Efficient Firm Dynamics in a Frictional Labor Market *

Leo Kaas† Philipp Kircher‡

Abstract

We develop and analyze a labor market model in which heterogeneous firms operate under decreasing returns and compete for labor by posting long-term contracts. Firms achieve faster growth by offering higher lifetime wages, which allows them to fill vacancies with higher probability, consistent with recent empirical findings. The model also captures several other regularities about firm size, job flows and pay, and generates sluggish aggregate dynamics of labor market variables. In contrast to existing bargaining models, efficiency obtains on all margins of job creation and destruction, and the model allows a tractable characterization over the business cycle.

JEL classification: E24; J64; L11

Keywords: Labor market search, multi-worker firms, job creation and job destruction

December 18, 2013

*We thank Rüdiger Bachmann, Matthieu Bellon, Steven Davis, Jan Eeckhout, Mike Elsby, Jason Faberman, William Hawkins, Matthias Hertweck, Iourii Manovskii, David Ratner, Richard Rogerson, and Ludo Visschers, David Comerford for research assistance, as well as seminar and conference audiences at BI Oslo, Carlos III, Chicago, Chicago Fed, CEMFI, Cologne Macroeconomics Workshop, Columbia University, Deutsche Bundesbank, EEA (Oslo), Essex Economics and Music, ETH Zurich, Labour Market Search and Policy Applications (Konstanz), Leicester, NBER Summer Institute (Boston), Philadelphia Fed, Royal Holloway, SAET (Faro), SED (Montreal, Seoul), St. Gallen, St. Louis Fed, Tor Vergata Rome, Toulouse, UC Los Angeles, UC San Diego, UC Santa Barbara, University of Pennsylvania, Verein fuer Socialpolitik, Vienna Macroeconomics Workshop (Rome) and Yale. Kircher gratefully acknowledges support from the National Science Foundation and the European Research Council.

†Department of Economics, University of Konstanz, leo.kaas@uni-konstanz.de

‡Department of Economics, University of Edinburgh, philipp.kircher@ed.ac.uk.

1 Introduction

Search models of the labor market following the Diamond-Mortensen-Pissarides framework have traditionally abstracted from the role of firms, concentrating on the concepts of jobs and vacancies (see, e.g., Rogerson et al. (2005)). While a recent wave of contributions include firm size through decreasing returns in production, they rely on the standard assumption that vacancies are filled at a common matching rate which depends on aggregate market conditions but is independent of the characteristics of the firm that posts the job. In this paper we propose an alternative theory in which heterogeneous firms compete for workers through their wage announcements, which naturally implicates differential job-filling rates across firms. This theory predicts several relations for the cross-section of firms and for the time-variation over the business cycle that seem to match with recent empirical findings. It leads to a very different view of the efficiency of the labor market compared to the previous literature, while retaining a high level of tractability even in the presence of aggregate shocks.

Recent empirical evidence highlights that the probability of filling jobs depends on the characteristics of the firm. In the cross-section, Davis et al. (2013) show that firms expand faster not only by posting more vacancies, but especially by filling these vacancies at higher rates; for example, the job-filling rate almost doubles as monthly employment growth increases from 10% to 20%. Across time, they back out an aggregate measure of "recruiting intensity" that moves pro-cyclically, leading to a lower level of matching efficiency for a given labor market tightness in downturns.

Our theory models firms through decreasing returns to labor as in Hopenhayn and Rogerson (1993). In the labor market, we follow the competitive search literature (e.g. Moen (1997)) where employers can publicly post long-term wage contracts to attract unemployed workers. When a firm attracts more workers to its vacancies, the matching rate increases. In our setting with large firms, we allow the firms to choose the number of vacancies alongside the posted wage contracts, and it is in fact optimal for them to use both margins. Therefore, matching rates are not an aggregate object but are firm-specific. Growing firms decide to offer better contracts if it is increasingly costly to hire additional workers, which arises, for example, when recruitment takes up time of the existing workers (Shimer (2010)), so that firms expand their workforce slowly over time. We argue that this feature not only generates varying job-filling rates at the micro level, but also gives rise to sensible aggregate dynamics. Particularly, important labor market variables, such as the job-finding rate, react with delay to aggregate shocks. While such sluggish adjustment is consistent with the evidence from vector autoregressions (e.g. Fujita and Ramey (2007)), it is hard to reconcile with the textbook search and matching model (Shimer (2005)). In a quantitative assessment, our model tracks well both the cross-sectional variation as well as the business-cycle variation of recruiting intensity described by Davis et al. (2013). It also leads to slow adjustment of the aggregate job-finding rate and other desirable business-cycle properties.

Our view that firms can attract workers to their vacancies is aimed to capture the features mentioned above and to provide a framework to think about job creation and job destruction of heterogeneous firms in frictional labor markets. It formulates an alternative that contrasts with the prevailing workhorse model based on random search and bilateral bargaining pioneered by Stole and Zwiebel (1996) and Smith (1999).[1] One obvious difference between the models is the rate at which firms fill their jobs. In the existing contributions, this is governed by the aggregate matching function, so that firms can only hire more if they post more

[1]Subsequent work adopts this approach to study, for example, unemployment and efficiency (Bertola and Caballero (1994), Acemoglu and Hawkins (2013)), labor and product market regulation (Koeniger and Prat (2007), Ebell and Haefke (2009)), business cycles (e.g., Elsby and Michaels (2013), Fujita and Nakajima (2013)), and international trade and its labor market implications (Helpman and Itskhoki (2010)).

vacancies, which conflicts with the evidence cited above. Our model naturally focuses on both recruiting margins, the number of vacancies and their filling rate. Competition for workers on the second margin also leads to very different normative implications. In the bargaining frameworks, firms hire excessively in order to depress the wages of all their workers, yielding a within-firm externality (see e.g. Smith (1999)).[2] In our setting, contracts are long-term, eliminating the inefficiency within the firm, and we show that across firms the wage posting leads to a modified Hosios (1990) condition which ensures that the decentralized economy creates and destroys jobs efficiently both on the *extensive* margins of firm entry/exit and on the *intensive* margins of firm expansion/contraction. While the Hosios condition is at the heart of many efficiency arguments in the competitive search literature, the subtle nature of search markets does not always render it sufficient to induce constrained efficiency, especially when choices along different margins interact.[3] We are not aware of a formal efficiency result for large firms operating under decreasing returns.[4] Finally, we establish that our environment is particularly tractable, even outside of steady state. While one could possibly add recruiting intensity to existing bargaining models, the complications arising from such settings, especially in the presence of aggregate shocks, make this difficult. Tractability in our model arises from free entry of firms and competitive search. When a firm decides whether to hire and what contracts to offer, it needs to know

[2]In contrast to one-worker bargaining models, the inefficiency cannot be corrected by an appropriate level of the bargaining power parameter. Even with wage commitments, the randomness of the search process generates an across-firm externality that impedes efficiency (see Hawkins (2010)).

