allocation of labor to those firms. I find that the allocation of labor to firms is significantly downgraded following the Great Recession in the UK. Firm-level regressions confirm that the positive relationship between firms' labor productivity and their employment growth rate is weaker post-recession. I see these facts as evidence that less productive firms represent a potential sullying channel during the UK Great Recession and interpret them through the lens of the calibrated model.

In the model, firms make hiring and exit decisions and commit to a long term state-contingent wage contract, which they optimize to retain workers. I provide conditions on the primitives of the model such that the optimal contract is increasing in the firm's own productivity after all histories in equilibrium. This monotonicity implies that job-to-job transitions are always productivity enhancing, since better contracts are offered by more productive firms at all points of the business cycle. It also simplifies the solution of the model in the presence of aggregate shocks. With on-the-job search, the optimal contract itself depends on the whole distribution of offered contracts through the rate at which workers quit firms, a daunting fixed-point problem. Instead, the fact that contracts are increasing in firm productivity makes the distribution of workers across productivity level sufficient to characterize the firm's policies out of steady-state. I approximate this distribution with a set of its moments to solve the full model with aggregate shocks.

I calibrate the model to match a set of labor market and firm dynamics moments from British data. In doing so, I specifically include moments capturing workers' transition rates in and out of unemployment and between employers, as well as moments disciplining the selection of firms upon entry. These moments include the firm exit rate, as well as the persistence and dispersion of labor productivity at the firm level, which I obtain from the firm-level data. While not being targeted directly in the calibration, the model does a very good job at replicating the large concentration of employment in the largest firms observed in the data. This is important since any measure of aggregate productivity derived from firm data is shaped by this high level of employment concentration.

Given the calibrated model, I feed in a sequence of aggregate shocks triggering a sharp and prolonged increase in unemployment, akin to the UK experience during the Great Recession. The model generates firm dynamics and labor market aggregates in line with the data. It also replicates the drop in worker allocation to firms measured in the firm-level data after 2008. In the simulated recession, the model captures a large part of the persistence of this effect: it accounts for about sixty percent of the overall reduction of workers to firms found in the data by 2015.

To understand the interplay between firm dynamics and search frictions in the medium term, I leverage the model to decompose labor productivity into three components: (i) aggregate shock, (ii) firm selection, (iii) worker reallocation. I can then assess the role of each component in driving aggregate labor productivity in the simulated recession. While firm selection has a large positive effect on labor productivity in the short run, I find that the worker reallocation component has a medium-term negative impact on labor productivity. On net, it dominates the firm selection effect three years after the start of the recession.

The reason why the allocation of workers to firms is downgraded following the shock comes from on-the-job search. Firms have two margins to control the rate at which they grow their workforce in the model: the rate at which they hire and the rate at which workers quit. While the hiring rate drops everywhere in the productivity distribution, the rate at which workers quit their job shrinks primarily on the lower part of the firm productivity distribution, as these workers now compete with a larger pool of unemployed workers. This second effect dominates in the calibrated model. As a result, low productivity firms do not shrink as fast in the aftermath of the recession. This mechanism is supported in the data by the reduction of the job-to-job transition rate and the lower association between firm labor productivity and firm growth after 2008.

Related literature. This work is first related to the literature that documents the reallocation effects of recessions using micro-level data. Starting with Davis and Haltiwanger (1992), economic downturns have been recognized as periods of heightened job reallocation. Moscarini and Postel-Vinay (2012) and Fort et al. (2013) further point out the existence of heterogeneous responses in employment growth respectively by firm size and firm age. All of these papers focus on employment changes, but do not link them to productivity explicitly. Closer to the exercise in this paper, Foster et al. (2016) relate job reallocation to productivity in US manufacturing and study how this relationship varies over the business cycle. They find that the reallocation of jobs triggered by the US Great Recession has been less productivity enhancing than in previous contractions. I stress that all of these studies are based on firm- or establishment-level data. As such, a notion of firm is a prerequisite to any theory of worker flows that aims to address these facts.

Second, this paper contributes to the growing literature that combines firm dynamics with search frictions in the labor market. Models of firm dynamics traditionally center on adjustment costs in capital inputs or in the customer base (Khan and Thomas, 2013; Clementi and Palazzo, 2016; Sedláček and Sterk, 2017), but they maintain the assumption that labor markets clear. Conversely, the macro labor literature stresses the role of search frictions in the pairing of workers to jobs, but without a theory to aggregate these matches into a notion of firm (Mortensen and Pissarides, 1994; Shimer, 2005; Lise and Robin, 2017).

Amongst the recent papers integrating firm dynamics and search frictions on the labor market, my work merges three unique features: firm dynamics, random search with on-the-job search, and aggregate shocks. Gavazza et al. (2018), Kaas and Kircher (2015), and Sepahsalari (2016) abstract from job-to-job flows and center on heterogeneity in the efficiency of hiring over the business cycle. Elsby and Michaels (2013) and, more recently, Elsby and Gottfries (2019) present models of firm dynamics in a random search environment. The main difference with my framework is that these two papers do not allow for firm entry and exit. The life cycle of the firm is key to my analysis to allow for a potential cleansing channel through firm exit. Schaal (2017), finally, focuses on the impact of uncertainty shocks in a related model cast in a directed search framework. With respect to the random search environment considered in this paper, directed search implies that firms are indifferent between contracts in equilibrium, and as such job-to-job transitions need not be productivity enhancing. In my model, by contrast, more productive firms poach workers from less productive firms in equilibrium, clarifying the link between voluntary quits and improvements to productivity.

Two central contributions for the present paper are Moscarini and Postel-Vinay (2013) and Coles and Mortensen (2016). I maintain the assumption that firms can commit to state-contingent contracts from the former and build on the size independence result in the latter to characterize the optimal contract in the presence of firm-specific shocks. With respect to these papers, the key addition in my framework is to make firm entry and exit endogenous. Lastly, two recent papers, Engbom (2019) and Gouin-Bonenfant (2019), rely on a similar constant returns to scale assumption to simplify the firm's problem. They study, respectively, the impact of aging on dynamism in the labor market and of the dispersion in firm productivity on the labor share. Both focus on transitions between steady-states and abstract from aggregate shocks.

Outline. Section 2 documents novel facts on firm dynamics and labor productivity from British firm-level data. Section 3 introduces the model. Section 4 defines the equilibrium. Section 5 describes the calibration and numerical solution. Section 6 analyzes the reallocation properties of the model during a recession and Section 7 concludes.

2 Firm Productivity and Labor Productivity during the Great Recession

To document the interaction between firm productivity at the micro level and labor productivity at the macro level, I construct an index of labor productivity aggregating from the ground up, starting from firm-level data. This paper uses the Business Structure Database (Office for National Statistics, 2019), a dataset with yearly information on the universe of British firms between 1998 and 2016. Importantly, these data cover about a decade before and after the Great Recession – officially starting in 2008Q2 and ending in 2009Q2 in the UK – thus allowing to decompose labor productivity before the onset of the recession and during the recovery period.

Aggregate labor productivity from firm-level data. The Business Structure Database (BSD) combines several administrative sources to derive the employment, sales, industry, and birth year for each active firm in a given year. This last variable, in particular, allows to decompose labor productivity by firm age, and thus to study the role played by new firms in the evolution of labor productivity. In the subsequent analysis, I exclude Health and Education, which are mostly public in the UK, as well as a few industries whose aggregates do not line up with official UK statistics (Finance and Insurance, Mining and Quarrying). The final sample is made up of more than 33 million firm-year observations. Further details regarding sample selection, the construction of these variables, and the validation of the aggregate series derived from the BSD against official UK statistics can be found in Appendix B.1.

I follow Bartelsman et al. (2013) in defining the following industry-level labor productivity index

$$LP_t := \sum_i \underbrace{ES_{i,t}}_{\text{employment share}} \times \underbrace{LP_{i,t}}_{\text{firm labor productivity}} \tag{1}$$

with the employment share, $ES_{i,t}$, and labor productivity measure, $LP_{i,t}$, at firm i in period t given by

$$ES_{i,t} := \frac{\text{employment}_{i,t}}{\sum_i \text{employment}_{i,t}}, \quad LP_{i,t} := \ln\left(\frac{\text{sales}_{i,t}}{\text{employment}_{i,t}}\right). \tag{2}$$

While the Business Structure Database reports sales, and not value-added as more conventionally used to define labor productivity, its key advantage is to make this information available throughout the firm's life span and for all firms in the economy, independently of their size or sector.[1] Decker et al. (2018), who also study the relation between firm productivity and worker reallocation, define labor productivity at the firm-level similarly. In addition, I compare the sales-based labor productivity measure to a value added one from a different dataset of British firms in Appendix B.2. I find that the two measures are very strongly correlated (correlation coefficient of .925), in line with findings based on US data (Foster et al., 2001).

The analysis is carried out within industries to account for price differences (in product or input) across sectors.[2] To further abstract from trends in industry shares over the sample period, these industry-level measures are aggregated using time-invariant industry weights. These weights are defined as the labor share of each industry in the pooled sample.[3]

Worker reallocation and productivity during the UK Great Recession. Figure 1 briefly outlines the UK experience during the Great Recession. It shows that the sharp and prolonged increase in unemployment comes at the same time as a persistent reduction in aggregate labor productivity growth, a pattern known in the UK as the "labour productivity puzzle." While previous work has analyzed the role of cross-sector reallocation in accounting

[1]To be precise, the BSD covers all firms either above the VAT registration threshold or with at least one employee liable for income tax. These thresholds are not overly restrictive as they represent low levels of economic activity.

[2]In my baseline analysis, I use the "division" level of the British Standard Industrial Classification (SIC). With about seventy categories, this level of decomposition is roughly equivalent to three-digits sub-sectors in the NAICS system. The graphs shown in this section are similar when using a thinner SIC level.

[3]The decline in manufacturing is the most noticeable trend in the UK over the period. The share of manufacturing in employment falls from 16 to 8 percent between 1998 and 2016.

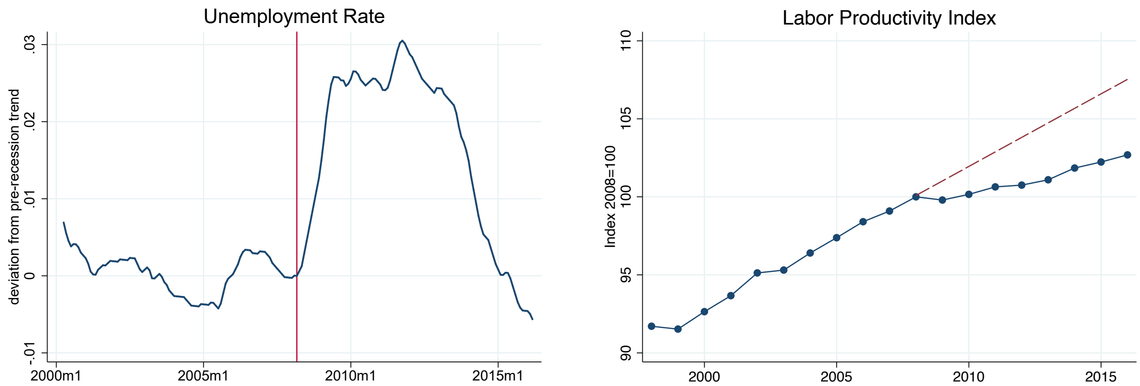

Figure 1: The Great Recession in the UK. Left: monthly unemployment rate (Office for National Statistics). Right: Labor productivity index, as defined in Equation (1). The index is computed separately for each industry and aggregated using time-invariant industry weights (see main text for details).

for the economy-wide pattern (Patterson et al., 2016), these data show that labor productivity growth also markedly slows down within sectors.

To assess the role of worker reallocation in accounting for the overall drop in labor productivity, I decompose LP_t as

$$\begin{aligned} LP_t &= \sum_i ES_{i,t} \times LP_{i,t} \\ &= \underbrace{\overline{LP}_t}_{\text{average firm productivity}} + \underbrace{\sum_i \left(ES_{i,t} - \overline{ES}_t\right)\left(LP_{i,t} - \overline{LP}_t\right)}_{\text{OP allocative efficiency measure}}, \end{aligned}$$

an equality referred to in the literature as the "OP decomposition" (Olley and Pakes, 1996).[4] In this last expression, the first term is the average (unweighted) productivity of firms in the economy. The second term measures how well labor is allocated to firms, as it increases as more firms with above average productivity also have a larger than average employment share.

The evolution of each of these terms after the Great Recession is depicted in Figure 2. They are both expressed in deviation from their respective pre-recession linear trend. Figure

[4]This equality follows directly from expanding the second term and noting that, by definition, $\sum_i ES_{i,t} = 1$.

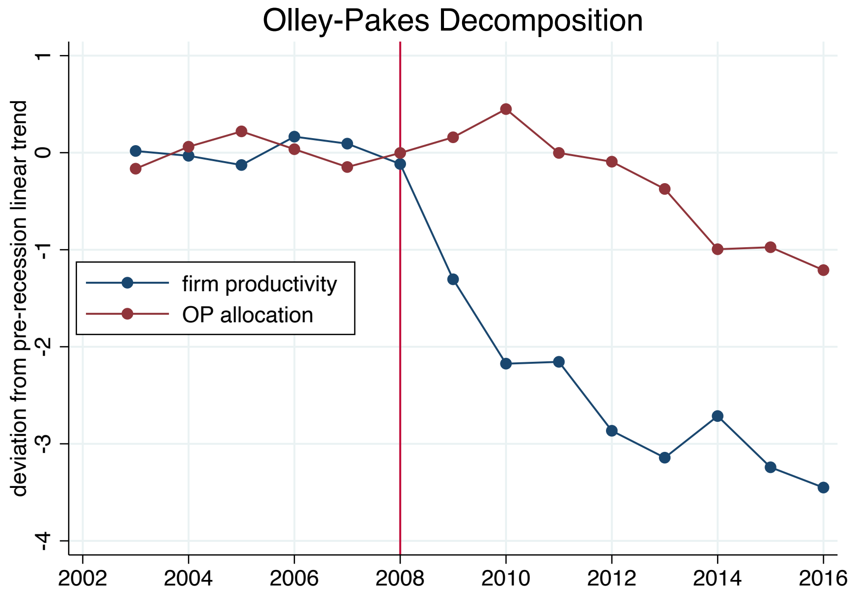

Figure 2: OP decomposition in the aftermath of the Great Recession. Each series is shown in deviation from its pre-recession linear trend.

2 shows that average firm productivity and the allocation of labor to firms have contributed to lower labor productivity growth in the aftermath of the recession. In particular, after a small increase right after the onset of the recession, the OP measure of allocation has kept moving down since 2010. By the end of the sample period, it represents about a fourth of the overall reduction with respect to the pre-recession trend in labor productivity.

Firm-level evidence. At the firm level, this aggregate pattern shows up as a lower association between firm labor productivity and their subsequent employment growth at the firm level. Table 1 shows regressions of the form

$$\Delta \ln n_{i,t+1} = \alpha LP_{i,t} + \beta \text{post}_t \times LP_{i,t} + \mu_{t,s} + \epsilon_{i,s,t}, \tag{3}$$

where $LP_{i,t}$ is firm labor productivity, "post" is an indicator for the years following the Great Recession, and $\mu_{s,t}$ a set of industry-year fixed effects. The coefficients (α, β) then measure the strength of the relationship between labor productivity and employment growth within

	$\Delta \ln n_{i,t+1}$ (continuers)	Firm Entry	Firm Exit
$LP_{i,t}$	0.065	-0.004	-0.020
	(0.000)	(0.000)	(0.000)
$LP_{i,t} \times \text{post}_t$	-0.017	-0.004	0.003
	(0.000)	(0.000)	(0.000)
Year-Industry FEs	Yes	Yes	Yes
Size Control ($\ln n_{i,t}$)	Yes	Yes	Yes
N (million)	22.01	27.08	27.78

Table 1: Reallocation during the Great Recession at the firm level. The specification is given in Equation (3). The first column is the change in employment in the next period (restricted to surviving firms). The second is a linear probability model for firms exiting in the next period (restricted to surviving and exiting firms). The third is a linear probability model for entering firms (restricted to firms entering in the current period and incumbents). Robust standard errors in parentheses.

an industry-year cell. The coefficient β shows that the positive association between firm labor productivity and employment growth drops by about a fourth post-recession, implying the average growth between a firm one standard deviation above and below the mean is about 2.3 percent lower. This finding suggests that employment growth is not as productivity enhancing as prior to the recession. The model developed in subsequent sections offers a rational in terms of search frictions for this firm behavior a the micro level. (Table 1 also reports results on the productivity of entering and exiting firms, showing minor differences between the pre- and post-recession periods.)

Job-to-job transitions. The slow down in labor productivity growth is further accompanied by a net reduction in job-to-job transitions, as shown in Figure 3. In the aftermath of the downturn, the rate of transitions between employers is about one-third lower than prior to the start of the episode. The recovery period is then characterized by both a much larger pool of unemployed workers and a drop in, potentially productivity enhancing, employer-to-employer transitions.