[3]Galenianos and Kircher (2009) study a setting where firms commit to wages but efficiency fails because of an intensive margin (search intensity) on the workers' side. Guerrieri (2008) introduces an intensive margin through moral hazard and finds efficiency in steady state but not out of steady state. These subtleties indicate a lack of an easily applicable general proof on which we could draw to establish efficiency in our context.

[4]Hawkins (2013) suggests such an outcome on the basis of a static model, but his results are complicated by the stochastic nature of the hiring process and they do not extend to dynamic settings with shocks. Menzio and Moen (2010) do not obtain efficiency because they focus on lack of commitment, and Garibaldi and Moen (2010) abstract from decreasing returns.

the workers' utility value of unemployment, as this defines the relevant outside option. This utility value generally depends on the distribution of other firms in the market, which is an infinite-dimensional object. In our setting, since workers can choose where to search for a job, they are indifferent between existing firms and new entrants, and the latter number adjusts to equate the marginal benefit to the entry costs, independent of the existing firms. This implies that only the current aggregate productivity enters the workers' utility value and hence the firms' optimization problem, eliminating the need for approximation techniques like those of Krusell and Smith (1998) that are usually necessary to study business cycles with heterogeneous firms (e.g. Elsby and Michaels (2013), Fujita and Nakajima (2013)). The fact that *individual* firms' policy functions jump with business cycle shocks does not imply, however, that important *aggregate* variables, such as the workers' job-finding rate, jump as well. To the contrary, the distribution of firms evolves slowly and many job openings are not governed by free entry. Hence, the aggregate job-finding rate and the vacancy-unemployment ratio feature a slow response to business-cycle shocks, as documented by Fujita and Ramey (2007) and Fujita (2011), as well as an imperfect correlation with aggregate productivity (Shimer (2005)).

The idea that policy functions are jump variables also feature in Pissarides (2000) for random search and in Shi (2009) and Menzio and Shi (2010, 2011) in competitive search, but in those settings there is entry at all wage contracts and the job-finding rate is a jump variable, perfectly correlated with aggregate productivity.[5] Since the link between firm-level dynamics and aggregate dynamics is important, we explore this feature in more detail in the quantitative section of this paper. Indeed we demonstrate that the calibrated model generates aggregate

[5] In Shi (2009) and Menzio and Shi (2010, 2011), firms are indifferent between all contracts and there is free entry at every contract. In our setting, the workers are indifferent between all wage contracts, but there is still free entry on the firms' side. This additional feature brings about the difference in some results, while retaining tractability.

labor market dynamics that are largely in line with the U.S. business cycle. In particular, aggregate measures of the vacancy yield and of the recruiting intensity show similar cyclicality and volatility as found by Davis et al. (2013).

Our work describes the recruitment behavior of firms competing for unemployed workers. One could envision additionally competition for employed workers. Burdett and Mortensen (1998), Postel-Vinay and Robin (2002) and Moscarini and Postel-Vinay (2013) explore this in random search environments, but the complexity of these models makes it difficult to study firm dynamics, as firms are usually assumed to face neither idiosyncratic nor aggregate shocks.[6] In the competitive-search literature, job-to-job mobility has been considered by Shi (2009), Menzio and Shi (2010, 2011), Garibaldi and Moen (2010) and recently Schaal (2010). Except for the last contribution, firm size in these models is not restricted by the operated technology, circumventing considerations induced by the difference between average and marginal product. Schaal (2010) differs from ours by assuming linear recruitment costs, which imply that firms immediately jump to their desired sizes, they are indifferent between all contracts and hence face identical job-filling rates, and there is no aggregate sluggishness.

To build intuition for our model and to highlight its features, we first outline a model without productivity shocks. In that setting we derive cross-sectional implications relating firm size and growth to pay and job-filling rates. Tractability and efficiency in the presence of shocks is established in Section 3, before we move to the quantitative analysis in Section 4. All proofs and some extensions are relegated to the Appendix.

[6] Moscarini and Postel-Vinay (2013) do allow for aggregate shocks, but their requirement of rank-preserving hiring prevents the study of firm entry and firm-specific shocks. To our knowledge, the only model that explicitly focuses on firm dynamics is Lentz and Mortensen (2010), which combines decreasing returns with on-the-job search, but again it has no idiosyncratic or aggregate shocks.

2 A Stationary Model of Firm Growth

2.1 The Environment

The model is set in discrete time and we consider a stationary environment. That is, there are neither idiosyncratic nor aggregate shocks in this section. The labor market within a given period operates in three stages. First, new firms enter and draw their productivity. Second, production and search activities take place. Third, vacancies and unemployed workers are matched, and a fraction of workers leave their firms. Afterwards some firms exit, and the next period starts. The following explains each part in turn.

The economy consists of a continuum of workers and firms. The mass of workers is normalized to one. Each worker is infinitely-lived, risk-neutral, and discounts future income with factor $\beta < 1$. A worker supplies one unit of labor per period when employed and receives income $b \geq 0$ when unemployed. Only unemployed workers search for employment, so there are no job-to-job transitions. On the other side of the labor market is an endogenous mass of firms. Firms are large relative to workers, in the sense that each active firm employs a continuum of workers. Firms are also risk neutral and have the same discount factor β.

An entrant firm pays setup cost $K > 0$ to start production. At this point it draws productivity x with probability $\pi_0(x)$ from the finite set $x \in X$. In each period, a firm produces output $xF(L)$ with $L \geq 0$ workers, where F is a twice differentiable, strictly increasing and strictly concave function satisfying $F'(0) = \infty$ and $F'(\infty) = 0$. Firms die with exogenous probability $\delta > 0$, in which case all workers are laid off into unemployment. Furthermore, each employed worker separates from the firm with exogenous probability $s \geq 0$. Thus, in this section, a firm's productivity stays constant throughout its life, and any worker's retention probability is exogenous at $\varphi \equiv (1 - \delta)(1 - s)$.