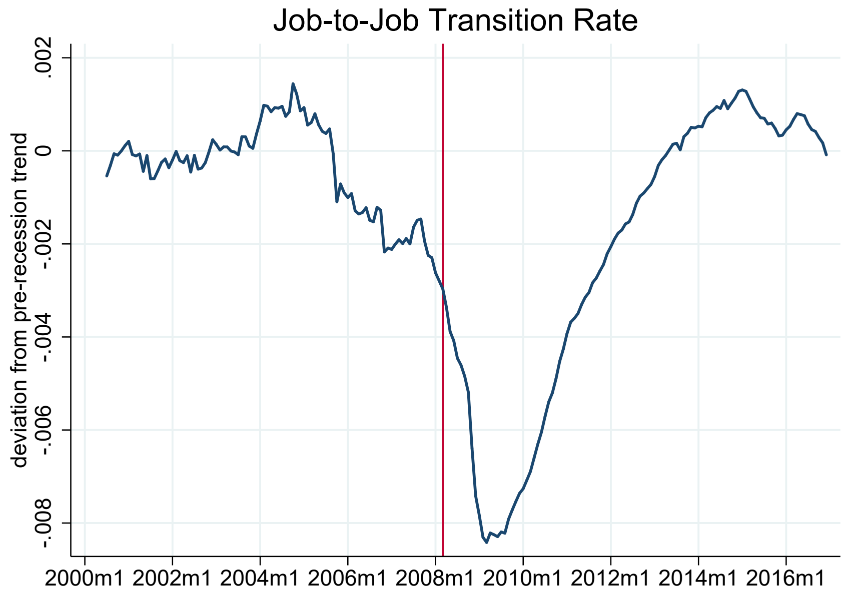

Figure 3: Job-to-job monthly transition rate derived from the British Household Panel Survey (Postel-Vinay and Sepahsalari, 2019). See Appendix B.3 for additional details on the construction of this series.

3 A Model of Firm Dynamics with On-the-Job Search

3.1 Environment

Time is discrete and the horizon is infinite. Aggregate productivity is driven by an economy-wide shock, ω_t, which follows a stationary first-order Markov process, $Q(\omega_{t+1}|\omega_t)$.

Agents. There are two types of agents in the economy: workers and firms. Both are risk-neutral, infinitely-lived, and maximize their pay-offs discounted with factor β. The labor force is represented by a continuum of working age individuals with measure one. These workers are ex-ante identical and supply one unit of labor in-elastically. There is an endogenously evolving measure of firms shaped by firm entry and exit. These firms face an idiosyncratic productivity shock evolving according to a distinct first-order Markov process denoted by $\Gamma(p_{t+1}|p_t)$.

Timing. Each period t can be divided into the six following phases:

1. Productivity shocks. Aggregate productivity, ω_t, and firm-specific productivity p_t are realized.

2. Entrepreneurial shock. With probability μ, workers become potential entrepreneurs. They draw an initial idea with productivity $p_0 \sim \Gamma_0$ and decide whether to enter.

3. Firm exit. Firms decide whether to stay on or discontinue their operations based on the realization of the productivity shocks. If they exit, all of their workers become unemployed.

4. Exogenous separations. Employees at continuing firms lose their jobs with exogenous probability δ.

5. Search. Recruitment at incumbent firms takes place. Firms post vacancies to hire. Both unemployed and employed workers search for jobs.

6. Production and payments. Unemployed workers have home production b. Firms produce with their employees after the search stage. Wages accrue to employed workers. Newly created businesses start producing with a single worker, the entrepreneur.

It is assumed that workers becoming unemployed due to firm exit or a δ-shock start searching in the next period. Similarly, potential entrepreneurs (workers hit by a μ-shock) quit their job and do not search in the current period. They become unemployed should they choose not to pursue their business opportunity.

A recursive formulation is used throughout the paper. All value functions in subsequent sections are written from the production and payments stage onward, taking expectation over the events occurring in period $t+1$, conditional on the information available at the end of period t. The measure of unemployed workers and incumbent firms, which are formally defined below, are recorded at the very start of the period, before the entrepreneurial shock occurs.

Contracts. Each firm designs and commits to an employment contract. This agreement between a firm and a worker specifies a wage payment contingent on the realization of some state variable, which is made precise once the agents' problems are formally introduced. The firm chooses this contract to maximize its long-run profits, taking other firms' contracts as given. In addition, it is assumed that firms are bound by an equal treatment constraint, which restricts them to offering the same contract to all of their employees, independently of when they are hired.[5] With full commitment, the discounted sum of future wage payments can be summarized by a contract value W_t, where t denotes the realization of the contractible state in the current period. In Section 4, I derive a closed-form expression for the optimal contract that makes the model straightforward to simulate.

Workers, on the other hand, cannot commit to a firm and are free to walk away at any point. Outside offers are their private information and are therefore not contractible. Given the equal treatment constraint, if the realization of the state entails a contract value below the value of unemployment, the firm loses its entire workforce and is forced to exit. This can equally be interpreted as the employment contract specifying firm exit after certain realizations of the state.

Search and matching technology. Search is random. The probability that a vacancy reaches a worker is denoted η_t. The probability that an unemployed worker draws an offer is denoted λ_t. Employed workers have less time to search; their probability to draw an offer is given by $s\lambda_t$, for some exogenous search intensity parameter $s < 1$. Denoting A_t the stock of vacancies and Z_t aggregate search effort (from employed and unemployed workers), flow accounting directly implies

$$\lambda_t Z_t = \eta_t A_t, \qquad \lambda_t, \eta_t \leq 1.$$

Following Burdett and Mortensen (1998), there is no bargaining: workers draw a take-

[5] Because all workers are ex-ante identical and since there is no learning on the job, this constraint can be interpreted as a non-discrimination rule.

it-or-leave-it offer from a distribution, F_t, which is endogenously determined in equilibrium. Since workers can only accept or decline these offers, their decision boils down to accepting better contracts.[6]

3.2 Incumbent firms

Production. Firms operate a constant returns to scale technology with labor as its only input. n_t denotes the measure of workers currently employed at the firm. The productivity factor is given by $\omega_t p_t$. ω_t stands for the aggregate component of productivity, which is common to all firms, while p_t represents the firm's idiosyncratic productivity. ω_t and p_t follow independent first-order Markov processes and are positive by assumption.

Hiring technology. Following Merz and Yashiv (2007) and Coles and Mortensen (2016), hiring is modeled as an adjustment cost, where the cost of hiring is spread equally amongst current firm employees. A firm of current size n_t hiring a total of H_t workers has a total recruitment bill of

$$n_t c\left(\frac{H_t}{n_t}\right) = n_t c(h_t), \qquad h_t := \frac{H_t}{n_t},$$

where c is assumed to satisfy $c' > 0$, $c'' > 0$ and $c(0) = 0$. As will become clear when writing down profits, a linear recruitment technology in the firm's employment at the time of hiring implies that the firm's problem is linear in n_t. This simplification makes the model more tractable, as the firm's policy functions do not depend on n_t.

In economic terms, this formulation of the firm's hiring cost should be seen as a screening and training cost for new hires. Similarly to the model developed in Shimer (2010), current employees are an input in the recruitment process, with additional hires decreasing the revenue from each worker all else equal. The fact that this cost is convex in hires per current

[6] Burdett and Mortensen (1998) use the term "wages" instead of "contract", since, in their stationary environment, a contract is a constant, non-renegotiable wage.

employees reflect the smaller disruption to production of recruiting, say, five new employees at a business with 100 workers than at a ten-worker one. Empirical evidence suggests that these training costs can be substantial: Gu (2019) finds that job adaptation, as assessed by employers, takes 22.5 weeks on average for US non-college workers.[7]

Once the firm has chosen its target number of hires, I assume that the actual vacancies corresponding to these hires, which I formally define once the notation for aggregates is introduced, are posted at no extra cost.

Discounted Profits. Because the firm fully commits to a contract upon entry, its profits can be written in recursive form by requiring that the firm offers at least the value of the current contract in equilibrium (Promise-Keeping constraint). Let $\overline{V}$ denote the value of this contract given the realization of the states this period. Let χ_t further denote the firm's decision to continue given the realization of the shocks at the start of period t.

A firm with current productivity p_t employing n_t workers has discounted profits

$$\begin{aligned}\Pi_t\left(p_t, n_t, \overline{V}\right) = \max_{\substack{h_{t+1}\geq 0\\ w_t\\ W_{t+1}}} \Bigg\{ & (\omega_t p_t - w_t) n_t \\ & + \beta E_t\Big[\chi_{t+1}\Big(- c(h_{t+1})(1-\mu)(1-\delta)n_t + \Pi_{t+1}(p_{t+1}, n_{t+1}, W_{t+1})\Big)\Big]\Bigg\}, \end{aligned} \tag{4}$$

where the firm's continuation decision is given by

$$\chi_{t+1} := \mathbb{1}\Big\{\big(W_{t+1} \geq U_{t+1}\big) \cap \big(\Pi_{t+1} \geq 0\big)\Big\}, \tag{5}$$

since the firm needs to offer at least U_{t+1} for its workers not to quit and I assume that it must make non-negative discounted profits. Anticipating on the results in Section 4, in equilibrium the firm's continuation decision can be expressed solely in terms of the firm's current idiosyncratic productivity, though this threshold evolves with the business cycle.

[7] The exact question in the Multi-City Study of Urban Equality asks about the time it takes a typical employee in an occupation to become fully competent. See Gu (2019) for details.

In addition, the firm's maximization problem is subject to the two following constraints. First, full commitment implies a Promise-Keeping constraint in the sequential form problem. The firm's choice of wages, w_t, and contract values in the next period, W_{t+1}, has to give workers a value of at least $\overline{V}$ in expectation. Second, the size of its workforce, conditional on the firm surviving, is defined as

$$n_{t+1} = \Bigg[1 - \underbrace{q_{t+1}(W_{t+1})}_{\text{quit rate}} + \underbrace{h_{t+1}}_{\text{hiring rate}}\Bigg] \underbrace{(1-\mu)(1-\delta)n_t}_{\text{remaining workers at search stage}}. \tag{6}$$

The measure of workers employed at the search stage is $(1-\mu)(1-\delta)n_t$, those not leaving to become potential entrepreneurs (rate μ) or exogenously becoming unemployed (rate δ). $q_{t+1}(W_{t+1})$ denotes the rate at which workers still employed at the search stage leave the firm to take on better jobs, conditional on the firm offering value W_{t+1}. The quit rate is given by $q_{t+1}(W_{t+1}) := s\lambda_{t+1}\overline{F}_{t+1}(W_{t+1})$, the rate at which workers employed at the firm find better jobs in the current period.[8] Equation (6) makes the firm's trade-off in controlling the growth of n_t explicit. It can either offer better contracts, thus limiting poaching, or intensify its hiring effort through h at a higher recruitment cost.

Linearity of Discounted Profits. It can be guessed and verified that discounted profits are linear in n_t. Define profit per worker as $n_t\pi_t(p_t, \overline{V}) := \Pi_t(p_t, n_t, \overline{V})$. By substituting this guess in the right-hand side of (4) and using the law of motion for employment, it can be shown that

$$\begin{aligned}\pi_t\left(p_t, \overline{V}\right) = \max_{\substack{h_{t+1}\geq 0\\ w_t\\ W_{t+1}}} \Bigg\{ &(\omega_t p_t - w_t)\\ &+ \beta E_t\Big[(1-\mu)(1-\delta)\chi_{t+1}\Big(-c(h_{t+1}) + (1-q(W_{t+1}) + h_{t+1})\pi_{t+1}(p_{t+1}, W_{t+1})\Big)\Big]\Bigg\},\end{aligned} \tag{7}$$

[8] I define $\bar{F}_t := 1 - F_t$. Note that conditional on $q_{t+1}(W_{t+1})$ and h_{t+1} Equation (6) holds exactly by a Law of Large Number argument since n_t is the measure of workers employed at the firm.

still subject to the Promise-Keeping constraint. See Appendix A.1.

It follows directly from Equation (7) that the firm's optimal policies do not depend on its current size n_t. In particular, there is no partial layoff in the model, since the continuation decision χ_t is the same at all n_t. Jobs are only terminated in the following four cases: (i) exogenous entrepreneurial shocks at rate μ, (ii) exogenous separations at rate δ, (iii) voluntary quits for better jobs at rate $s\lambda_t\overline{F}_t(W_t)$, (iv) firm exit.

To sum up, firms are defined in the model by a recruitment technology – the cost function c – and a "contract policy" – the state-contingent contract W_{t+1} it offers to all its workers. While firm size does not enter directly the firm's policy functions, it is still well-defined in the model. This is because these policies pin down, conditional on survival, the growth rate of employment. Even if two firms with the same idiosyncratic productivity in a given period grow at the same rate, the accumulation of firm-specific shocks generates a firm-size distribution in the cross-section. The model actually replicates the Pareto tail of the empirical firm size distribution very well. I return to this point when calibrating the model in Section 5.

3.3 Firm Entry

Firm entry is governed by the decision of workers to become entrepreneurs. I assume that unemployed and employed workers draw a business idea with probability μ from an exogenous distribution Γ_0 at the start of each period t. This distribution gives the initial (firm-specific) productivity of entering businesses. I further make the assumption that employed workers cannot go back to their previous job when hit by such an "entrepreneurial shock". They must either enter the market with their new idea or become unemployed.

The decision of potential entrepreneurs to start a new business then weighs the value of starting up a firm against the value of unemployment. Entering entrepreneurs are assumed to get the full surplus $S_t(p_t) := \pi_t(p_t, \overline{V}) + \overline{V}$ of the match.[9] They then decide to enter

[9] Note that given the value of a firm is linear in n_t and given the equal treatment constraint, the surplus of the firm and all of its workers is simply $\Pi_t(p_t, n_t, \overline{V}) + n_t\overline{V} = n_t S_t(p_t)$.

if $U_t \leq S_t(p_0)$ for some initial draw p_0 from Γ_0. If they choose not to take this business opportunity, they fall back into unemployment until next period (they do not search in the current period). If they choose to enter, it is assumed that entrepreneurs have their business purchased by some outside investors (not modeled), and become the first workers at these firms.

Similarly to Gavazza et al. (2018), firms need to have positive employment to operate the recruiting technology. There is no meaningful notion of a firm with zero worker in this framework, and entrepreneurs therefore become the first workers at newborn firms. With a continuum of workers, the interpretation of this entry process is that a measure μ of workers, both employed and unemployed, becomes potential entrepreneurs in each period. They then create firms at which they become the first workers, and these firms have employment n_0. I normalize $n_0 = 1$, so that the measure of entering firms is equal to that of starting entrepreneurs.[10] These firms then move on to the production stage, and become incumbent firms from the next period onward.

3.4 Value of Employment and Unemployment

First, let Q_t denote the value of a potential entrepreneur, a worker hit by a μ-shock,

$$Q_t := \int \max\big(S_t(p), U_t\big) d\Gamma_0(p).$$

An unemployed worker has home production b and receives job offers with probability λ_{t+1}, conditional on not being hit by an entrepreneurial shock, μ. The value of being unemployed is then

$$U_t = b + \beta E_t \Big\{ \mu Q_{t+1} + (1-\mu)\Big[(1-\lambda_{t+1})U_{t+1} + \lambda_{t+1} \int \max\big(W', U_{t+1}\big) dF_{t+1}(W')\Big]\Big\}. \quad (8)$$

[10]In the British firm data, more than three quarters of entering firms report employment equals to one, where employment is defined as "employees and working proprietors."

Similarly to unemployed workers, employees are hit with probability μ by an "entrepreneurial shock", in which case they leave their present job to explore this idea. Otherwise, employed workers can search on the job with exogenous relative intensity $s < 1$. They separate with exogenous probability δ. Employed workers earn wages w_t in the current period, and are promised a state-contingent value W_{t+1} in the next period. Recall that due to the commitment structure, the firm exits and all of its workers become unemployed after some realizations, when it cannot offer its workers more than their reservation value, U_{t+1}, summarized by the indicator $\chi_{t+1}(W_{t+1})$.

Taken together, these shocks give rise to the following value function for the employed worker

$$W_t = w_t + \beta E_t \Bigg\{ \mu Q_{t+1} + (1-\mu)\Bigg[\Big((1-\chi_{t+1}) + \delta\chi_{t+1} \Big) U_{t+1} \\ + \chi_{t+1}(1-\delta)\Big(\big(1 - q_{t+1}(W_{t+1})\big) W_{t+1} + s\lambda_{t+1} \int \max\left(W', U_{t+1}\right) dF_{t+1}(W') \Big) \Bigg] \Bigg\}. \quad (9)$$

3.5 Joint Firm-Worker Surplus

The firm and worker problems can be summarized in a single expression combining (7) and (9). I show in Appendix A.2 that the following expression for $S_t := \pi_t(p_t, \overline{V}) + \overline{V}$ can be obtained after rearranging these two equations

$$S_t(p) = p_t\omega_t + \beta E_t \Bigg\{ \mu Q_{t+1} \\ + (1-\mu)\Bigg[(1 - \chi_{t+1}(p_{t+1}))U_{t+1} + \chi_{t+1}(p_{t+1})\Big(\delta U_{t+1} + (1-\delta)\psi_{t+1}(p_{t+1}) \Big) \Bigg] \Bigg\}. \quad (10)$$

In this last expression, $\psi_{t+1}(p_{t+1})$ denotes the joint value of a firm-worker pair, conditional on the firm not exiting, which writes

$$\psi_{t+1}(p_{t+1}) := \max_{\substack{h_{t+1}\geq 0\\ W_{t+1}}} \Bigg\{ -c(h_{t+1}) + (1-q_{t+1}(W_{t+1}))S_{t+1}(p_{t+1}) \\ + h_{t+1}(S_{t+1}(p_{t+1}) - W_{t+1}) + (1-\delta)s\lambda_{t+1}\int_{W_{t+1}}^{\infty} W' dF_{t+1}(W')\Bigg\}. \quad (11)$$

This simplification directly follows from the assumptions that the firm fully commits to its workers and that utility is transferable, since both firms and workers are risk-neutral. Conditional on survival, the optimal contract and hiring rate maximize Equation (11), and, importantly, the resulting contract fully internalizes the Promise-Keeping constraint.

3.6 Aggregation

Search Effort, Vacancies, and Offer Distribution. Let $\nu_t(p,n)$ denote the cumulative measure of firms with productivity less than p and workforce less than n at the start of period t, before workers are hit by "entrepreneurial" shocks and firm exit takes place. Aggregate search effort is the measure of searching workers, both unemployed and employed,

$$Z_t := (1-\mu)\left[u_t + s(1-\delta)\int \chi_t(p) n d\nu_t\right], \quad (12)$$

where the unemployment rate is $u_t := 1 - \int n d\nu_t$. This expression excludes potential entrepreneurs and displaced workers, who do not search in period t by assumption.

Let $a_t(p,n)$ denote the vacancies posted by a continuing firm with productivity p and workforce n. Total vacancy posting aggregates the vacancies of all active firms in the economy

$$A_t := \int \chi_t(p) a_t(p,n) d\nu_t. \quad (13)$$

Finally, the cumulative density of offered contracts is the sum of vacancies offering a contract less than some contract value W over the total posted vacancies (since search is random)

$$F_t(W) := A_t^{-1} \int \mathbb{1}\left\{W_t(p) \leq W\right\} \chi_t(p) a_t(p,n) d\nu_t. \tag{14}$$

Firm Vacancy Posting. To close the model, we need to specify vacancy posting by firms, $a_t(p,n)$. Since there is no cost of posting vacancies by assumption, the firm simply posts as many as required by its target hiring rate, $h_t(p)$. $a_t(p,n)$ is then implicitly defined by the accounting equation

$$h_t(p) \underbrace{(1-\mu)(1-\delta)n}_{\text{remaining workers at search stage}} = \underbrace{a_t(p,n)}_{\text{vacancies}} \underbrace{\eta_t}_{\text{contact rate}} \underbrace{Y_t(W_t(p)),}_{\text{acceptance rate}} \tag{15}$$

where η_t is the probability that this vacancy reaches a worker and $Y_t(W_t(p))$ is the chance it is accepted. This probability is determined by whether the worker reached by the vacancy is currently employed at a firm offering less than $W_t(p)$ in the current period.[11]

4 Rank-Monotonic Equilibrium

This section formalizes the definition of equilibrium used in the remainder of the paper. I provide conditions on the cost of hiring function such that the optimal contract is increasing in the current realization of idiosyncratic productivity after all histories. I label these equilibria as "Rank-Monotonic" in the rest of the paper. This characterization is similar in spirit to the "Rank-Preserving" property defined in Moscarini and Postel-Vinay (2013) in the sense that the optimal contract is increasing in firm-specific productivity in both cases. However, while in their framework with constant productivity this property implies that more productive firms are always larger along the equilibrium path – it preserves the rank of firms in the firm-size distribution – in Moscarini and Postel-Vinay (2013), idiosyncratic productivity shocks

[11] I provide a full expression for $Y_t(W_t(p))$ in Appendix A.3.

break the direct link between a firm's rank in the productivity distribution and firm size in my framework. Though more productive firms are still growing faster and therefore more likely to be large in equilibrium – contracts are monotonic in a firm's productivity – my model also allows for new, fast-growing entering start-ups. These firms show up in the model as firms entering near the top of the productivity distribution, which will grow fast while being initially small.

This property drastically simplifies the numerical solution of the model since, (i) there is no need to compute the full distribution of offered contracts, a daunting fixed-point problem as the optimal contract itself depends on this distribution, (ii) the optimal contract has a closed-form solution.

4.1 Recursive Equilibrium

Given the Markov structure of the shocks, attention can be restricted to recursive equilibria in which the state-space relevant to the firm's decision is made of the two shocks and the measure of firms in the (n, p)-space, ν, the latter being sufficient to compute all aggregates in the model. In addition, Equation (7) makes clear that the firm's current size is not part of this state-space. More formally:

Definition 1 *A Recursive Equilibrium is a triple of policy functions (V,h,χ) and a pair of value functions (S,U) that depend on the current realization of aggregate productivity, the current realization of idiosyncratic productivity, and the measure of firms at the start of the period. Given that all firms follow the policies given by (V,h,χ), these functions satisfy:*

1. *Equations* (12)-(15) *hold with* $\chi_{t+1}(p) = \chi(p, \omega, \nu)$, $h_{t+1}(p) = h(p, \omega, \nu)$, *and* $V_{t+1}(p) = V(p, \omega, \nu)$;

2. *The contract and hiring functions solve the maximization problem in* (11). *The continuation decision is given by* $\chi(p, \omega, \nu) = \mathbb{1}\{V(p, \omega, \nu) \geq U(\omega, \nu)\}$;

3. S *and* U *solve, respectively,* (10) *and* (8).

4.2 Rank-Monotonic Equilibrium

A Rank-Monotonic Equilibrium (RME) adds the following requirement to the optimum contract:

Definition 2 *A Rank-Monotonic Equilibrium is a Recursive Equilibrium such that the optimal contract, $V(p, \omega, \nu)$, is weakly increasing in p for all ω and ν.*

Result 1 further provides sufficient conditions on the cost of hiring function such that a Recursive Equilibrium is in fact Rank-Monotonic.

Result 1 *Assume that the hiring cost function is twice differentiable, increasing and convex. Assume the Markov process for idiosyncratic productivity satisfies first-order stochastic dominance ($\Gamma(.|p') \leq \Gamma(.|p)$ for $p' > p$ with strict inequality for some productivity level). Then:*

1. *The firm-worker surplus defined by Equation* (10) *is increasing in p;*

2. *Any equilibrium is Rank-Monotonic if*

$$\frac{c''(h)h}{c'(h)} \geq 1, \quad \forall h \geq 0.$$

The proof is in Appendix A.4. Similarly to the result in Moscarini and Postel-Vinay (2013), Result 1 is not an existence statement, but a characterization of the properties of the optimal contract conditional on the existence of such an equilibrium.

The condition on the cost function in Result 1 is an additional convexity requirement. Firms use the retention margin – through offering better contracts – only in the extent the hiring technology is sufficiently costly. With identical workers and no learning on the job, the model could potentially generate a large amount of churning at the top of the productivity distribution if employers have little incentives to promise their worker higher values to retain them. Given the conditions in Result 1, hiring costs become so high for larger h that firms find it optimal to use both the retention and hiring margins to control their optimal growth rate.

The rest of the paper centers on Rank-Monotonic equilibria. When taking the model to the data, I restrict the parameter space to ensure that the convexity requirement on the cost of hiring function in Result 1 is satisfied.

4.3 Additional Characterization of RMEs

Because the optimal contract is increasing in p after every history in a Rank-Monotonic Equilibrium, several aggregates can be recast as functions of p, which allows to further characterize the optimal contract. I start by defining the measure of workers employed at firms less productive than p at the start of period t

$$L_t(p) := \int_{\tilde{p} \leq p} n d\nu_t(\tilde{p}, n).$$

Note that this last measure fully summarizes acceptance/quit decisions at each level of productivity in a Rank-Monotonic Equilibrium since the optimal contract is increasing in p. Firms will poach workers from firms with productivity below them and lose workers to firms with productivity above them.

Optimal policies. First, since both the firm-worker surplus and the optimal contract are increasing in p in a Rank-Monotonic Equilibrium, the entry and exit thresholds coincide. The firm's continuation policy can be written $\chi(p, \omega, L) = \mathbb{1}\{S(p, \omega, L) \geq U(\omega, L)\}$. I denote $p_E(\omega, L)$ the corresponding entry and exit productivity threshold, which is implicitly defined by $S(p_E, \omega, L) = U(\omega, L)$.

Second, Appendix A.5 shows that the optimal contract takes the following form

$$V(p,\omega,L)=\frac{uU(\omega,L)+s(1-\delta)\displaystyle\int_{p_E}^{p}S(\tilde{p},\omega,L)dL(\tilde{p})}{u+s(1-\delta)\Big(L(p)-L(p_E)\Big)}. \tag{16}$$

The optimal contract is therefore a weighted average between the value of unemployment and the firm-worker surplus, where the weights are given by, respectively, the measure of workers in unemployment and the measure of workers searching this period at firms with productivity less than p. This expression is reminiscent of the Nash-Bargaining solution used in classic search models, which breaks down the firm-worker surplus between each party with a constant exogenous weight (e.g., Mortensen and Pissarides, 1994). The difference in my setting is that the weights are fully endogenous and evolve with the distribution of workers over the business cycle.

Third, the optimal hiring rate follows directly from inverting the derivative of the cost function in the firm's corresponding first-order condition from Equation (11)

$$c'(h(p,\omega,L))=S(p,\omega,L)-V(p,\omega,L).$$

Distribution of offered contracts. In a RME, the acceptance rate for a firm with current productivity p can be expressed as a function of the measure of workers employed at firms with current productivity below p. The distribution of offered contracts can then be simplified as

$$\lambda_t F_t\left(V(p)\right)=\int_{p_E}^{p}\frac{h_t(\tilde{p})}{u_t+s(1-\delta)\Big(L_t(\tilde{p})-L_t(p_E)\Big)}dL_t(\tilde{p}). \tag{17}$$

The derivations can be found in Appendix A.5.

Employment Law of Motion. Taken together, these policies imply the following law of motion for the measure of employed worker

$$L_t^P(p) = \mu \int_{p_E}^{p} \chi_t(\tilde{p}) d\Gamma_0(\tilde{p}) + (1-\mu)\chi_t(p) \Big[L_t(p)\rho_t(V_t(p)) + u_t \lambda_t F_t(V_t(p)) \Big], \tag{18}$$

where L_t^P denotes the measure of workers at firms with productivity less than p at the end of period t (at the production stage). The first term corresponds to entering entrepreneurs with initial draws less than p. The two terms in the square brackets give, first, the fraction of workers retained at firms less than p and the inflow from unemployment. The end of period and beginning of next period measures are directly linked by

$$\frac{dL_{t+1}(p)}{dp} = \int_{\underline{p}}^{\overline{p}} \frac{dL_t^P(\tilde{p})}{d\tilde{p}} d\Gamma(p|\tilde{p}),$$

which corresponds to the "re-shuffling" of workers across productivity levels due to the firm-specific shocks.

To sum up, knowing the value functions S and U for all values of the aggregate shock and the measure of employment across firm productivity is enough to simulate the model in the presence of aggregate shocks.[12] The firm's optimal policies admit closed-form solutions conditional on these value functions, and these policies in turn determine the law of motion for workers across firm productivity.

5 Calibration

This section presents the calibration and simulation procedure. Though the size independence and Rank-Monotonic equilibrium results simplify the firm's problem, solving for the firm's policies still requires to keep track of the measure of workers across firm idiosyncratic

[12] It is also possible to define the net surplus as $\phi(p, \omega, L) := S(p, \omega, L) - U(\omega, L)$ and express the firm's policies as a function of this single value function, which I do in practice when simulating the model. To economize on notation, all the corresponding expressions are relegated to Appendix A.6.

productivity levels, L_t. I then proceed in two steps to calibrate the model.

I start by solving the model without aggregate shocks, and target some key labor market and firm dynamics moments from British data to calibrate the main parameters. In doing so, I focus on a Stationary Rank-Monotonic Equilibrium. Formally:

Definition 3 *A Stationary Rank-Monotonic Equilibrium is a triple of policy functions* (V, h, χ)*, a pair of value functions* (S, U)*, and a measure of workers across firm productivity* L*, that depend on the current realization of the firm's productivity* p*. These functions satisfy the following requirements:*

1. *The conditions for a Rank-Monotonic Equilibrium in Definition 2 are satisfied;*
2. *The law of motion for the measure of worker induced by the firm's optimal policies* (18) *is constant and equal to* L*.*

I return to the full model with aggregate shocks in a second step, and describe how the measure of workers is approximated out of steady-state in Section 6.

5.1 Parametrization

A period t is set to a month. I specify the Markov processes for idiosyncratic productivity shocks as $\ln p_{t+1} = \rho_p \ln p_t + \sigma_p \epsilon^p_{t+1}$ with $\epsilon^p_{t+1} \sim \mathcal{N}(0, 1)$. Such a process satisfies first-order stochastic dominance conditional on past realizations, which is required for the equilibrium to be Rank-Monotonic (Result 1). The productivity of initial ideas, Γ_0 is assumed to follow a log-normal distribution with mean μ_0 and standard deviation σ_0. The functional form for the cost of hiring function is guided by the conditions derived in Result 1. I calibrate the parameters in the following cost function

$$c(h) = \frac{(c_1 h)^{c_2}}{c_2}, \tag{19}$$

which satisfies the convexity requirements in Result 1 provided $c_2 \geq 2$. I enforce this condition when searching over the parameter space. Taken together, these functional form assumptions give the following vector of eleven parameters to calibrate

$$\left\{\beta, \delta, c_1, c_2, s, \mu, b, \rho_p, \sigma_p, \mu_0, \sigma_0\right\}.$$

5.2 Calibration strategy

The discount factor, β, is set in line with a 5% annual discount rate. This leaves ten parameters to calibrate, which I pin down by targeting an equal number of moments from the data. My choice of moment targets reflects both the search and firm dynamics components of the model. To discipline worker transitions in and out of unemployment and between employers, I target the unemployment to employment (UE), employment to unemploment (EU), and job-to-job (EE) average monthly transition rates in the UK over the pre-crisis period (2000-2007). These series are derived from the British Household Panel Survey (BHPS) following the methodology described in Postel-Vinay and Sepahsalari (2019).

To discipline the life cycle of firms, I target the firm exit rate, as well as the auto-correlation and inter-quartile range of labor productivity. I use the measure of labor productivity defined in Equation (2) (log sales over employment), deviated from year-industry averages. These moments are computed directly from the Business Structure Database (BSD), and are therefore yearly measures. In addition, I also include moments that relate specifically to the dynamics of young firms. Firms are labeled as "young" if they are less than ten years old, since this cut-off implies an equal share of young and old firms on average. These moments are the share of young firms, the share of workers employed by young firms, and the exit rate and inter-quartile range of labor productivity at young businesses. They are also derived from the BSD.

To compute the moments implied by the model, I solve for a Stationary Rank-Monotonic Equilibrium, given a vector of candidate parameters. This yields a distribution of firms

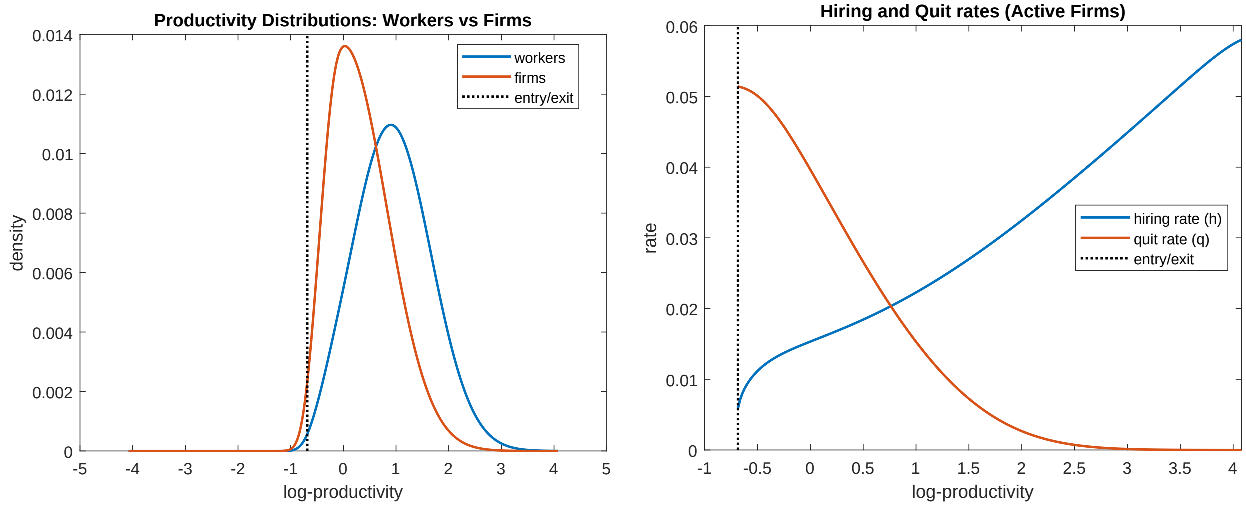

Figure 4: Properties of Stationary Rank-Monotonic Equilibrium at estimated parameters. Left panel: beginning of period productivity distributions of firms and workers. Right panel: monthly employment growth rate implied by the firm's policies.

and workers across productivity levels, as well as an entry/exit productivity threshold and a monthly employment growth rate for surviving firms. Figure 4 depicts the obtained distributions and employment growth rate at the estimated parameters. The monthly transition rates can then be computed directly based on this equilibrium. For instance, the monthly probability to find a job when unemployed implied by the model is given by

$$\underbrace{\mu \int_{\tilde{p} \geq p_E} d\Gamma_0(\tilde{p})}_{\text{successful entrepreneurs}} + \underbrace{(1-\mu)\lambda.}_{\text{successful job search}}$$

However, because the moments relating to firm dynamics are derived from yearly data, their model counterpart are obtained by simulating a panel of firms. I simulate a cohort of 150,000 entrants (roughly the size of a typical cohort in the Business Structure Database) for twenty years and aggregate the output from that simulation exactly like the data. Note that monthly turnover at firm i and in month t is defined as $p_{i,t}n_{i,t}$ and summed over a year to get a model equivalent to the turnover concept in the BSD and compute the labor productivity measure defined in Equation (2). Since this last productivity measure is in logs, the actual units of sales are irrelevant to my calibration.