Search for new hires is a costly activity. A firm with workforce L and productivity x that posts V vacancies incurs recruitment costs $C(V, L, x)$. Apart from twice differentiability, we assume that a firm's output net of recruitment costs is strictly increasing in (L, x) and strictly concave in (V, L). In particular, this requires that C is strictly convex in V. Popular functional form are

$$C(V, L, x) = xF(L) - xF(L - hV) + k(V) \quad \text{or} \quad C(V, L, x) = \frac{c}{1+\gamma}\left(\frac{V}{L}\right)^{\gamma} V \ . \tag{1}$$

In the first specification, $k(V)$ captures some convex monetary costs (see e.g. Cooper et al. (2007)) and hV captures labor input in recruitment (see e.g. Shimer (2010)). Even in the absence of monetary costs and despite linearity of the labor input, this leads to convex costs because of decreasing returns in production.[7] The second, constant-returns specification, which is borrowed from Merz and Yashiv (2007), assumes that the average cost per vacancy increases in the vacancy rate (i.e. vacancies divided by employment) and it also allows larger firms to hire a given number of workers at lower costs.[8] In either setting, firms cannot instantaneously grow large simply by posting enough vacancies at constant marginal cost. For some proofs of cross-sectional relationships derived below (Proposition 1 and subsequent corollaries), we focus on cost functions such as those in (1)

[7]Clearly no more workers can be engaged in hiring than are present at the firm. To get the hiring process started for entrant firms, we need to assume that a new firm is endowed with initial labor input of the entrepreneur L_e so that the actual labor input is $\tilde{L} = L_e + L$. Recruitment activities are then constrained by $hV \leq L + L_e$, and Inada conditions on F ensure that this constraint never binds. A similar adjustment is needed for the second specification in (1) to avoid division by zero at entrant firms (see Section 4).

[8]To be precise, Merz and Yashiv (2007) specify and estimate convex adjustment costs (at the aggregate level) that depend on hires rather than vacancies. Relatedly, Blatter et al. (2012) estimate hiring costs on Swiss firm-level data and also find evidence for convexity. Costs that depend on hires better reflect training costs and could additionally be introduced into our framework. Costs that depend on the number of job openings rather capture recruiting costs and are more common in the search and matching literature.

which satisfy the following properties on cross-derivatives:

$$\text{(C)} \qquad C_{12} \leq 0 \ , \ C_{13} \geq 0 \ , \quad \text{and} \ \ C_{12}C_{13} + C_{11}[F' - C_{23}] \geq 0 \ .$$

In order to attract workers, a recruiting firm announces a flat flow wage income w to be paid to its new hires for the duration of the employment relation. The assumption that the firm offers the same wage to all its new hires turns out not to entail a restriction; see the discussion following equation (6) below. Further, because of risk neutrality, only the net present value that a firm promises to the worker matters. Flat wages are one way of delivering these promises.[9]

Unemployed workers direct their job search towards the most attractive offers: they can observe all wage offers and choose for which wage to search. At any wage, job seekers and vacancies are matched according to a matching function. In particular, a firm fills its vacancies with probability m only if it offers a wage that attracts $\lambda(m)$ unemployed job seekers per vacancy.[10] Standard assumptions on the matching function guarantee that this function is twice differentiable, strictly increasing and strictly convex in m, with $\lambda(0) = 0$, $\lambda'(0) \geq 1$ and $\lambda'(1) = \infty$.[11] It is increasing since firms achieve a higher matching probability only if more workers are searching for their vacancies. It is convex since it becomes increasingly difficult to improve matching prospects any further when more workers are

[9]This is a theory of the present value of offered wages. Constant wages can be viewed as the limiting case of risk-neutral firms and risk-averse workers, as risk aversion vanishes. But other payment patterns are conceivable; for further discussion about this issue, see Section 3.4.

[10]Note that we adopt the standard assumption in the literature on large firms in search models that each job has its own matching probability, i.e., applicants from one job cannot be hired at another job in the same firm, which arises, for example, if different jobs require different qualifications. Only few papers explore the idea that workers are literally identical and can be hired for another job than the one they applied for (see Burdett et al. (2001), Hawkins (2013) and Lester (2010)).

[11]Function λ is simply the inverse of the standard reduced-form matching function $\tilde{m}$: $[0, \infty) \to [0, 1)$ that maps the realized unemployed-vacancy ratio $\tilde{\lambda}$ into the hiring probability. Typically, $\tilde{m}$ is assumed to be strictly increasing and strictly concave, and $\tilde{m}(\tilde{\lambda}) \leq \min(1, \tilde{\lambda})$ guarantees that $\tilde{m}'(0) \leq 1$. Therefore, we can define $\lambda(m) = \tilde{m}^{-1}(m)$, and the properties in the text follow.

attracted to the job. The workers' matching probability is $m/\lambda(m)$, which is strictly decreasing.

To understand what wage $w(m)$ a firm has to offer in order to achieve matching probability m, note that in a stationary environment an unemployed worker who is seeking for a particular wage in one period is willing to search for that wage in every period. Let U denote the discounted present value from such job search. It is given by the following asset value equation:[12]

$$(1-\beta)U = b + \underbrace{\frac{m}{\lambda(m)}\beta(1-\delta)\frac{w(m)-(1-\beta)U}{1-\beta\varphi}}_{\equiv\rho} \ . \tag{2}$$

It states that the flow value of unemployment equals the current period unemployment income b together with an option value from searching, denoted by ρ. The search value is the probability of finding a job multiplied with the worker's discounted job surplus. Since workers have a choice where to search for a job, their flow value from unemployment must be equal in all markets that attract workers. Therefore, ρ is a global value that is common to all markets, which means that a firm has to offer the following wage to achieve matching rate $m > 0$:

$$w(m) \equiv b + \rho + \frac{1-\beta\varphi}{\beta(1-\delta)}\frac{\lambda(m)}{m}\rho \ . \tag{3}$$

This relation says that a firm can only recruit workers when its wage offer matches the workers' unemployment value $(1-\beta)U = b+\rho$ plus a premium which is needed to attract workers to jobs with filling rate m. This premium is increasing in m, which is a crucial insight. The relationship between job-filling rates and wage offers is standard in the competitive search literature.

[12]Bellman equations for employed and unemployed workers are $E = w + \beta[\varphi E + (1-\varphi)U]$ and $U = b+\beta[m\lambda(m)^{-1}(1-\delta)E+(1-m\lambda(m)^{-1}(1-\delta))U]$. Equation (2) follows by substituting the first into the second.

2.2 The Firms' Recruitment Policies

Consider the problem of a firm that takes the search value of unemployed workers and the associated relationship (3) as given. Later, the search value will be determined as an equilibrium object that depends on the number of firms and their wage offers.

Let $J^x(L, W)$ be the profit value of a firm that has productivity x, employs L workers and is committed to a wage bill of W. An entrant firm's profit value is then $J^x(0, 0)$. The firm's recruitment choice involves deciding the number of posted vacancies V as well as the job-filling probability m, which requires a wage offer of $w(m)$. Its recursive profit maximization problem is expressed as

$$\begin{aligned} J^x(L,W) &= \max_{(m,V)\in[0,1]\times\mathbb{R}_+} xF(L) - W - C(V,L,x) + \beta(1-\delta)J^x(L_+,W_+) \ , \\ & \quad \text{s.t. } L_+ = L(1-s) + mV \ , \ W_+ = W(1-s) + mw(m)V \ . \end{aligned} \tag{4}$$

The first line reflects the value of output net of wage and recruitment costs, plus the discounted value of continuation with an adjusted workforce and its associated wage commitment. The second line says that employment next period consists of the retained workers and the new hires. For the wages, since separations are random they reduce the wage bill proportionally, and new commitments are added for the new hires.

This problem can be simplified by noting that wages are commitments that have to be fulfilled as long as the worker does not separate, irrespective of future recruitment decisions. This has the implication that $J^x(L, W) = J^x(L, 0) - W/(1 - \beta\varphi)$, which eliminates the wage bill as a state variable, so that (4) readily yields

$$J^x(L,0) = \max_{(m,V)\in[0,1]\times\mathbb{R}_+} xF(L) - C(V,L,x) - D(m)V + \beta(1-\delta)J^x(L_+,0) \ ,$$
$$\text{s.t. } L_+ = L(1-s) + mV \ , \tag{5}$$

where $D(m) \equiv mw(m)\beta(1-\delta)/(1-\beta\varphi)$ captures the cost of increasing the matching probability by raising wage costs. Note that D is increasing and strictly convex since the matching function $\lambda(m)$ has these properties. Problem (5) makes it readily apparent that a firm has two channels to hire workers in a given period. It can increase the number of vacancies and associated costs C, or it can increase the filling rate of each job and associated costs D. The optimality conditions for the control variables in (5) are derived rigorously in the Appendix, but we provide some intuition here for the main trade-offs. The optimal choices for the number of vacancies and their matching probability are governed by one intratemporal and one intertemporal optimality condition.

Regarding the intratemporal optimality condition, consider a firm that aims to hire H workers in this period. It faces the problem of choosing the number of vacancies and the job-filling probability to minimize costs $C(V,.) + D(m)V$ subject to $H = mV$. The first-order condition for this problem is

$$C_1(V,L,x) = D'(m)m - D(m) = \rho[m\lambda'(m) - \lambda(m)] \ . \tag{6}$$

This links the marginal recruitment costs to the marginal increase in wage costs associated with increasing the job-filling probability.

Relationship (6) offers a number of insights. It defines the optimal policy for vacancy postings $V = V^x(m,L)$ as a function of the job-filling rate and firm size. Because of convex recruitment costs, this policy function is increasing in m; thus, vacancy postings and job-filling rates are complementary tools in the firm's recruitment strategy. This captures the basic stylized fact highlighted by Davis et al. (2013) that firms use both more vacancies as well as higher job-filling rates to achieve faster growth.[13] In contrast, under constant marginal recruitment costs ($C_1(V, L, x) = c$), as assumed in much of the literature, the job-filling rate would be constant and independent of firm characteristics, while all employment adjustment is instantaneous and is achieved through the number of vacancies. Finally, note that equation (6) balances the wage costs for new hires against recruitment costs at a unique point, which shows why a firm would not want to offer different wages at a given point in time even if this were permissable.

The firm also decides how to structure hiring over time. This is governed by an intertemporal optimality condition which reads

$$xF'(L_+) - C_2(V_+, L_+, x) - b - \rho = \frac{\rho}{\beta(1-\delta)} \Big[\lambda'(m) - \beta\varphi\lambda'(m_+) \Big] \, . \qquad (7)$$

Here L_+, V_+, and m_+ are employment, vacancy postings and the job-filling rate in the next period. The left-hand side of (7) gives the marginal benefit of a higher workforce in the next period. If this is high, then the firm rather hires more now than to wait and hire next period, as expressed by the right-hand side which is increasing in the current job-filling rate m and decreasing in m_+. In particular, a more productive firm wants to achieve fast growth by offering a more attractive contract now rather than later, thus raising the current job-filling rate. Equation (7) implicitly defines the optimal job-filling policy $m^x(L)$. Starting from $L = 0$, this determines the firm's growth path through $L_+ = L(1-s) + m^x(L)V$, where $V = V^x(m^x(L), L)$ comes from the static optimality condition (6).

An illustration how a firm grows over time is provided in Figure 1 which shows the phase diagram in (L, m) space for the firm's problem with recruitment costs $C(V, L, x) = xF(L) - xF(L - hV) + cV$ for which the optimality conditions become

[13] The first equation in (6) suggests that this argument holds in a broader class of models in which firms can influence job-filling rates. In our model, job-filling rates are increased via higher wage offers which reflects the allocative role of wages in the labor market.

especially tractable.[14] Initially the firm is small and the optimal job-filling rate exceeds the long-run rate m^*. This rate is the firm's policy after it converges to its long-run optimal size $L^* > 0$ where it only conducts replacement hiring. The downward-sloping saddle path depicts the firm's policy function $m^x(L)$ and describes the adjustment process to the long-run optimal size, along which the firm spreads recruitment costs over time. This is in contrast to a model with linear recruitment costs in which firms would jump directly to (L^*, m^*). In terms of comparative statics, this example also shows that the stationary firm size and the job-filling rates along the transition depend positively on x: a more productive firm grows larger and offers higher lifetime wages on its transition to the long-run employment level. The following proposition and its corollaries provide broader