The model fit to the targeted moments is shown in Table 2. Overall, the model replicates these statistics well, with the exception of the exit rate at young firms and the persistence of labor productivity, which are both slightly lower in the model than their empirical counterpart. The model can still account for about half of the difference in firm exit between young and old businesses.

I show how each parameter is related to each moment in Figure 5.[13] The figure depicts the absolute elasticity of each moment to each parameter around its estimated value. Though each parameter drives more than one moment, the following broad groups can be derived from the figure. The main parameters determining the transition rates are the job destruction rate (δ), the cost of hiring function parameters (c_1, c_2), and relative search effort s. The exit rate moments are primarily driven by the persistence of idiosyncratic productivity (ρ_p) and the mean relative productivity of entrants (μ_0). The dispersion and correlation of labor productivity are mainly determined through the flow-value of unemployment (b) and the standard deviation of the shocks (σ_p, σ_0). Finally, the share of young firms and employment at young firms are primarily responding to the rate of arrival of business ideas (μ), with the relative search intensity (s) and persistence of idiosyncratic productivity (ρ_p) also playing a role.

The estimated parameters are listed in Table 3. The estimated job destruction rate is low, since the bulk of EU transitions come from firm exit in the model.[14] There is no clear benchmark in the literature for the hiring cost function parameters because this functional form has seldom been used. I find that the implied average hiring cost as a fraction of monthly sales is 5.3%. Among the studies using a related specification, Merz and Yashiv (2007) estimates the exponent to be approximately cubic, but in a pure adjustment cost model without search frictions, while Moscarini and Postel-Vinay (2016) use a highly convex function (exponent = 50) in their baseline calibration, but with this cost applying to the

[13] I also report slices of the objective function for each parameter in Appendix C.2.

[14] A potential strategy to further discipline this feature would be to get an estimate of the fraction of EU transitions coming from firm exit. This information is not readily available in UK data.

Moment	Model	Data
Worker transitions (monthly)		
UE	0.067	0.069
EU	0.004	0.004
EE	0.020	0.020
Firm dynamics(yearly)		
Exit Rate	0.133	0.128
$\rho(LP_{i,t}, LP_{i,t-1})$	0.632	0.798
$IQR(LP_{i,t})$	0.669	0.678
Average Employment	12.8	12.8
Firm Share Young	0.621	0.560
Exit Rate Young	0.141	0.176
$IQR(LP_{i,t})$ Young	0.645	0.616

Table 2: Targeted moments.

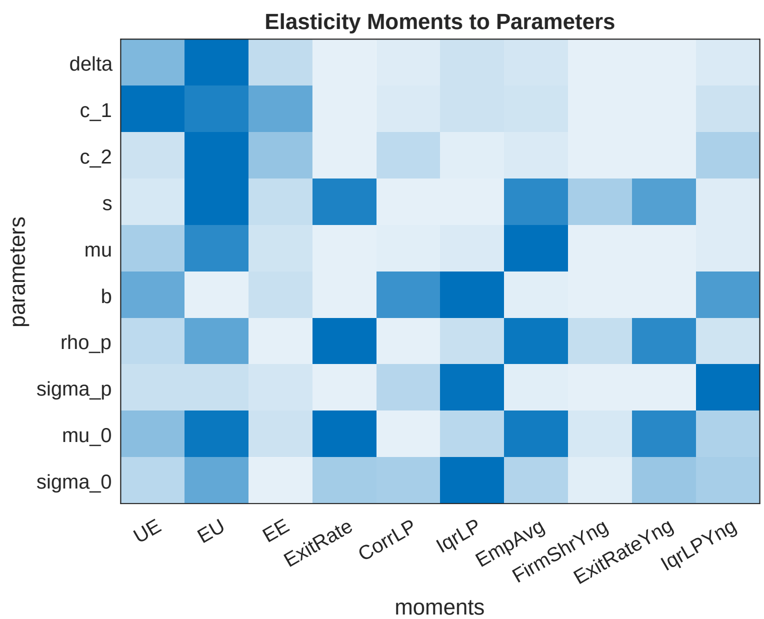

Figure 5: Elasticity of each moment to each parameter. Each cell corresponds to $\left|\frac{d \ln \text{moment}_j}{d \ln \text{parameter}_i}\right|$ for each parameter in row i and each moment in column j. A darker shade of blue indicates a larger absolute elasticity. The elasticities are computed by solving the model in a small neighborhood around the parameters and fitting a line through each parameter-moment series in logs.

Parameter	Description	Value
Pre-calibrated		
β	discount factor (≈5% annual)	0.996
Estimated		
δ	prob. job destruction (×100)	0.085
c_1	hiring cost:	45.855
c_2	$c(h) = (c_1 h)^{c_2}/c_2$	4.977
s	relative search effort	0.802
μ	prob. of start-up (×100)	0.082
b	flow value of unemployment	0.502
ρ_p	firm productivity process:	0.982
σ_p	$\ln p_{t+1} = \rho_p \ln p_t + \sigma_p \epsilon^p_{t+1}$	0.153
μ_0	initial productivity draw:	-0.200
σ_0	$\ln p_0 \sim \mathcal{N}(\mu_0, \sigma_0)$	0.050

Table 3: Parameter estimates.

number of actual hires and not the hiring rate. The relative search effort (s) of employed worker is large compared to traditional estimates obtained from US data. This reflects the fact the EE transition rate is much larger relative to the UE transition rate in British data (respectively .02 and .07 monthly in the British Household Panel Survey) than in US data (respectively .02 and .21 monthly in the Survey of Income and Program Participation). The flow-value of unemployment represents 19% of the average wage in the economy, which is about half the value used in Shimer (2005).[15] The idiosyncratic shock parameters, finally, imply a large degree of persistence of idiosyncratic productivity and a much larger dispersion of idiosyncratic productivity post-entry than pre-entry. As such post-entry shocks are a key driver of the life-cycle of the firm in the model.

5.3 Firm size distribution

Though the firm size distribution is not included in the set of targeted moments, the model still generates the large concentration of employment in the largest firms observed in the

[15]Hornstein et al. (2011) show that lower values of b in the Burdett and Mortensen (1998) model yield a mean to min wage ratio more in line with the data.

data. Figure 6 displays the normalized employment size (employment at the firm divided by average firm employment in the economy) and the associated complementary CDF (the firm's rank in terms of employment size) in the model and the data on a log-log scale. It shows that the model can replicate the log-linear relationship between firm employment and tail probability, a well-documented empirical feature of the firm size distribution. The resulting Pareto coefficient, estimated for the sample of firms larger than average size, is 1.066 in the data and 1.03 in the model.

This feature of the model can be rationalized within the framework developed by the literature on power laws in economics (e.g., Gabaix, 1999). This line of research stresses several characteristics of the underlying process driving the size of individual units – firm employment in my setting, but typically the population of cities – that lead to a Pareto tail in steady-state. First, the growth rate of individual units is modeled through an evolving, but size independent growth rate (Gibrat's Law). Second, these individual units must be exposed to a birth-death process (Reed, 2001).

Without going into the technical details underpinning these results, I note that the evolution of firm size in my model is consistent with these requirements.[16] First, as shown in Section 3, the constant returns to scale assumption implies that the firm's policies are independent of employment. Conditional on survival, the growth rate at a firm with current productivity realization p is given by $(1-\mu)(1-\delta)\big(1-q(V(p))+h(p)\big)$, irrespective of its current employment. Second, the entry-exit threshold naturally generates firm birth and death.

6 Business Cycle

In a Stationary Equilibrium, the distribution of workers across firm productivity is stationary and consistent with the firm's optimal policies by definition. But in the presence of aggregate

[16] Gouin-Bonenfant (2019) also gets a similarly good fit to the firm size distribution in a search model with similar properties.

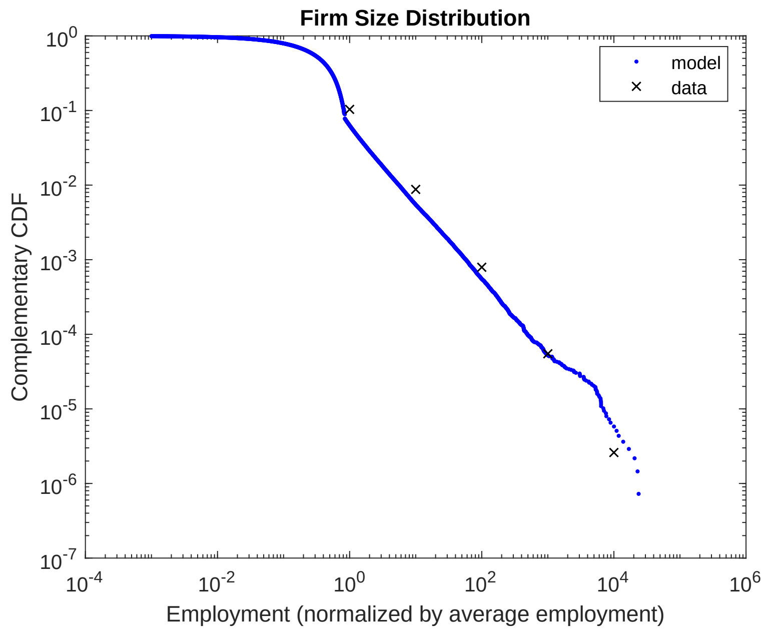

Figure 6: Firm size distribution at the calibrated parameters. Normalized employment is defined as employment at the firm divided by average firm employment in the economy. The complementary CDF at firm employment n is given by $\Pr(\text{employment}_{i,t} > n)$. The data series comes from the BSD and is computed separately for each year and averaged across years. The model series is obtained by simulating a cohort of entrants for one hundred years.

shocks, this distribution evolves over time and enters the firm's state space (see Definition 2). This extra state variable comes with a technical hurdle since the distribution of workers across firm productivity is an infinitely dimensional object.

In this section, I start by describing the approximation used to solve the model out of steady-state. I then proceed with a series of exercises highlighting the interplay of firm dynamics and search frictions in accounting for labor productivity following the Great Recession in the UK.

6.1 Solving the Model with Aggregate Shocks

I now reintroduce aggregate shocks in the model. In the spirit of Krusell and Smith (1998), the measure of workers out of steady-state is approximated by a set of its moments. This measure is summarized by the unemployment rate, $u_t = 1 - \int dL_t(p)$, and a vector of moments

$\boldsymbol{m}_t$ from the normalized measure of workers $L_t / \int dL_t(p)$.[17]

In addition, simulating the full model with aggregate shocks requires to solve for the firm's policy functions for all values of the aggregate shock and the distribution of workers, L_t. Given the approximation of L_t, the state-space relevant to the firm now reduces to ω_t, u_t, and $\boldsymbol{m}_t$. I then approximate the firm-worker surplus and the unemployed worker's value function out of steady-state with a polynomial.[18] For instance, the value function for workers in unemployment is approximated as

$$\ln U(\omega_t, L_t) - \ln \bar{U} \approx \tilde{U}(\omega_t, \tilde{u}_t, \tilde{\boldsymbol{m}}_t; \theta_U)$$

where $\tilde{x}_t$ denotes a variable in log-deviation from steady-state and θ_U is a vector of coefficients to be solved for. The firm-worker surplus is similarly approximated, using a separate polynomial at each idiosyncratic productivity node. The solution algorithm proceeds by repeatedly simulating the model until the coefficients converge. Additional details regarding the implementation of this algorithm can be found in Appendix C.3.

An alternative approach to simulate heterogeneous agents models with aggregate shocks is to use the perturbation method proposed by Reiter (2009). Such linearization techniques have been successfully applied to firm dynamics models (Sedláček and Sterk, 2017; Winberry, 2016). However, my simulations suggest that this first-order approximation is highly inaccurate in the context of my model due to the discontinuity implied by the firm's entry and exit threshold. I therefore choose the simulation-based approach outlined here and report accuracy tests for my proposed algorithm in Appendix C.4.

Finally, the Markov process for aggregate productivity shocks is assumed to follow $\ln \omega_{t+1} = \rho_\omega \ln \omega_t + \sigma_\omega \epsilon^\omega_{t+1}$ with $\epsilon^\omega_{t+1} \sim \mathcal{N}(0,1)$. The parameters in this process $(\rho_\omega, \sigma_\omega)$ are chosen to replicate the model-simulated persistence and volatility of unemployment in the UK between 1971 and 2018. They are shown in Table 4.

[17] Recall that there is a measure one of workers, so $u_t + \int dL_t(p) = 1$ by definition.

[18] I approximate the value functions and not the firm's policies directly since the latter are not smooth functions of the aggregate states due to the entry/exit threshold.

ω parameters		Persistence		Volatility	
ρ_ω	0.965	Data	0.937	Data	0.081
σ_ω	0.110	Model	0.920	Model	0.084

Table 4: Parameters aggregate shock (ω_t). Persistence and volatility are, respectively, the first autocorrelation and standard deviation of HP-filtered (log) unemployment in the UK between 1971Q1 and 2018Q4. The model series are obtained from simulating the model and aggregating and filtering its output similarly to the data.

6.2 The Great Recession in the Model

To understand the reallocation effects of a large recession in the model, I input a sequence of aggregate shocks that triggers a sharp rise in unemployment, akin to the UK experience during the Great Recession. I show that the model can generate a reduction in the allocation of workers to firms that is in line with the patterns documented in the data. I then leverage the model to further account for changes in labor productivity following the recession in terms of firm selection and worker reallocation.

Aggregate responses. Figure 7 displays key model generated aggregates along with the respective data series. The top left panel shows the Great Recession counterfactual I run in the model: I input a sequence of aggregate shocks designed to replicate the sharp increase in unemployment observed at the onset of the recession. Aggregate productivity is then left to revert back to its steady-state level according to the process given in Table 4.

The remaining plots benchmark the implied responses in the number of active firms, the labor market transition rates, and vacancies against the corresponding data series.[19] All data series are shown in deviation from their pre-recession linear trend. Overall, the model does a decent job at replicating the overall pattern of these aggregates, keeping in mind that these series are not targeted and that the model is calibrated on the pre-recession period based on its stationary solution.

[19]While it is not necessary to specify a matching function to solve the model since it can be solved using the identity $\lambda_t Z_t = \eta_t A_t$, a functional from assumption is required to back out vacancies. I use the standard Cobb-Douglas form $\kappa A_t^\alpha Z_t^{1-\alpha}$ where I normalize $\kappa = 1$ and set the elasticity to vacancies to .5.

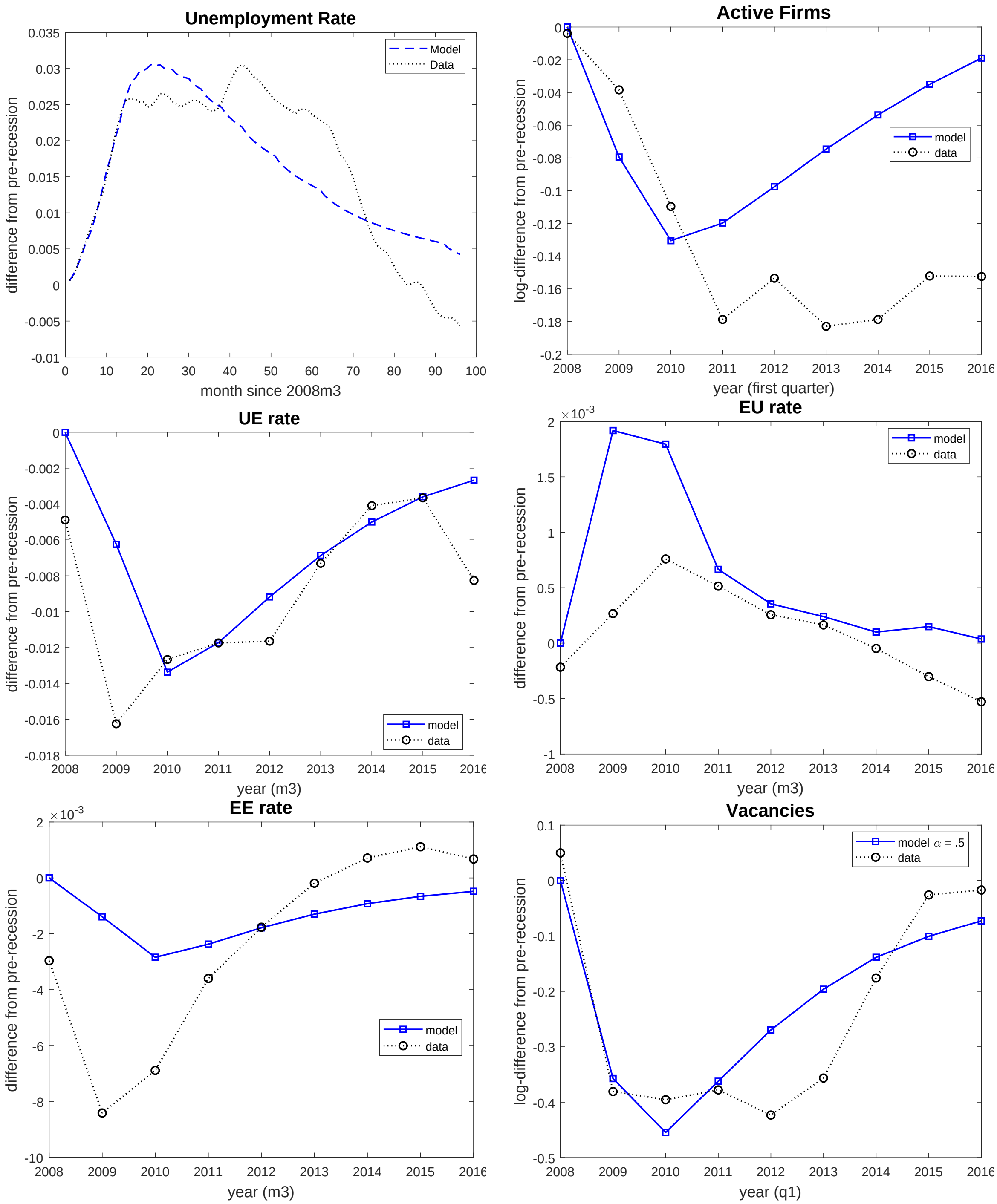

Figure 7: Aggregate model responses to a sequence of productivity shocks triggering the increase in unemployment depicted in the top left panel. See Appendix B.4 for the source of additional series.

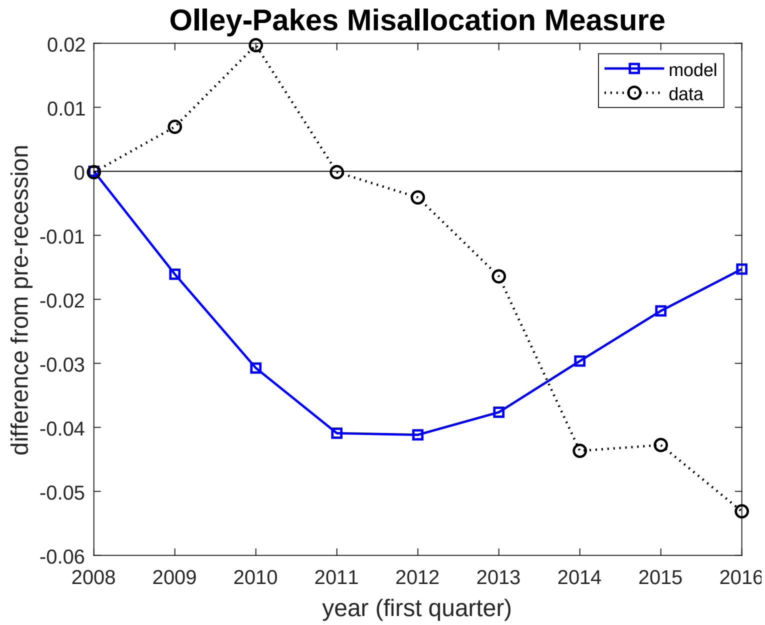

Figure 8: Olley-Pakes misallocation measure during the simulated recession. The data is computed in deviation from their pre-recession linear trend. The model series is computed simulating a cohort of firms over the course of the recessionary episode and aggregating its output similarly to the data.

As an additional validation, I study the reallocation of workers implied by the model in the simulated recession. The measure of worker reallocation used is similar to the empirical part of the paper and given by

$$\sum_i \left(ES_{i,t} - \overline{ES}_t\right)\left(LP_{i,t} - \overline{LP}_t\right),$$

an expression that increases with a higher share of workers employed at the most productive firms. Figure 8 benchmarks the model response against the data in deviation from their pre-recession linear trend. It shows that the model generates a drop in this measure that is, overall, similar in magnitude to that observed in the data. Though it recovers more quickly than the data, it still accounts for more than fifty percent of the overall response by 2015, seven years after the start of the recession.

The Olley-Pakes decomposition through the lens of the model. Recall that the labor productivity index used in the empirical part of the paper is given by

$$LP_t = \sum_i ES_{i,t} \times LP_{i,t},$$

where $ES_{i,t}$ and $LP_{i,t}$ denote, respectively, the employment share and labor productivity at firm i in period t. In the notation of the model, this expression rewrites

$$LP_t = \int \underbrace{\ln\left(\frac{\omega_t p n}{n}\right)}_{\text{firm productivity}} \underbrace{dn\overline{\nu}_t^P(p,n)}_{\text{employment ``share''}} = \ln\omega_t + \int \ln(p) d\overline{L}_t^P(p),$$

where the superscript "P" denotes the production stage (end of period) and a bar denotes a normalized measure.[20] This last equality makes clear that aggregate labor productivity is determined by the aggregate shock and the employment-weighted distribution of firm productivity, $\overline{L}_t^P$, an object shaped by firm dynamics and search frictions in equilibrium.

Again LP_t can be further decomposed into a firm productivity component and a worker reallocation component. The equality

$$LP_t = \sum_i ES_{i,t} \times LP_{i,t} = \overline{LP}_t + \sum_i \left(ES_{i,t} - \overline{ES}_t\right)\left(LP_{i,t} - \overline{LP}_t\right)$$

can be written in the model as

$$LP_t = \underbrace{\ln\omega_t}_{\text{aggregate shock}} + \underbrace{\int \ln(p) dK_t^P(p)}_{\text{firm selection}} + \underbrace{\int \ln(p) dL_t^P(p) - \int \ln(p) dK_t^P(p)}_{OP_t := \text{measure of allocation}}. \tag{20}$$

In this expression, the first term gives the direct impact of the aggregate shock, the second term captures changes in the distribution of firms across productivity level. Finally, the last term corresponds to the allocation of worker to firms. In the model, it relates directly

[20] So $\overline{L_t}^P(p) := \int_{\tilde{p}\leq p} dL_t^P(\tilde{p}) / \int dL_t^P(p)$. Besides, because a model period is a month, this expression is monthly labor productivity.

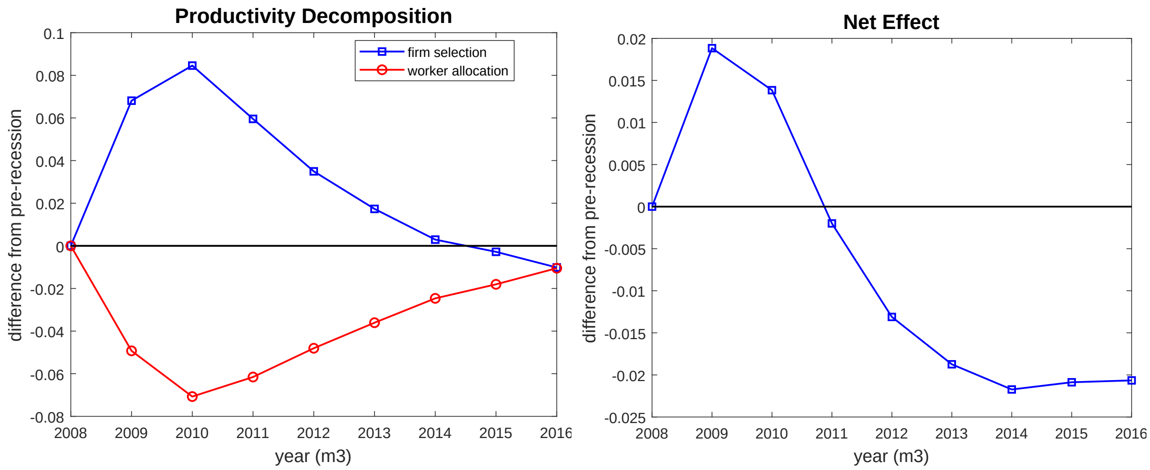

Figure 9: Labor productivity decomposition in the model.

to the difference between the distribution of workers across firm productivity (L_t^P) and the distribution of firms across productivity (K_t^P), two objects jointly determined in equilibrium in the model.

Through the lens of the model, aggregate labor productivity is then made up of two endogenous terms: firm selection and worker reallocation (the last two terms in Equation 20). I plot the evolution of these two components over the course of the simulated recession in Figure 9. It shows that they represent opposite forces shaping labor productivity. However, while they are initially of the same magnitude, the worker reallocation term exhibits more persistence. It is still negative eight years after the start of the recession. On net, the worker reallocation effect dominates in the medium term, as shown on the right panel of Figure 9.

Inspecting the worker reallocation mechanism. I illustrate the main worker reallocation mechanism at the micro level in Figure 10. On top of the firm selection effect, which shifts the entry threshold upward, how well labor is allocated to firms also depends on which firms grow faster following the shock. Figure 10 shows changes in the quit rate, hiring rate, and net employment rate with respect to their pre-recession level along the firm-specific productivity dimension.

The figure shows that while the hiring rate drops at all productivity levels with respect

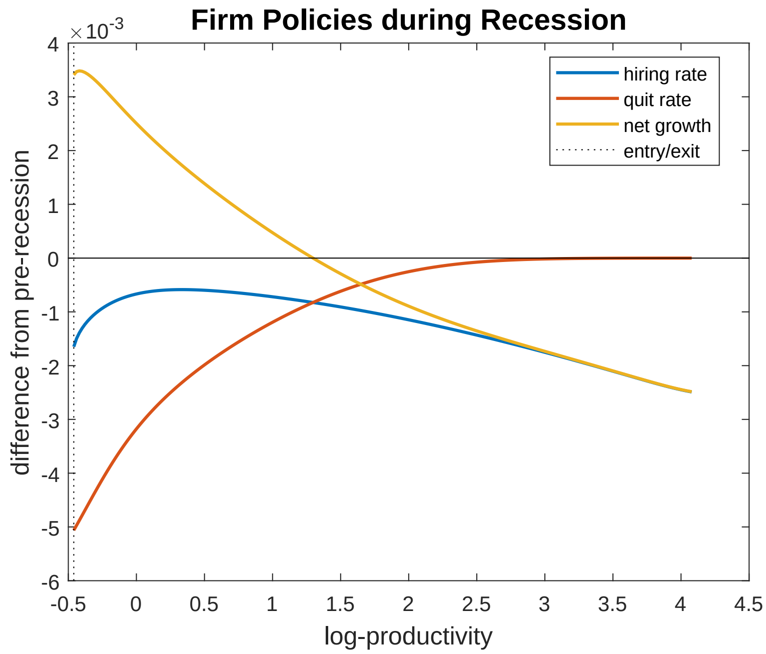

Figure 10: Firm policies during the simulated recessionary episode. Policies are averaged across simulation periods and presented in deviation from the stationary equilibrium. Idiosyncratic productivity is truncated to only include firms above the entry/exit threshold at all point in time after the shock.

to the pre-recession period, the quit rate drops even more at the bottom of the productivity distribution. This is because, in a random search environment, the probability for workers to draw an offer from a high-productivity firm is reduced as they compete with more unemployed workers. Since voluntary quits are always productivity enhancing in equilibrium, this reduction in the quit rate contributes to dampening labor productivity.

The fact that the resulting net employment growth rate increases – in relative terms, since these firms are still shrinking, but not as fast as they would in normal times – at the bottom of the productivity distribution during the shock is consistent with the firm-level data. In Table 1, I find that the relationship between labor productivity at the firm level and employment growth becomes less positive in the aftermath of the recession. While this relationship cannot be decomposed further into hires, quits, and layoffs without matched employer-employee data, the drop in the quit rate at low quality firms is consistent with the evidence described in Haltiwanger et al. (2018) for the US. These authors find that job-to-job transitions out of the bottom rung of the wage ladder – where firms are ranked based on wages and not productivity as in Figure 10 – decline by eighty-five percent during the Great Recession.

6.3 Policy Experiment: Unemployment-contingent Benefits

The trade-off between firm selection and worker reallocation during a recession can be further illustrated in the following policy experiment. In the spirit of unemployment insurance extensions in the US, I allow the value of non-employment, b, to depend on the unemployment rate.[21] Specifically, the value of non-employment is assumed to vary with unemployment benefits according to

$$\ln b_t - \ln \overline{b} = \kappa \times (\ln u_t - \ln \overline{u}),$$

where $\overline{b}$ and $\overline{u}$ denote, respectively, the value of non-employment and the unemployment rate in the stationary equilibrium and κ is the elasticity of unemployment benefits to the unemployment rate. I tentatively set $\kappa = .3$ and solve the model again using the same sequence of aggregate shocks as in the benchmark economy.

Figure 11 compares the response of the model under the unemployment-contingent benefit policy to the baseline constant b model over the course of the simulated Great Recession. With respect to labor productivity, such a policy has two opposite effects. First, it makes the selection effect more stringent. Unemployment increases by three percentage points at its peak in the baseline model and by almost four and a half points under the alternative policy. This effect is reflected in the firm selection term, which is also more positive since the entry threshold is higher.

Second, unemployment-contingent benefits magnify the worker reallocation effect result-

[21] The actual policy makes the duration, and not the level, of unemployment benefits contingent on the unemployment rate. I focus on the level of these benefits to avoid the need to introduce an extra state variable for unemployed workers off and on benefits. See Rujiwattanapong (2019) for a model fully capturing the unemployment insurance extension mechanism.

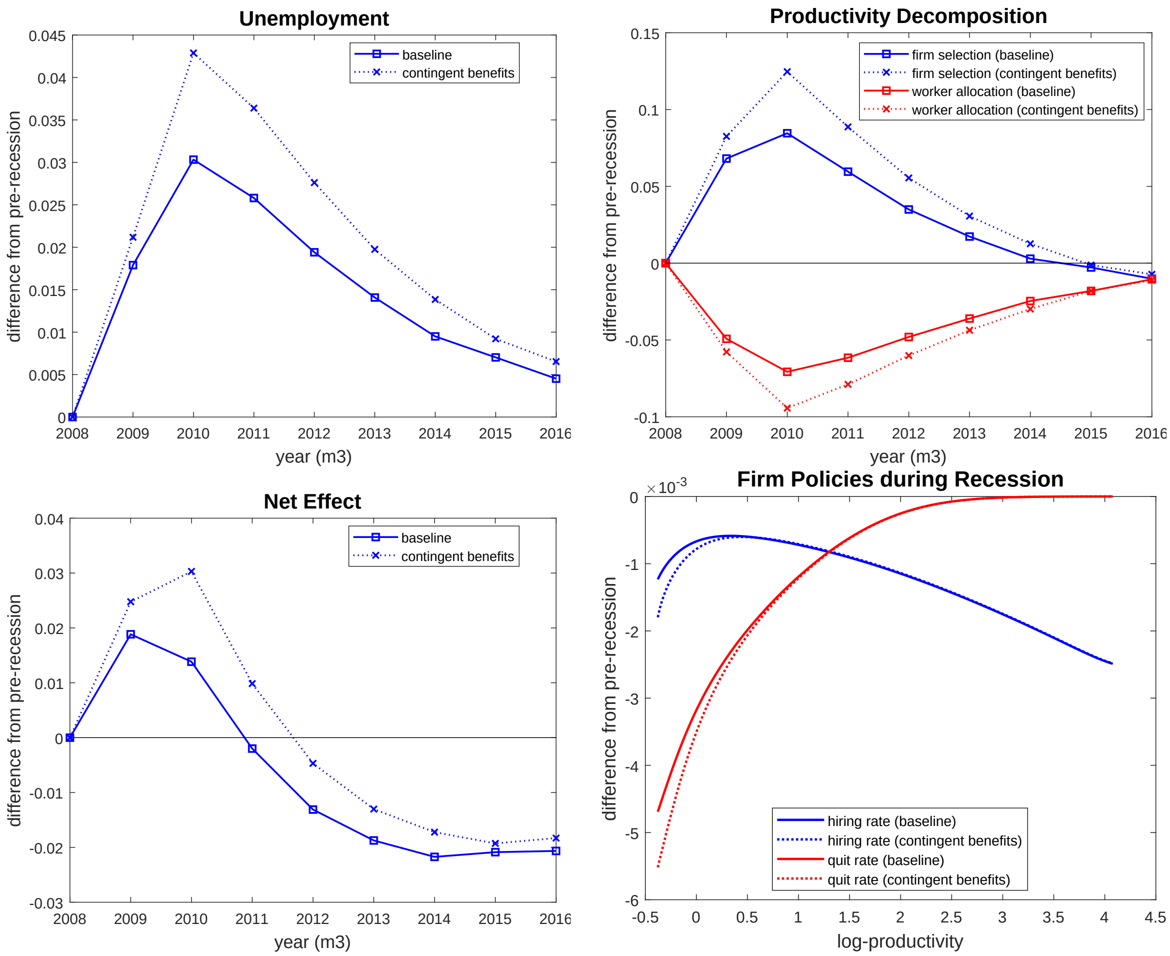

Figure 11: Baseline vs Unemployment-contingent benefits model. Model response to the simulated recession using the same sequence of aggregate shocks.

ing from search frictions. As can be seen from the firm's policies, the quit rate drops even more at the bottom of the distribution in this case: workers employed at these firms must compete with more unemployed workers to climb up the contract-productivity ladder. While the net effect of the policy is still positive in my calibration, the model does suggest that such policies can also have negative consequences on labor productivity by decreasing the pace of worker reallocation to more productive units.