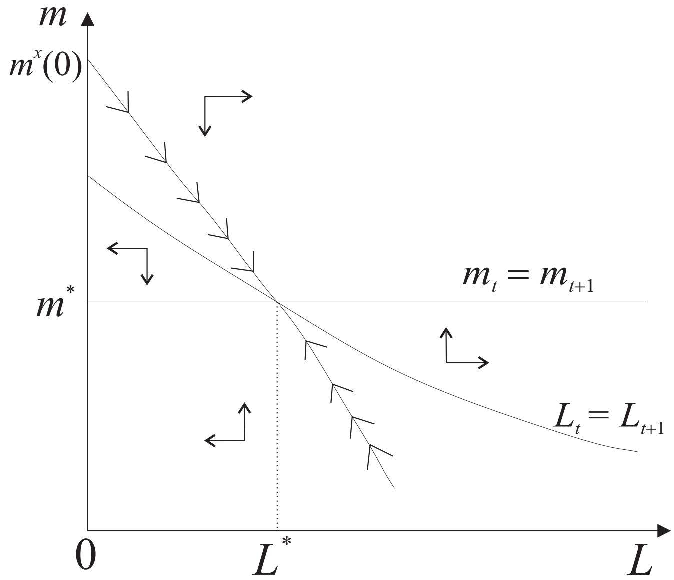

Figure 1: The firm's optimal recruitment policy follows the declining saddle path.

[14]In Lemma 3 of the Appendix we show that equations (6) and (7) simplify to only one equation linking m_t and m_{t+1} which is independent of L_t. This equation has a unique long-run job-filling probability $m^* > 0$ if h is low enough, and m_t converges to m^* from any initial value $m_0 > 0$. Employment adjusts according to $L_{t+1} = L_t(1-s) + m_t V^x(m_t, L_t)$. Using (6), it is easy to see that the curve $L_{t+1} = L_t$ is downward-sloping in (L, m) space, so that the saddle path lies above this curve when $L_t < L^*$.

comparative statics results. The job-filling rate is linked via (3) to the wage offer, so that the findings carry over to the net present value of wages to new hires.[15]

Proposition 1: *Consider recruitment cost functions satisfying property (C). The firm's value function $J^x(L, W)$ is strictly increasing and strictly concave in its workforce L, strictly increasing in productivity x, strictly supermodular in (x, L) and decreasing in the worker's search value ρ. The job-filling rate $m^x(L)$ is strictly increasing in productivity x and strictly decreasing in the workforce L. Posted vacancies $V^x(m, L)$ are increasing in L and strictly increasing in the desired job-filling rate m.*

Since these results hold for any search value ρ, they also apply when this value is determined in general equilibrium. These results imply relationships between size, productivity, pay, and hiring:

Corollary 1: *Consider recruitment cost functions satisfying property (C). Conditional on size, more productive firms pay higher lifetime wages and have a higher job-filling rate. Conditional on productivity, younger/smaller firms pay higher lifetime wages and have a higher job-filling rate.*

In the Appendix we also prove the following connection to firm growth rates.

Corollary 2: *If recruitment costs are given by either specification in (1) with parameter h sufficiently small, more productive firms have a higher growth rate, conditional on size; and larger/older firms have a lower growth rate, conditional on productivity.*

While it already follows from (6) that vacancy postings and job-filling rates are positively related, the two corollaries link these policies to the firm's growth rate. They point out that job-filling rates and firm growth rates are positively correlated, depending positively on x and negatively on L. This cross-sectional

[15] These characterization results depend crucially on the supermodularity of the value function, which renders this proof non-trivial. While standard techniques (Amir (1996)) can be applied when the cost function is independent of firm size and productivity, this is not true in general, as we discuss in the Appendix.

relationship has been highlighted recently by Davis et al. (2013), and we further explore in Section 4 how well our model captures this quantitatively. Furthermore, since higher job-filling rates are directly associated with higher earnings for new hires, the two corollaries also imply that faster-growing firms offer higher lifetime wages. Belzil (2000) documents such patterns after controlling for size and worker characteristics; he shows that wages, particularly those of new hires, are positively related to a firm's job creation. Our findings that younger firms grow faster (conditional on survival) and pay higher wages (to workers with the same characteristics) are consistent with the evidence (see Haltiwanger et al. (2013), Brown and Medoff (2003) and Schmieder (2013)). Moreover, a positive wage-size relation emerges in our model if the dispersion in productivity is large enough.[16]

2.3 Firm Entry, General Equilibrium, and Efficiency

Free entry of firms implies that no entrant makes a positive profit, that is,

$$\sum_{x \in X} \pi(x) J^x(0,0) \leq K \ , \tag{8}$$

with equality if entry is positive. This condition implicitly pins down the workers' job surplus ρ and therefore the relationship between wages and job-filling rates. In a stationary equilibrium, a constant mass of N_0 firms enter the market in every period, so that there are $N_a = N_0(1-\delta)^a$ firms of age a in any period. Let $(L_a^x, m_a^x, V_a^x, w_a^x)_{a \geq 0}$ be the employment/recruitment path for a firm with productivity x. Then, a firm of age a has L_a^x employed workers, and $\lambda(m_a^x)V_a^x$ unemployed workers are searching for jobs with offered wage w_a^x. Therefore, the mass of entrant firms N_0 is uniquely pinned down from aggregate resource feasibility:

[16]We note that enough productivity dispersion is also required in models with intra-firm bargaining, and even more so because wages of *all* workers decline in a growing firm.

$$1 = \sum_{a \geq 0} N_0 (1-\delta)^a \sum_{x \in X} \pi(x) [L_a^x + \lambda(m_a^x) V_a^x] \ . \tag{9}$$

This equation says that the unit mass of workers is either employed or unemployed. We now define a stationary equilibrium.

Definition: *A stationary competitive search equilibrium is a list* $\left(\rho, N_0, (L_a^x, m_a^x, V_a^x, w_a^x)_{x \in X, a \geq 0}\right)$ *with the following properties. Unemployed workers' job search strategies maximize utility: (3) holds for all* (w_a^x, m_a^x)*. Firms' recruitment policies are optimal:* $(L_a^x, m_a^x, V_a^x)_{a \geq 0}$ *solve (5) for all* $x \in X$*. There is free entry of firms: (8) and* $N_0 \geq 0$ *hold with complementary slackness. Aggregate resource feasibility (9) holds.*

Since the firms' behavior has already been characterized, it remains to explore equilibrium existence and uniqueness.