7 Conclusion

I develop a model with three key features: (i) on-the-job search, (ii) firm dynamics, (iii) aggregate shocks. Firms with heterogeneous productivities compete to attract and retain workers in a frictional labor market. In equilibrium, job-to-job transitions are always productivity enhancing, as more productive firms offer better contracts. I use the model to analyze how firms' recruiting behaviors at the micro level drive the evolution of aggregate labor productivity at the macro level in the aftermath of a recession.

The central insight of the model is that search frictions dampen labor productivity following a large aggregate shock. On-the-job search causes the quit rate – the rate at which workers voluntary leave their current job to take a better one – to drop on the lower part of the productivity distribution after a recession. Search frictions then hamper the reallocation of workers from less to more productive firms.

In an experiment designed to replicate the increase in unemployment observed during the UK Great Recession, I find that this channel can account for a large portion of the drop in the allocation of workers to firms measured in British firm-level data. Through the lens of the model, this negative worker reallocation effect dominates the positive firm selection effect implied by the aggregate shocks in the medium term.

References

Barlevy, G. (2002). The sullying effect of recessions. *The Review of Economic Studies 69*(1), 65–96.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review 103*(1), 305–34.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Clementi, G. L. and B. Palazzo (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics 8*(3), 1–41.

Coles, M. G. and D. T. Mortensen (2016). Equilibrium labor turnover, firm growth, and unemployment. *Econometrica 84*(1), 347–363.

Davis, S. J. and J. Haltiwanger (1992, 08). Gross Job Creation, Gross Job Destruction, and Employment Reallocation*. *The Quarterly Journal of Economics 107*(3), 819–863.

Decker, R. A., J. C. Haltiwanger, R. S. Jarmin, and J. Miranda (2018). Changing business dynamism and productivity: Shocks vs. responsiveness. Technical report, National Bureau of Economic Research.

den Haan, W. J. (2010). Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control 34*(1), 79 – 99. Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty.

Elsby, M. W. and R. Michaels (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics 5*(1), 1–48.

Elsby, M. W. L. and A. Gottfries (2019). Firm Dynamics, On-the-Job Search and Labor Market Fluctuations. Working Paper.

Engbom, N. (2019). Firm and worker dynamics in an aging labor market. Technical report.

Fort, T. C., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2013). How firms respond to business cycles: The role of firm age and firm size. *IMF Economic Review 61*(3), 520–559.

Foster, L., C. Grim, and J. Haltiwanger (2016). Reallocation in the great recession: cleansing or not? *Journal of Labor Economics 34*(S1), S293–S331.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001). Aggregate productivity growth: Lessons from microeconomic evidence. In *New developments in productivity analysis*, pp. 303–372. University of Chicago Press.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly journal of economics 114*(3), 739–767.

Gavazza, A., S. Mongey, and G. L. Violante (2018). Aggregate recruiting intensity. *American Economic Review 108*(8), 2088–2127.

Gouin-Bonenfant, E. (2019). Productivity dispersion, between-firm competition, and the labor share. Technical report.

Gu, R. (2019). Specific capital, optimal firm insurance, and the dynamics of the postgraduate wage premium. Technical report.

Haltiwanger, J. C., H. R. Hyatt, L. B. Kahn, and E. McEntarfer (2018). Cyclical job ladders by firm size and firm wage. *American Economic Journal: Macroeconomics 10*(2), 52–85.

Hornstein, A., P. Krusell, and G. L. Violante (2011). Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review 101*(7), 2873–98.

Kaas, L. and P. Kircher (2015, Oct). Efficient firm dynamics in a frictional labor market †. *American Economic Review 105*(10), 3030–3060.

Khan, A. and J. K. Thomas (2013). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy 121*(6), 1055–1107.

Krusell, P. and A. A. Smith, Jr (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy 106*(5), 867–896.

Lise, J. and J.-M. Robin (2017). The macrodynamics of sorting between workers and firms. *The American Economic Review 107*(4), 1104–1135.

Merz, M. and E. Yashiv (2007). Labor and the market value of the firm. *American Economic Review 97*(4), 141.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *The 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. *The American Economic Review 102*(6), 2509–2539.

Moscarini, G. and F. Postel-Vinay (2013, October). Stochastic Search Equilibrium. *The Review of Economic Studies 80*(4), 1545–1581.

Moscarini, G. and F. Postel-Vinay (2016). Wage posting and business cycles: A quantitative exploration. *Review of Economic Dynamics 19*, 135–160.

Office for National Statistics (2019). Business Structure Database, 1997-2018: Secure Access. `http://doi.org/10.5255/UKDA-SN-6697-10`. 10th Edition, SN:6697.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Patterson, C., A. Şahin, G. Topa, and G. L. Violante (2016). Working hard in the wrong place: A mismatch-based explanation to the uk productivity puzzle. *European Economic Review 84*, 42–56.

Postel-Vinay, F. and A. Sepahsalari (2019). Labour mobility and earnings in the uk, 1992-2016. Technical report.

Reed, W. J. (2001). The pareto, zipf and other power laws. *Economics letters 74*(1), 15–19.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control 33*(3), 649–665.

Rujiwattanapong, W. S. (2019). Unemployment dynamics and endogenous unemployment insurance extensions.

Schaal, E. (2017). Uncertainty and unemployment. *Econometrica 85*(6), 1675–1721.

Sedláček, P. and V. Sterk (2017). The growth potential of startups over the business cycle. *American Economic Review 107*(10), 3182–3210.

Sepahsalari, A. (2016). Financial market imperfections and labour market outcomes.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 25–49.

Shimer, R. (2010). *Labor markets and business cycles.* Princeton, N.J: Princeton University Press.

Winberry, T. (2016). Lumpy investment, business cycles, and stimulus policy. *Job market paper, Princeton University/University of Chicago Booth School of Business*.

A Additional Derivations and Proofs

A.1 Size-independent Discounted Profits

We want to guess and verify that a solution to the functional equation (4) has the form $n_t\pi_t(p_t, \overline{V})$. That is, we want to show that

$$\Pi_{t+1}(p_{t+1}, n_{t+1}, W_{t+1}) = n_{t+1}\pi_{t+1}(p_{t+1}, W_{t+1}) \implies \Pi_t(p_t, n_t, \overline{V}) = n_t\pi_t(p_t, \overline{V}).$$

Start from (4), still subject to the Promise-Keeping constraint and the law of motion for its workforce. Plugging in the guess in (4) gives

$$E_t\Big[\chi_{t+1}(p_{t+1})\Big(- c(h_{t+1})(1-\mu)n_t + \Pi_{t+1}(p_{t+1}, n_{t+1}, W_{t+1})\Big)\Big]$$
$$= E_t\Big[\chi_{t+1}(p_{t+1})\Big(- c(h_{t+1})(1-\mu)n_t + n_{t+1}\pi_{t+1}(p_{t+1}, W_{t+1})\Big)\Big].$$

Now substitute the law of motion for the firm's workforce in the last expression. Note that with a continuum of workers, it is assumed to hold exactly condition on the firm surviving and $\rho_{t+1}(W_{t+1}),\ h_{t+1}$. This substitution would still work with a discrete number of workers as long as the law of motion holds in expectation, so the Law of Iterated Expectations can be applied conditioning on the realization of the shocks at the start of the period. Substituting n_{t+1} then yields

$$n_tE_t\Big[\ \chi_{t+1}(p_{t+1})\Big(- c(h_{t+1}) + (\rho_{t+1}(W_{t+1}) + h_{t+1})\pi_{t+1}(p_{t+1}, W_{t+1})\Big)\Big].$$

Using this last expression in the main profit equation, it follows directly that firm profits are linear in n_t, as shown in (7).

A.2 Derivation Match Surplus

Recall that the joint value of a match is defined as $S_t(p_t) := \pi_t(\overline{V}) + \overline{V}$. Rearranging the Promise-Keeping constraint gives an expression for w_t

$$w_t = \overline{V} - \beta E_t \Bigg\{ \mu Q_{t+1} + (1-\mu)\Big[(1-\chi_{t+1}(p_{t+1}))U_{t+1}$$
$$+ \chi_{t+1}(p_{t+1})\Big(\delta_{t+1}U_{t+1} + (1-\delta_{t+1})(1 - s\lambda_{t+1}\bar{F}_{t+1}(W_{t+1}))W_{t+1} + s\lambda_{t+1}\int_{W_{t+1}}^{\infty} \theta dF_{t+1}(\theta)\Big)\Big]\Bigg\}.$$

Substituting w_t in the expression for firm profit per worker (7) gives

$$\begin{aligned}
S_t(p) &:= \pi_t(p, \overline{V}) + \overline{V} \\
&= -\overline{V} + \max_{\substack{h_{t+1}\geq 0 \\ W_{t+1}}} \Bigg\{ p_t\omega_t \\
&+ \beta E_t\Big[\mu Q_{t+1} + (1-\mu)\Big((1-\chi_{t+1}(p_{t+1}))U_{t+1} \\
&+ \chi_{t+1}(p_{t+1})\big(\delta U_{t+1} + (1-\delta)((1 - s\lambda_{t+1}\bar{F}_{t+1}(W_{t+1}))W_{t+1} + s\lambda_{t+1}\int_{W_{t+1}}^{\infty} \theta dF_{t+1}(\theta)) \\
&- c(h_{t+1}) + (\rho_{t+1}(W_{t+1}) + h_{t+1})\pi_t(p_{t+1}, W_{t+1})\big)\Big)\Big]\Bigg\} + \overline{V}.
\end{aligned}$$

Finally, taking the max operator inside the expectation and grouping terms yields

$$\begin{aligned}
S_t(p) = p_t\omega_t + \beta E_t\Bigg[&\mu Q_{t+1} + (1-\mu)\Big((1-\chi_{t+1}(p_{t+1}))U_{t+1} \\
&+ \chi_{t+1}(p_{t+1}) \max_{\substack{h_{t+1}\geq 0 \\ W_{t+1}}} \big\{ -c(h_{t+1}) + \rho_{t+1}(W_{t+1})S_{t+1}(p_{t+1}) + h_{t+1}(S_{t+1}(p_{t+1}) - W_{t+1}) \\
&+ (1-\delta)s\lambda_{t+1}\int_{W_{t+1}}^{\infty} \theta dF_{t+1}(\theta)\big\}\Big)\Bigg].
\end{aligned}$$

A.3 Definition Acceptance Rate

Define G_t the share of workers employed at firms offering contract value less than W in the current period

$$G_t(W) := \frac{\displaystyle\int \mathbb{1}\left\{W_t(p) \leq W\right\} \chi_t(p) n d\nu_t}{\displaystyle\int \chi_t(p) n d\nu_t}.$$

The acceptance rate at some offered W is then given by

$$Y_t(W) := \frac{u_t + s(1-\delta) G_t(W) \displaystyle\int \chi_t(p) n d\nu_t}{u_t + s(1-\delta) \displaystyle\int \chi_t(p) n d\nu_t},$$

where the numerator is the (intensity-weighted) measure of workers currently employed at firms offering contracts less than W and the denominator is the total measure of such workers.

A.4 Proof Rank-Monotonic Equilibrium

The outline of the proof is similar to that in Moscarini and Postel-Vinay (2013, 2016). The key difference is that the firm's problem can be considered separately for each worker since $\Pi_t\left(p, n, \overline{V}\right) = n\pi_t\left(p, \overline{V}\right)$. There is therefore no need to show super-modularity of the firm-worker surplus in its productivity and own size. It is enough to show that the firm-worker surplus is increasing in p, which implies that the optimal contract is also increasing in p, conditional on some convexity requirements of the cost of hiring function. We want to prove the two following statements:

1. Conditional on S being increasing in p, $\frac{c''(h)h}{c'(h)} \geq 1, \forall h \geq 0$ is sufficient to guarantee that V is increasing in p;
2. The firm-worker surplus mapping defined by (10) implies that S is increasing in p.

Taking each point in order:

1. Sufficient conditions on c for a RME Conditional on the firm surviving, the maximization problem associated with (10) defines the optimal contract and hiring rate after all histories. At any interior maximum, the following first-order conditions are associated with (11)

$$
\begin{aligned}
[h]: &\quad c'(h) = S(p) - W \\
[W]: &\quad \rho'(W)(S(p) - W) = h,
\end{aligned}
$$

where I have dropped the time subscripts, but S and ρ implicitly depend on calendar time in what follows. In addition, at any maximum, the associated Hessian matrix, H, is negative-definite, which requires

$$
\det(H) = -c''(h)\Big(\rho''(W)\big(S(p) - W\big) - \rho'(W)\Big) - 1 > 0.
$$

The two FOCs can be combined to give the following expression in W

$$
-c'\big(\rho'(W)\big(S(p) - W\big)\big) + S(p) - W = 0
$$

and totally differentiating that last expression with respect to p gives

$$
\frac{dW}{dp} = \frac{\frac{\partial S(p)}{\partial p}\big(c''(h)\rho'(W) - 1\big)}{\det(H)}.
$$

In this last expression, the denominator is positive at any maximum. By assumption, the firm-worker surplus is increasing in p, so $\frac{\partial S(p)}{\partial p} \geq 0$. Noting that the two FOCs can be combined to give $\rho'(w) = \frac{h}{c'(h)}$, it follows that

$$
\frac{dW}{dp} \geq 0 \iff c''(h)\rho'(W) \geq 1 \iff \frac{c''(h)h}{c'(h)} \geq 1.
$$

2. Firm-worker surplus increasing in p In this part of the proof, we want to show that S is increasing in p, which was assumed in the previous part. I follow the proof strategy outlined in (Moscarini and Postel-Vinay, 2013, Appendix A) and start by showing that the mapping defined by (10) maps from the space of differentiable, bounded and increasing functions into itself, conditional on a constant measure of firms $\nu(p,n)$. With this condition, the Continuous Mapping Theorem can be applied, so the net-surplus defined by the mapping exists, is unique, and increasing in p.

In a second step, the condition on ν is relaxed. In this case, the Continuous Mapping Theorem cannot be applied, as S is no longer defined on $\mathcal{R}^N$. But, since it is known that S is increasing in p in the restrictive case and that this solution is unique, we know that every candidate solution of the unrestricted mapping should have the property as well.

In the remainder of the proof, we then fix the beginning of period measure of firms to some value. We want to show that the mapping defined by (10) maps from the space of differentiable, bounded and increasing functions into itself. Differentiability in p follows directly from noting that the expectation in (10) is differentiable in p as long as the conditional probability density of future productivity is. This can be assumed. Since the support of p is convex and closed, it also follows that the mapping defined in (10) maps into the set of bounded functions.

Finally, to show that the mapping is increasing in p, first note that, for continuing firms, the envelope condition on the firm's optimization problem (11) gives

$$\frac{d\psi_{t+1}(p)}{dp} = \frac{\partial \psi_{t+1}(p)}{\partial p} = \Big(\rho_{t+1}(V^*) + h^*\Big)\frac{\partial S_{t+1}(p)}{\partial p} \geq 0,$$

where V^*, h^* denote optimal policies. The term inside the expectation in the firm-worker surplus (10) is then weakly increasing in p: constant on the part of the support of p_{t+1} where the firm exits, and weakly increasing otherwise.

To complete the proof, an additional assumption is needed on the idiosyncratic productivity shock. Namely, it has to be assumed that given a higher realization of productivity in the current period, the conditional Cumulative Distribution Function of future productivity satisfies first-order stochastic dominance.

With this assumption, conditional on any two distinct previous realizations of p, the conditional densities of future idiosyncratic productivity satisfy a single-crossing property. Let p_0 denote this crossing point and let $p_1,\ p_2$ be two values in $[\underline{p}, \overline{p}]$ such that $p_2 > p_1$, then

$$S_t(p_2) - S_t(p_1) = \omega_t(p_2 - p_1) + \beta(1-\mu)\Bigg(E_t\Big[\kappa_{t+1}(p)\ |p_2\Big] - E_t\Big[\kappa_{t+1}(p)\ |p_1\Big]\Bigg),$$

where $\kappa_{t+1}(p)$ is a notation for the terms inside the expectation

$$\kappa_{t+1}(p) := (1 - \chi_{t+1}(p))U_{t+1} + \chi_{t+1}(p)\Big(\delta U_{t+1} + \psi_{t+1}(p)\Big).$$

(The μQ_{t+1} terms are independent of the previous value of p, so they cancel.) Showing that S_t is increasing in p now amounts to show that the difference in expectation in the last expression is non-negative. This difference can be rewritten

$$\int_{\underline{p}}^{\overline{p}} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp,$$

where $\gamma(p|p_i)$ is the density of the p-shock conditional on p_i.

Now, given the crossing-point p_0, we can rewrite

$$\begin{aligned}&\int_{\underline{p}}^{\overline{p}} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp \\ &= \int_{\underline{p}}^{p_0} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp + \int_{p_0}^{\overline{p}} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp\end{aligned}$$

and, since $E_t\Big[\kappa_{t+1}(p)\Big]$ is weakly increasing in p, we get the following inequalities

$$\int_{\underline{p}}^{p_0} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp \geq E_t\Big[\kappa_{t+1}(p)\Big]\int_{\underline{p}}^{p_0}\Big(\gamma(p|p_2) - \gamma(p|p_1)\Big)dp$$

and

$$\int_{p_0}^{\overline{p}} E_t\Big[\kappa_{t+1}(p_0)\Big]\Big(\gamma(p|p_2)-\gamma(p|p_1)\Big)dp \geq E_t\Big[\kappa_{t+1}(p_0)\Big]\int_{p_0}^{\overline{p}}\Big(\gamma(p|p_2)-\gamma(p|p_1)\Big)dp.$$

Finally, summing up the last two inequalities, we get

$$E_t\Big[\kappa_{t+1}(p)\ |p_2\Big]-E_t\Big[\kappa_{t+1}(p)\ |p_1\Big]=\int_{\underline{p}}^{\overline{p}} E_t\Big[\kappa_{t+1}(p)\Big]\Big(\gamma(p|p_2)-\gamma(p|p_1)\Big)dp \geq 0,$$

which shows that $S_t(p_2) \geq S_t(p_1)$ for $p_2 > p_1$.

A.5 RME Contracts

This Appendix proves that the RME contract has the form given in (16). Before turning to the actual proof, I first show that the contract offer distribution, F_t, rewrites

$$\begin{aligned}F_t(W) &:= A_t^{-1}\int \mathbb{1}\left\{W_t(p)\leq W\right\}\chi_t(p)a_t(p,n)d\nu_t\\ &= \int \mathbb{1}\left\{W_t(p)\leq W(p)\right\}\frac{\chi_t(p)(1-\mu)n}{Z_t\lambda_t Y_t(W)}d\nu_t,\end{aligned}$$

where the substitution follows from the firm's vacancy posting position (15) and the equality $\eta_t A_t = \lambda_t Z_t$. Besides, in a RME, contracts are strictly increasing in p, so we have

$$G_t(W_t(p)) = \frac{\displaystyle\int_{\underline{p}}^{p}\chi_t(p')dL_t(p')}{\displaystyle\int_{\underline{p}}^{\overline{p}}\chi_t(p')dL_t(p')} = \frac{L_t(p)-L_t(p_E)}{L_t(\overline{p})-L_t(p_E)},$$

where p_E denotes firm's entry/exit threshold and the acceptance rate can now be simplified as

$$Y_t(V_t(p)) = \frac{u_t+s(1-\delta)\Big(L_t(p)-L_t(p_E)\Big)}{u_t+s(1-\delta)\Big(L_t(\overline{p})-L_t(p_E)\Big)}.$$

Finally, plugging this last expression into the contract offer distribution evaluated at $V(p)$ gives Equation (17)

$$\lambda_t F_t\left(V(p)\right) = \int_{p_E}^{p} \frac{h_t(p')}{u_t + s(1-\delta)\Big(L_t(p') - L_t(p_E)\Big)} dL_t(p').$$

To get (16), start from the first-order condition with respect to the optimal contract from (11) for active firms at some productivity level p

$$[W]: \quad \rho'(W)(S(p) - W) = h,$$

where I drop the time subscripts on ρ, S, but these functions depend implicitly on ω and L. The derivative of the retention rate is given by

$$\rho'(W) = s(1-\delta)\lambda \frac{dF(W)}{dW},$$

and, in a Rank-Monotonic Equilibrium, the derivative of the offer function can be expressed from (17) as

$$\lambda \frac{dF(W)}{dW}\frac{dW}{dp} = \frac{hl(p)}{u + s(1-\delta)\big(L(p) - L(p_E)\big)}.$$

Combining these three expressions yields the following first-order differential equation in W

$$\frac{dW}{dp} + \frac{s(1-\delta)l(p)}{u + s(1-\delta)\big(L(p) - L(p_E)\big)} W = \frac{s(1-\delta)l(p)}{u + s(1-\delta)\big(L(p) - L(p_E)\big)} S(p).$$

with boundary condition $W(p_E) = U$. Noting that

$$\frac{d\ln\Big(u + s(1-\delta)\big(L(p) - L(p_E)\big)\Big)}{dp} = \frac{s(1-\delta)l(p)}{u + s(1-\delta)\big(L(p) - L(p_E)\big)},$$

the corresponding integrating factor is then

$$\exp \int \frac{s(1-\delta)l(p)}{u+s(1-\delta)\big(L(p)-L(p_E)\big)}dp = u+s(1-\delta)\big(L(p)-L(p_E)\big).$$

Along with the boundary condition, this yields (16) in the main text

$$W(p) = \frac{uU + s(1-\delta)\displaystyle\int_{p_E}^{p} S(p')dL(p')}{u+s(1-\delta)\big(L(p)-L(p_E)\big)}.$$

A.6 Derivations Net Surplus

This Appendix shows that the model can be recast in a single value function by subtracting the unemployed worker's value function to the firm-worker surplus. I omit it from the main text not to clutter the description of the model. However, this more compact formulation is used in solving and simulating the model since the firm's policies can all be expressed as a function of the net surplus.

Net Surplus Equation The net firm-worker surplus is defined as $\phi_t(p) := \pi_t + \overline{V} - U_t := S_t(p) - U_t$. Adding and subtracting U_{t+1} in (10), the firm-worker surplus can be rewritten

$$\begin{aligned} S_t(p) = p_t\omega_t + \beta E_t \Bigg[& U_{t+1} + \mu Q_{t+1} \\ & + (1-\mu)\Bigg(\chi_{t+1}(p_{t+1}) \max_{\substack{h_{t+1}\geq 0 \\ W_{t+1}}} \big\{ -c(h_{t+1}) + \rho_{t+1}(W_{t+1})\phi_{t+1}(p_{t+1}) \\ & + h_{t+1}(\phi_{t+1}(p_{t+1}) - (W_{t+1}-U_{t+1})) + (1-\delta)s\lambda_{t+1} \int_{W_{t+1}}^{\infty} \theta - U_{t+1} dF_{t+1}(\theta)\big\}\Bigg)\Bigg]. \end{aligned}$$

Using the same strategy, the unemployed worker's value can similarly be rearranged as

$$U_t = b + \beta E_t \Big[U_{t+1} + \mu Q_{t+1} + (1-\mu)\lambda_{t+1} \int \max\{\theta - U_{t+1}, 0\}\, dF_{t+1}(\theta) \Big].$$

The net surplus can then be expressed as

$$\phi_t(p) := S_t(p) - U_t = p_t\omega_t - b + \beta(1-\mu)E_t\Bigg[\chi_{t+1}(p_{t+1})\Big\{\tilde{\psi}_{t+1}(p) - \lambda_{t+1}\int_0^\infty \theta d\tilde{F}_{t+1}(\theta)\Big\}\Bigg] \tag{21}$$

where $\tilde{F}_{t+1}$ defines the offer distribution for the firm's contract net of the value of unemployment, and $\tilde{\psi}_t(p)$ is the firm's optimization problem in net surplus form

$$\tilde{\psi}_t(p) := \max_{\substack{h\geq 0\\ V}}\Big\{-c(h) + \rho_t(V)\phi_t(p) + h\Big(\phi_t(p) - V\Big) + (1-\delta)s\lambda_t\int_V^\infty \theta d\tilde{F}_t(\theta)\Big\}$$

where, the firm now picks a contract V net of the value of unemployment.

Firm policies as a function of ϕ in a RME Since $\phi = S - U$ and U does not depend on p, ϕ is also increasing in p for every candidate equilibrium. In a Rank-Monotonic Equilibrium, the corresponding net contract follows by subtracting $U(\omega, L)$ in (16), which gives

$$V(p,\omega,L) - U(\omega,L) := \tilde{V}(p,\omega,L) = \frac{s(1-\delta)\int_{p_E}^{p}\phi(\hat{p},\omega,L)dL(\hat{p})}{u + s(1-\delta)\Big(L(p) - L(p_E)\Big)}, \tag{22}$$

The optimal hiring rate can also be expressed as solving

$$c'(h(p,\omega,L)) = \phi(p,\omega,L) - \tilde{V}(p,\omega,L),$$

and the entry/exit decision as $\chi(p,\omega,L) = \mathbb{1}\{\phi(p,\omega,L) \geq 0\}$.

B Data Appendix

B.1 Firm Data: The Business Structure Database

Variables. I gather the definitions of the main analysis variables here. Note that a given variable is potentially drawn from multiple sources depending on whether the enterprise is selected to be part of a survey in the last year. [22]

- Employment: Sum of employees and working proprietors. This variable comes from different sources, but, for the majority of firms, employment is derived from income tax data – which is deduced directly from pay in the UK. For these firms, the employment figure corresponds to either the last or four last available quarters when the snapshot is taken, between March and April each year.

- sales: Income from the "sale of good or services to third parties". These figures are net of VAT, but include other taxes (alcohol, tobacco). For the majority of businesses, these sales figures are drawn from VAT returns for the past financial year, which ends in early April each year.

- Industry: The Standard Industry Classification is updated twice over the sample period, in 2003 and 2007. Since this classification is given, in most cases, in the contemporaneous vintage, I convert all industries to the 2007 classification by i) directly assigning their SIC07 industry code for firms that survive until then, ii) creating crosswalks from the SIC92 to SIC03 and SIC03 to SIC07, based on the firms surviving across present both before and after each respective update comes into effect.

- Age: I follow Fort et al. (2013) in deriving a firm's age from establishment data. Each establishment has a birth year which corresponds to when they first appear on the registers. I define firm age as the age of the firm's oldest establishment when the firm

[22] I am grateful to Davide Melcangi and the research support team at ONS for clarifying the timing of some of these variables.

first becomes active. It then ages naturally from this point onward, for each extra year in the sample. An advantage of this definition is that it avoids artificially classifying as "young" firms appearing in the data as a result of mergers, changes of ownership, etc.

- Labor Productivity: As discussed in the main text, labor productivity is defined as the logarithm of sales over employment.

Validation with national statistics aggregates. To assess the accuracy of the aggregates derived from the Business Structure Database, I compare some of these aggregates series with the closest official series from the Office for National Statistics. Figure 12 reports two such benchmarks: employment and sales. The corresponding ONS series are, respectively, workforce jobs and "domestic output at basic prices".[23] As shown in Figure 12, some sectors have trends at odds with the official series, especially for the sales variable. I proceed by excluding the following aggregate sectors: B (mining and quarrying), K (finance and insurance), M (professional and technical services), R (arts and entertainment). I also drop sectors O-Q (public administration, education, and health), the last two being mostly public in the UK.

Analysis sample. I finally apply a set of restrictions to construct a panel of firms over available survey years. I drop all firms that report sales or employment zero in any given year. I also drop firms which do not report hiring anyone over all survey year. Figure 13 shows aggregate employment and sales for the analysis sample and the official aggregates from the Office for National Statistics.

B.2 Labor Productivity Measure

As discussed in the main text, the Business Structure Database only makes turnover available for each firm. Figure 14 benchmarks the turnover-based labor productivity measure against

[23] "domestic output at basic prices" relates to sales in the BSD, as it corresponds to an industry's gross output (not net of intermediary consumptions).

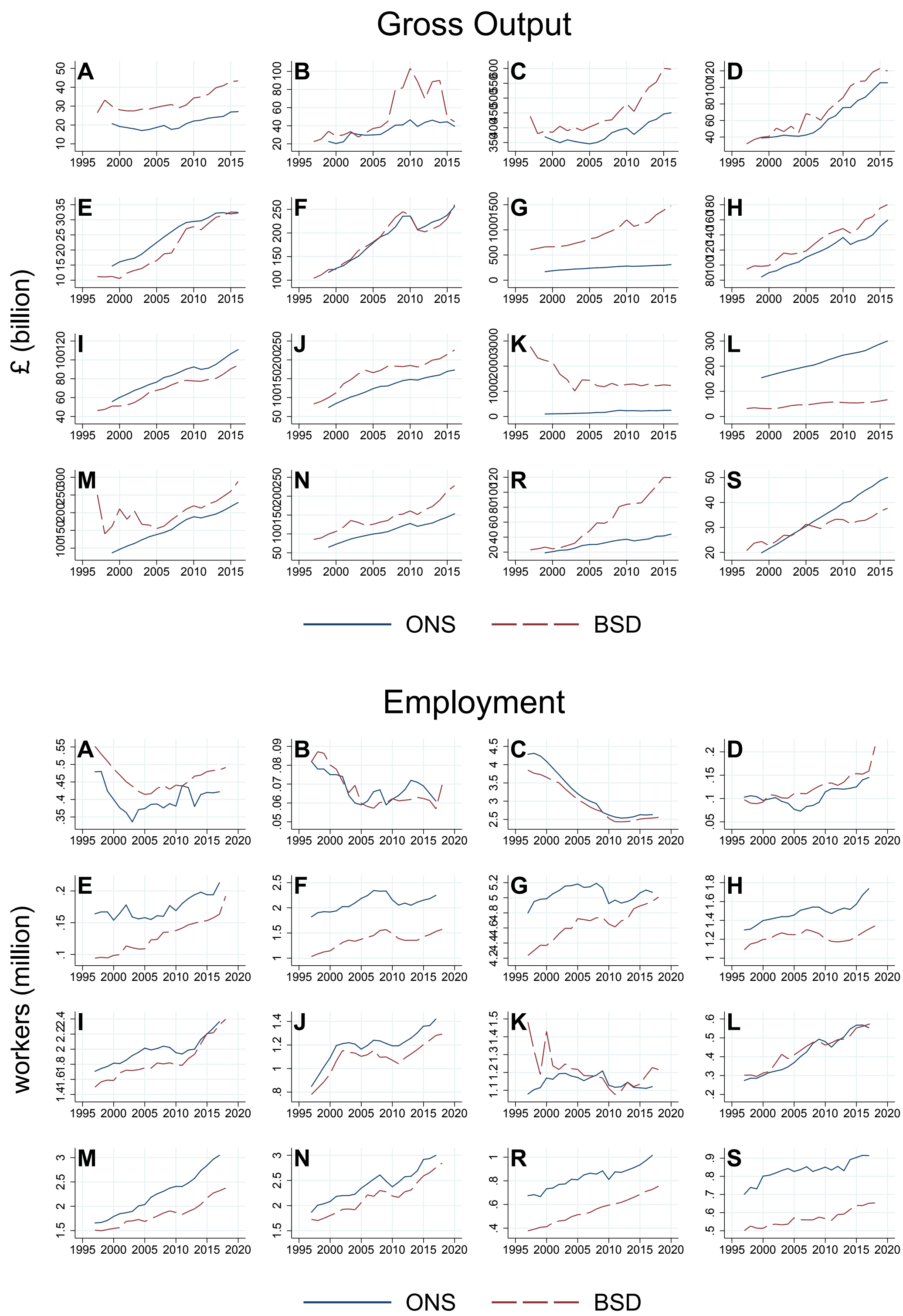

Figure 12: Benchmark with official statistics by broad industry. Comparison with official aggregate series by broad industry in the SIC07 classification (in bold, top left corner). The corresponding SIC07 industries are given in Table 5. See main text for details on the definition of these series.

SIC07 Section	Description
A	Agriculture, Forestry and Fishing
B	Mining and Quarrying
C	Manufacturing
D	Electricity, Gas, Steam and Air Conditioning Supply
E	Water Supply; Sewerage, Waste Management and Remediation Activities
F	Construction
G	Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles
H	Transportation and Storage
I	Accomodation and Food Service Activities
J	Information and Communication
K	Financial and Insurance Activities
L	Real Estate Activities
M	Professional, Scientific and Technical Activities
N	Administrative and Support Service Activities
O	Public Administration and Defence; Compulsory Social Security
P	Education
Q	Human Health and Social Work Activities
R	Arts, Entertainment and Recreation
S	Other Service Activities

Table 5: Description of SIC07 broad industries.