Proposition 2: *A stationary competitive search equilibrium exists and is unique. There is strictly positive firm entry provided that* K *is sufficiently small.*

The previous section already outlined that this model generates sensible relationships between productivity, size, growth, and job-filling rates. It is relevant to understand whether these patterns are actually socially efficient, especially since existing models with intra-firm bargaining always entail inefficiencies, as discussed in the introduction. To this end, consider a social planner who decides at each point in time about firm entry, vacancy postings and job-filling rates for all firms. The planner takes as given the numbers of firms that entered in some earlier period, as well as the employment stocks of all these firms. Formally, the planner's state vector is $\sigma = (N_a, L_a^x)_{a \geq 1, x \in X}$ where N_a is the mass of firms of age $a \geq 1$, and L_a^x is employment of a firm with productivity x and age a. The planner maximizes the present value of output net of opportunity costs of employment and net of the costs of entry and recruitment, subject to the economy's resource constraint. With $\sigma_+ = (N_{a,+}, L_{a,+}^x)_{a \geq 1, x \in X}$ denoting the state vector in the next period, the recursive formulation of the social planning problem is

$$
\begin{aligned}
S(\sigma) \;&=\; \max_{N_0,(V_a^x,m_a^x)_{a\geq 0}} \left\{ \sum_{a\geq 0} N_a \sum_{x\in X} \pi(x) \Big[xF(L_a^x) - bL_a^x - C(V_a^x, L_a^x, x) \Big] \right\} \\
&\qquad\qquad -KN_0 + \beta S(\sigma_+) \qquad\qquad (10) \\
&\text{s.t. } L_0^x = 0,\; L_{a+1,+}^x = (1-s)L_a^x + m_a^x V_a^x,\; a \geq 0,\; x\in X \;, \\
&N_{a+1,+} = (1-\delta)N_a \;,\; a\geq 0 \;, \\
&\sum_{a\geq 0} N_a \sum_{x\in X} \pi(x) \Big(L_a^x + \lambda(m_a^x) V_a^x \Big) \leq 1 \;.
\end{aligned}
$$

We say that a solution to problem (10) is *socially optimal.*

Proposition 3: *The stationary competitive search equilibrium is socially optimal.*

The efficiency of equilibrium can be linked to a variant of the well-known Hosios (1990) condition.[17] It says that efficient job creation requires that the firm's surplus share for the marginal vacancy is related to the elasticity of the matching function. Write the workers' search value $\rho = \frac{m}{\lambda(m)} S^w$ as the product between the match probability and the worker's job surplus S^w. Then, equation (6) can be rewritten as

$$C_1(V, L, x) = \frac{1-\varepsilon_{m,\lambda}}{\varepsilon_{m,\lambda}} m S^w \;,$$

where $\varepsilon_{m,\lambda} = \frac{\lambda(m)}{\lambda'(m)m} \in [0,1]$ is the matching-function elasticity.

3 Productivity Shocks and Firm Dynamics

In this section we show that our framework can be extended to include much richer dynamics, both at the firm level and in the aggregate, while retaining tractability and efficiency. We include both firm-specific and aggregate productivity shocks to explore not only two margins of job creation (firm entry and firm growth),

[17]See also the Hosios condition in a large-firm model with intra-firm bargaining in Hawkins (2010).

but also the two margins of job destruction (endogeneous firm exit and firm contraction). This extension allows us to study in the next section to which extent the model can quantitatively account for the micro-level heterogeneity in the firms' recruitment behavior and how it performs over the business cycle.

Assume now that output of a firm with L workers is $xzF(L)$ where $x \in X$ is idiosyncratic productivity and $z \in Z$ is aggregate productivity. Both x and z follow Markov processes on finite state spaces X and Z with respective transition probabilities $\pi(x_+|x)$ and $\psi(z_+|z)$. An entrant firm pays fixed cost $K(z)$, possibly dependent on the aggregate state, and draws an initial productivity level $x_0 \in X$ with probability $\pi_0(x_0)$. For a firm of age $a \geq 0$, let $x^a = (x_0, \ldots, x_a) \in X^{a+1}$ denote the history of idiosyncratic productivity, and let $z^t = (z_0, \ldots, z_t)$ be the history of aggregate states at time t with corresponding probability $\psi(z^t)$.

We assume that an active firm incurs a fixed operating cost $f \geq 0$ per period, which is required to obtain a non-trivial exit margin. In this section we are as agnostic as possible about the recruitment cost function; we only assume that $C(V, L, xz)$ is strictly increasing and convex in posted vacancies. Firms exit with exogenous probability $\delta_0 \geq 0$ which is a lower bound for the actual exit rates $\delta \geq \delta_0$. Similarly, workers quit a job with exogenous rate $s_0 \geq 0$ which provides a lower bound for the actual separation rates $s \geq s_0$.

The timing within each period is analogous to the previous section. *First*, aggregate and idiosyncratic productivities are revealed and new firms enter. *Second*, firms produce and they decide about hires, layoffs, and possibly about exiting at the end of the period. And *third*, workers and firms are matched. We start to describe and characterize the planning problem before we show its equivalence to a competitive search equilibrium.

3.1 The Planning Problem

The planner decides at each point in time about firm entry and exit, layoffs and hires (i.e. vacancy postings and matching probabilities) for all firm types, knowing that matching probability m requires $\lambda(m)$ unemployed workers per vacancy. In a given aggregate history z^t, we denote by $N(x^a, z^t)$ the mass of firms of age a with idiosyncratic history x^a. $L(x^a, z^t)$ is the employment stock of any of these firms. At every history node z^t and for every firm type x^a, the planner decides an exit probability $\delta(x^a, z^t) \geq \delta_0$, a separation rate $s(x^a, z^t) \geq s_0$, vacancy postings $V(x^a, z^t) \geq 0$, and a matching probability $m(x^a, z^t)$.[18] The numbers of firm types change between periods t and $t+1$ according to

$$N(x^{a+1}, z^{t+1}) = [1 - \delta(x^a, z^t)]\pi(x_{a+1}|x_a)\psi(z_{t+1}|z_t)N(x^a, z^t) \ , \tag{11}$$

and the workforce at any of these firms adjusts to

$$L(x^{a+1}, z^{t+1}) = [1 - s(x^a, z^t)]L(x^a, z^t) + m(x^a, z^t)V(x^a, z^t) \ . \tag{12}$$