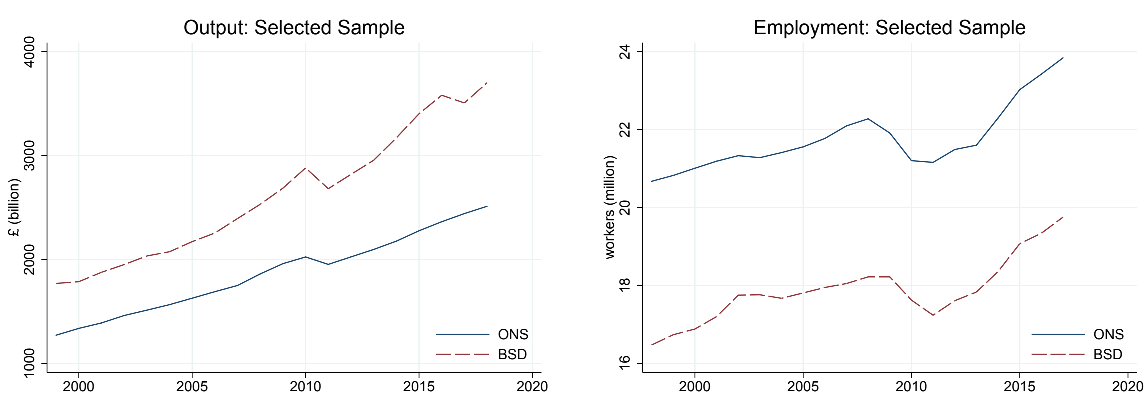

Figure 13: Aggregates from analysis sample.

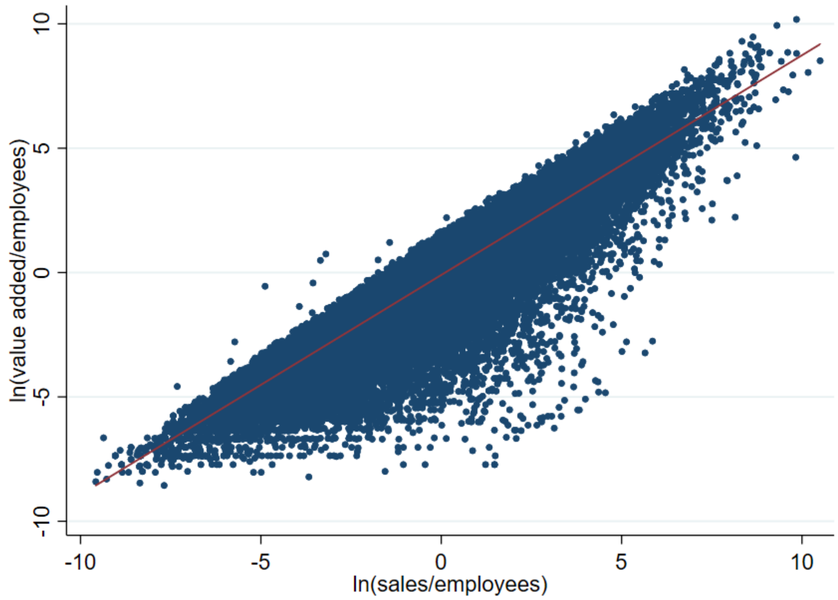

Figure 14: Correlation between turnover and value added labor productivity measures. These measures are defined, respectively, as the log of turnover per employee and turnover minus cost of sales per employee, deviated from industry means.

a value-added based labor productivity measure. The data is from Fame, a commercial database of company information, including detailed balance sheet data, for the UK and Ireland. The figure shows that, within industries, these two labor productivity measures are strongly associated.

B.3 Labor Market Transitions

The labor market transition rates are taken from Postel-Vinay and Sepahsalari (2019). They are derived from the British Household Panel Survey (BHPS) and its successor Understanding Society (UKHLS). Note that because of the transition from the BHPS to UKHLS, there is a gap in the series between August 2008 and December 2009, which is smoothed over using moving averages.[24]

[24] I am grateful to the authors for sharing these series, and to Pete Spittal for explaining how the transition between the two surveys affects them.

B.4 Additional Macro Series

Several additional series are taken directly from the Office for National Statistics website:

- Unemployment rate (aged 16 and over, seasonally adjusted): MGSX
- UK vacancies - total: AP2Y

C Numerical Solution

C.1 Stationary solution

As shown in Appendix A.6, the firm's policies can be expressed in terms of a single value function, the net surplus given in Equation (21). A Stationary Rank-Monotonic Equilibrium (see Definition 3) can similarly be defined as a fixed-point in the net surplus, ϕ and the measure of workers, L. The algorithm below is given in terms of net firm-worker surplus for concision.

Discretization. In a Rank-Monotonic Equilibrium, all heterogeneity in the model arises through p. I discretize idiosyncratic productivity using Tauchen's procedure with $N_p = 400$ points. This yields a $\{p_1, \ldots, p_{N_p}\}$ grid and the associated transition matrix for p.

This discretization can seen as the relevant policy or value function being constant on some (small) half-open interval. This provides an intuitive way to integrate against the measure of workers, Lt, by replacing the integral by the appropriate employment share weighted sum. For instance, the net optimal contract (22) at some productivity node p_k can be approximated as

$$\tilde{V}(p_k) = \frac{s(1-\delta)\displaystyle\int_{p_1}^{p_k} \chi(p')\phi(p')dL(p')}{u + s(1-\delta)\Big(L(p_k) - L(p_E)\Big)}$$

$$= \frac{s(1-\delta)\displaystyle\sum_{i=2}^{k}\int_{p_{i-1}}^{p_i} \chi(p')\phi(p')dL(p')}{u + s(1-\delta)\Big(L(p_k) - L(p_E)\Big)}$$

$$\approx \frac{s(1-\delta)\displaystyle\sum_{i=2}^{k}\chi(p_{i-1})\phi(p_{i-1})\int_{p_{i-1}}^{p_i} dL(p')}{u + s(1-\delta)\Big(L(p_k) - L(p_E)\Big)},$$

where the last integral in the approximation is simply the fraction of workers employed at firms in the interval between p_{i-1} and p_i.

Algorithm stationary equilibrium. Given this discretization, I iterate on the following steps:

1. Guess initial values for ϕ and L on the grid for idiosyncratic productivity. In line with the RME result, I start with some increasing function for the net surplus. In practice, I set $L = 0$ (all workers initially unemployed) as a first step.

2. Conditional on values for ϕ and L, the agents' optimal policies can be computed. For example, the activity threshold, p_E, is the point at which ϕ becomes positive. The optimal contract can be computed from (22).

3. The net surplus equation and the law of motion for employment shares imply new values for ϕ and L on the grid. Note that the net surplus equation gives an update for ϕ in the previous period, while that for the employment mass yields next period's employment for each productivity level. But this does not matter since the algorithm solves for a stationary equilibrium.

4. The final step consists in computing the Euclidean norm to check the convergence of L and ϕ. If this is the case, the pair (ϕ, L) represents a stationary equilibrium. Otherwise, go back to point 2 with the updated values until convergence.

C.2 Estimation

The parameters are calibrated by targeting the moments listed in Table 2. In practice, I minimize the distance between the model generated moments and their empirical counterpart using the following objective function

$$(\mathbf{m}_{\text{data}} - \mathbf{m}_{\text{model}}(\theta))^{\intercal}\, \mathbf{\Lambda}\, (\mathbf{m}_{\text{data}} - \mathbf{m}_{\text{model}}(\theta))$$

where θ denotes the parameter vector, $\mathbf{m}_{\text{data}}$ the vector of data moments, and $\mathbf{m}_{\text{model}}(\theta)$ the corresponding model generated vector of moments. Each moment is rescaled by the inverse of the square of its empirical value: $\mathbf{\Lambda} = \text{diag}(1/\mathbf{m}^2_{\text{data}})$. Figure 15 further shows slices of the objective function around the estimated parameter values.

C.3 Aggregate shocks solution

As explained in the main text, the simulation algorithm in the presence of aggregate shocks relies on two approximations. First, the measure of employment at firms of different productivity is summarized by a set of (un-centered) moments and the unemployment rate:

$$\begin{aligned}
u_t &= 1 - \int_{\underline{p}}^{\overline{p}} \chi_t(p) dL_t(p) \\
m_t^1 &= \int_{\underline{p}}^{\overline{p}} \chi_t(p) \ln p \frac{dL_t(p)}{1-u_t} \\
m_t^2 &= \int_{\underline{p}}^{\overline{p}} \chi_t(p) (\ln p)^2 \frac{dL_t(p)}{1-u_t} \\
&\text{etc.}
\end{aligned} \tag{23}$$

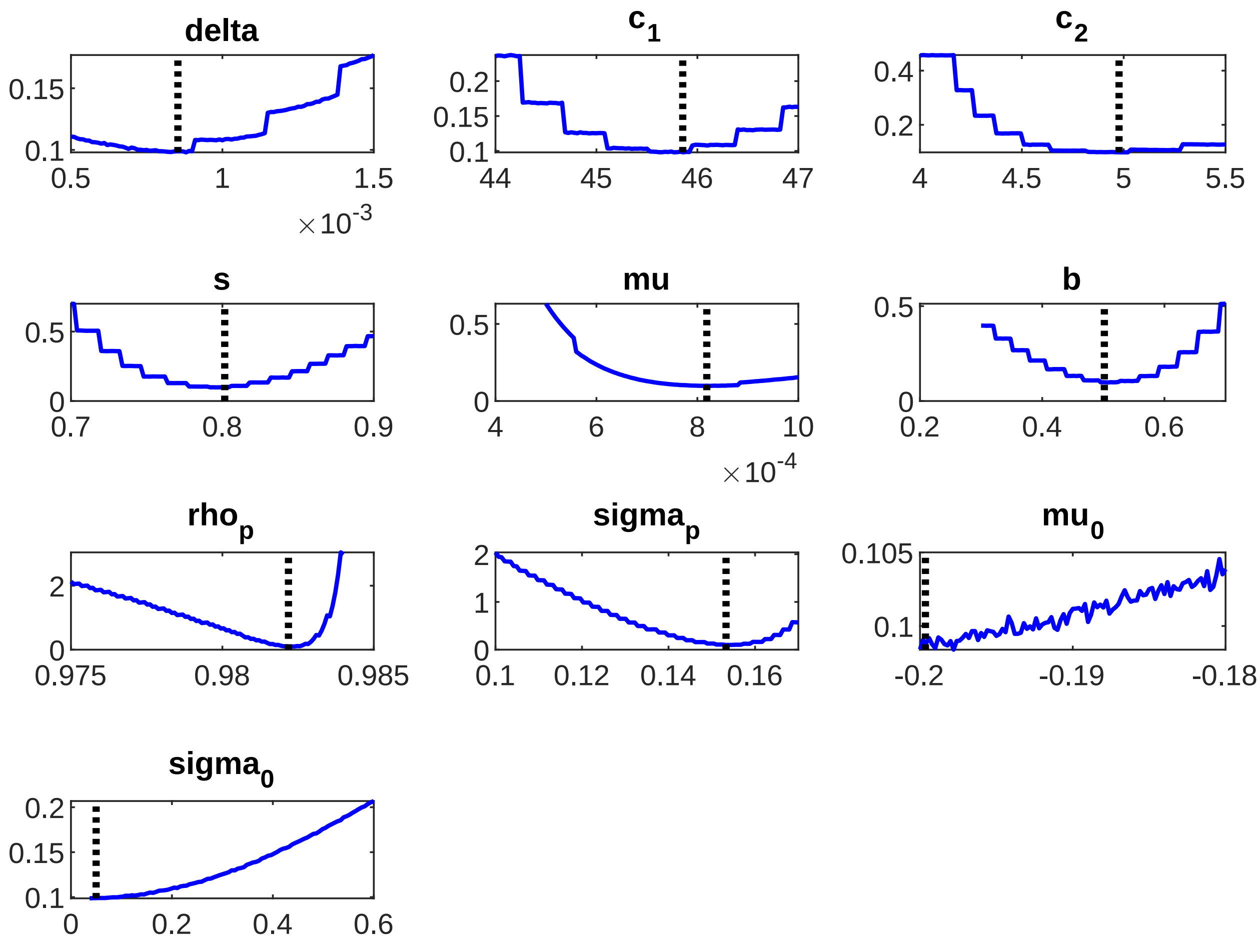

Figure 15: Slices of objective function for each parameter. Vertical dotted line denotes estimated parameter value.

In practice, given the Gaussian nature of the idiosyncratic shock, the stationary distribution of employment on the log-productivity grid has a normal shape – up to the truncation at its lower end implied by the entry and exit threshold. I then center on the first two moments in my implementation.

Second I parameterize the value functions for the firm-worker surplus, S_t, and the unemployed worker, U_t, with a polynomial.[25] Because preserving the monotonicity of S_t (especially around the entry threshold) is central to the procedure, I use a separate polynomial for each Tauchen node p_i. The value functions are approximated outside of steady-state as

$$\ln S(p_i, \omega_t, L_t) - \ln \overline{S}(p_i) \approx \tilde{S}(p, \omega_t, \tilde{\boldsymbol{m}}_t; \theta_{p_i}) \quad p_i \in \{p_1, ..., p_{N_p}\}$$

and

$$\ln U(\omega_t, L_t) - \ln \overline{U} \approx \tilde{U}(\omega_t, \tilde{\boldsymbol{m}}_t; \theta_U)$$

where $\tilde{\boldsymbol{m}}_t$ denotes the vector of moments in (23) in log-deviation from steady-state.

The algorithm then solves for the coefficients by iterating on the four following steps:

1. Draw a sequence of aggregate productivity shocks and guess an initial value for the coefficients of $\tilde{S}$ and $\tilde{U}$. I initialize them at zero in practice.

2. Simulate the measure of employment forward, starting from the stationary solution. Conditional on the current value of $\boldsymbol{\theta}$, agents make optimal hiring and contract offer decisions given the current states, which induces a law of motion for employment at each productivity level. The simulated measure of workers is approximated by a set of moments as described above.

3. Update $\tilde{S}$ and $\tilde{U}$, conditional on the simulation of L_t obtained in the previous step. This requires to take an expectation over future realizations of the aggregate shock.

[25] I choose to parameterize these value functions separately instead of the net surplus since they are positive by definition.

The aggregate shocks is discretized using Tauchen procedure with $N_\omega = 19$ nodes in practice.

4. Run a regression of $\tilde{S}$ and $\tilde{U}$ on the state variables to update the coefficients. Go back to step 2 and iterate until convergence.

I find the coefficients by running separate regressions for the firm-worker surplus at each p-node on the variables in the state-space. I omit the constant, thus imposing that the steady-state holds exactly at each node. Since these regressors are sometimes close to collinear, I make the procedure more robust by using ridge regression to regularize the problem. For instance, the coefficients for the unemployed worker's value function are found by solving

$$\min_{\theta_U} \sum_t (\ln U_t - \ln \overline{U} - \tilde{U}(\omega_t, \tilde{\boldsymbol{m}}_t; \theta_U))^2 + \zeta \sum_i \theta_{U_i}^2$$

where θ_{U_i} denotes individual elements of θ_U, $\zeta > 0$ is the associated regularization parameter, and

$$\tilde{U}(\omega_t, \tilde{\boldsymbol{m}}_t; \theta_U) = (\ln \omega_t - 0)\, \theta_U^\omega + \left(\ln m_t^1 - \ln \bar{m}^1\right) \theta_U^{m_1} + \left(\ln m_t^2 - \ln \bar{m}^2\right) \theta_U^{m_2}.$$

The regularization parameter, ζ¿0, ensures that the matrix of regressors is invertible by adding to it a ζ-diagonal matrix. I finally allow for less than full updating by appropriately dampening the obtained coefficients. I proceed similarly for each polynomial of the firm-worker surplus. Note that with these parametric assumptions, the coefficients $\{\theta_U, \theta_{p_1}, \dots, \theta_{p_{N_p}}\}$ are elasticities, which gives some intuition about the appropriate convergence condition.

C.4 Accuracy tests

The accuracy of the procedure is assessed through the tests proposed in den Haan (2010), adapted to the current setting. I compute the firm-worker surplus, $S_t(p)$ and unemployment

	Absolute Error (in %)	
Variable	Mean	Max
Value Functions		
S_t	0.060	0.284
U_t	0.027	0.128
Moments L_t ($\boldsymbol{m}_t$)		
u_t	1.177	4.274
m_t^1	0.048	0.358
m_t^2	0.034	0.238

Table 6: Accuracy Tests

value, U_t in two different ways. Given a sequence of aggregate shocks $\{\omega_s\}_{s=1}^T$, $S_t(p)$ and U_t can be obtained either using their respective approximation based on θ_p and θ_u, or computed directly solving the model backward in time and explicitly taking an expectation over ω_{t+1} in each period.

Table 6 reports these statistics for an alternative sequence of shocks, different to the one used to solve for the coefficients. I report the average and maximum absolute percent error between the approximation and explicit solutions, i.e. $100(y_t^{\text{approx.}} - y_t^{\text{explicit}})$, taken at each point in time and each node, where $y_t^{\text{approx.}}$ denotes $\tilde{S}(p, \omega_t, \hat{\boldsymbol{m}}_t; \theta_p)$ or $\tilde{U}(\omega_t, \hat{\boldsymbol{m}}_t; \theta_U)$ as appropriate.