At time $t = 0$, the planner takes as given the numbers of firms that entered the economy in some earlier period, as well as the employment stock of each of these firms. Hence, the state vector at date 0 is summarized by the initial firm distribution $(N(x^a, z^0), L(x^a, z^0))_{a\geq 1, x^a \in X^{a+1}}$. In a given history z^t, the planner also decides the mass of new entrants $N_0(z^t) \geq 0$, so that

$$N(x_0, z^t) = \pi_0(x_0)N_0(z^t) \text{ and } L(x_0, z^t) = 0 \ . \tag{13}$$

[18] To save on notation, we do not allow the planner to discriminate between workers with different firm tenure. Given that there is no learning-on-the-job, there is clearly no reason for the planner to do so. Nonetheless, the competitive search equilibrium considered in Section 3.4 allows firms to treat workers in different cohorts differently, which is necessary because firms offer contracts sequentially and are committed to these contracts. See the proof of Proposition 5 for further elaboration of this issue.

The sequential planning problem is

$$\max_{\delta,s,V,m,N_0} \sum_{t\geq 0,z^t} \beta^t \psi(z^t) \Bigg\{ -K(z_t)N_0(z^t) + \sum_{a\geq 0,x^a} N(x^a,z^t)\Big[x_a z_t F(L(x^a,z^t)) \\ -bL(x^a,z^t) - f - C(V(x^a,z^t), L(x^a,z^t), x_a z_t)\Big]\Bigg\} \ , \tag{14}$$

subject to the dynamic equations for N and L, namely (11), (12) and (13), and subject to the resource constraints, for all $z^t \in Z^{t+1}$,

$$\sum_{a\geq 0,x^a} N(x^a,z^t)\Big[L(x^a,z^t) + \lambda(m(x^a,z^t))V(x^a,z^t)\Big] \leq 1 \ . \tag{15}$$

This constraint says that the labor force (employment plus unemployment) cannot exceed the given unit mass of workers. We summarize a solution to the planning problem by a vector $(\mathbf{N}, \mathbf{L}, \mathbf{V}, \mathbf{m}, \mathbf{s}, \boldsymbol{\delta})$, with $\mathbf{N} = (N(x^a,z^t))_{a,t\geq 0}$ and similar notation for the other variables.

3.2 Characterization of the Planning Solutions

We show that there is a convenient characterization of a planning solution which says that hiring, layoff and exit decisions follow a recursive equation at the level of the individual firm. Specifically, for any existing firm, the planner maximizes the social value of the firm, taking into account the social value of each worker tied to the firm. This social worker value is given by the multiplier on the resource constraint (15) which we denote by $\mu(z^t)$ and which generally depends on the initial firm distribution and on the full state history z^t.

A particularly powerful characterization can be obtained under the provision that firm entry is positive in all states of the planning solution. When this is the case, the social value of a worker (and thus firm-level value and policy functions) depend only on the current aggregate state but are independent of the state history and of the firm distribution.

To gain intuition for the independence from the distribution of existing firms, envision any period in which the planner can assign unemployed workers either to existing firms or to new firms. If there are many existing firms, there are fewer workers left to be assigned to new firms. Nevertheless, the social value of any worker that is assigned to a new firm does not change: Each new firm has an optimal hiring policy, and if less workers are assigned to new firms, then proportionally less new firms will be created, leaving the marginal value of each worker unchanged. Therefore, efficient hiring by existing firms requires their marginal social benefit of hiring to be equal to the social benefit at the new firms which depends on the aggregate state alone.

To see the independence of value functions from the firm distribution formally, suppose there are n aggregate states $z \in Z = \{z_1, \dots, z_n\}$, and let μ_i be the social value of a worker in state z_i. Write $M = (\mu_1, \dots, \mu_n)$ for the vector of social values. Let $G(L, x, i; M)$ be the social value of a firm with employment stock L, idiosyncratic productivity x and aggregate productivity z_i, satisfying the Bellman equations

$$\begin{aligned} G(L,x,i;M) = \max_{\delta,s,V,m} \quad & xz_iF(L) - bL - f - \mu_i[L + \lambda(m)V] - C(V, L, xz_i) \\ & +\beta(1-\delta)E_{x,i}G(L_+, x_+, i_+; M) \ , \end{aligned} \tag{16}$$

where maximization is subject to $L_+ = (1-s)L + mV$, $\delta \in [\delta_0, 1]$, $s \in [s_0, 1]$, $m \in [0,1]$ and $V \geq 0$. The interpretation of this equation is rather straightforward. A firm's social value encompasses flow output net of the opportunity cost of employment, net of fixed costs and recruitment costs, and net of the social cost of workers tied to the firm in this period; these workers include the current workforce L and also $\lambda(m)V$ unemployed workers who are assigned to this firm.

Positive entry in all aggregate states requires that the expected social value of a new firm is equal to the entry cost,

$$\sum_{x \in X} \pi_0(x) G(0, x, i; M) = K(z_i) \ . \tag{17}$$

This characterization of planning solutions by (G, M) is particularly helpful for numerical applications. Despite considerable firm heterogeneity, the model can be solved by a recursive problem on a low-dimensional state space (16) and the (simultaneous) solution of a finite-dimensional fixed-point problem (17). Importantly, the distribution of firms is irrelevant for this computation. After the corresponding policy functions have been calculated, firm entry follows as a residual of the economy's resource constraint and *does depend* on the distribution of existing firms: in every period with aggregate state i, each existing firm with productivity x and size L attracts $V(L, x, i)\lambda(m(L, x, i))$ job seekers according to the policy functions, while a number $N_0(z^t)$ of new firms enter to absorb the remaining job seekers. Since job-finding prospects differ between firms, the aggregate job-finding rate therefore also depends on the firm-size distribution, as does the evolution of aggregate employment, output and job flows. As we see in the next section, these aggregate variables in fact adjust with delay to aggregate shocks. Because of the dependence of N_0 on the distribution of employment among existing firms, it cannot generally be guaranteed that the planning solution has positive entry in all state histories. Therefore, this property can only be checked ex-post in simulations of the model. Analytically, we prove that any solution of (16)–(17) which gives rise to positive entry in all state histories describes indeed a solution to the planner's problem. We also find that a unique solution of these equations exists for small aggregate shocks:

Proposition 4:

(a) Suppose that a solution of (16) and (17) exists with associated allocation $\mathbf{A} = (\mathbf{N}, \mathbf{L}, \mathbf{V}, \mathbf{m}, \mathbf{s}, \boldsymbol{\delta})$ *satisfying* $N(z^t) > 0$ *for all* z^t*. Then* $\mathbf{A}$ *is a solution of the sequential planning problem (14).*

(b) If $K(z)$*,* f*, and* b *are sufficiently small and if* $z_1 = \ldots = z_n = \overline{z}$*, equations (16) and (17) have a unique solution* (G, M)*. Moreover, if the transition matrix* $\psi(z_j|z_i)$ *is strictly diagonally dominant and if* $|z_i - \overline{z}|$ *is sufficiently small for all* i*, equations (16) and (17) have a unique solution.*

3.3 Recruitment and Layoff Strategies

The reduction of the planning problem to (16) permits a straightforward characterization of the optimal layoff and hiring strategies. For a growing firm, it follows from the first-order conditions for m and V, similar to equation (6), that

$$C_1(V, L, xz_i) = \mu_i[m\lambda'(m) - \lambda(m)] \ . \tag{18}$$

As in the previous section, this equation implies an increasing relation between matching probabilities and the number of posted vacancies at the firm. With higher m, the planner is willing to post more vacancies at higher marginal recruiting cost. Denote the solution to equation (18) by $V = V(m, L, x, i)$, which is positive for $m > \underline{m}(L, x, i)$. The planner's optimal choice of m for firm (L, x) in aggregate state i satisfies[19]

$$\beta(1 - \delta_0) E_{x,i} \frac{dG}{dL}(L_+, x_+, i_+; M) = \mu_i \lambda'(m) \ ,$$

with $L_+ = L(1 - s_0) + mV(m, L, x, i)$. Therefore, the firm hires if and only if

[19] Note that $\delta = \delta_0$ and $s = s_0$ if the firm hires workers.

$$\beta(1-\delta_0)E_{x,i}\frac{dG}{dL}(L(1-s_0),x_+,i_+;M) > \mu_i\lambda'(\underline{m}(L,x,i)) \ . \tag{19}$$

Conversely, the planner wants the firm to lay off workers if

$$E_{x,i}\frac{dG}{dL}(L(1-s_0),x_+,i_+;M) < 0 \ . \tag{20}$$

The two conditions (19) and (20) show how the firm's strategy depends on its characteristics. Small and productive firms recruit workers and grow, whereas large and unproductive firms dismiss workers and shrink. There is also an open set of characteristics where firms do not adjust their workforce (cf. Bentolila and Bertola (1990) and Elsby and Michaels (2013)).

3.4 Decentralization

As in Section 2, a competitive search equilibrium gives rise to the same allocation as the planning solution. Consider firms that offer workers a sequence of state-contingent wages, to be paid for the duration of the match. They also commit to cohort-specific and state-contingent retention probabilities. Contracts are contingent on the idiosyncratic productivity history of the firm at age k, x^k, and on the aggregate state history z^t at time t. Formally, a contract offered by a firm of age a at time T takes the form

$$\mathcal{C}_a = \Big(w_a(x^k,z^t),\varphi_a(x^k,z^t)\Big)_{k>a,t=T+k-a} \ ,$$

where $w_a(x^k,z^t)$ is the wage paid to the worker in history (x^k,z^t), conditional on the worker being still employed by the firm in that instant. $\varphi_a(x^k,z^t)$, for $k>a$, is the probability of retaining the worker at the end of the period, so $1-\varphi_a(x^k,z^t)$ is the separation probability.

In Appendix B, we describe the workers' and the firms' search problems and we define a competitive search equilibrium, analogously to the stationary model. We also prove the following welfare theorem, extending Proposition 3.

Proposition 5: *A competitive search equilibrium is socially optimal.*

It is not hard to see that a wage commitment is sufficient for a firm to implement its desired policy, even if it cannot commit to retention rates. Given risk neutrality, the firm can set the wages following any future history exactly equal to the reservation wage (i.e. the flow value of unemployment) which is the sum of unemployment income and the worker's shadow value, $b + \mu(z^t)$. It can achieve any initial transfer to attract workers through a hiring bonus. In this decentralization, the costs of an existing worker are always equal to his social value in the alternative: unemployment and search for another job. Since the flow surplus for any retained worker equals his shadow value, the firm's problem in this case coincides with the planner's problem, so that firing and exiting will be exactly up to the socially optimal level even though the firm does not commit to retention rates. Workers do not have any incentive to quit the job unilaterally, either, because they are exactly compensated for their social shadow value from searching. If the workers also cannot commit to stay, this is the unique wage policy that overcomes the commitment problem on both sides of the market and implements the socially efficient outcome. Alternatively, even a slight degree of risk aversion on the workers' side would give rise to flat wage profiles to offer insurance. This clarifies that the current model determines surplus sharing only, whereas the time path of payments depends on additional details, like the ability to commit to specific actions (see Schaal (2010) for a related point).

4 Quantitative Exploration

The previous sections outlined that this model can capture important features at the micro level (e.g. varying job-filling rates) and it is tractable for study-ing business cycle dynamics with potentially sluggish adjustment of aggregate variables. In this section we calibrate our model to the U.S. labor market in order to investigate the how well it is able to account for the main features in the data quantitatively. We first explore the model's cross-sectional properties, showing among other results how it generates differential job-filling rates as in Davis et al. (2013). We then show that the same parameterizations give rise to aggregate sluggishness and other business cycle features.

We briefly sketch the model calibration, referring to Appendix C for details and for the parameter values. We calibrate the model at weekly frequency and choose the firm-specific productivities to match firm and employment shares in the three size classes $1-49$, $50-499$, and ≥ 500.[20] For the recruitment technology, we choose the employment-scaled form[21] $c(V) = \frac{c}{1+\gamma}(\frac{V}{L})^{\gamma}V$. In our benchmark calibration we take a cubic function ($\gamma = 2$). While this relates to Merz and Yashiv (2007) who estimate a similar cubic hiring technology,[22] we take an agnostic view about this parameter value. Therefore, we compare the benchmark results with those obtained with a nearly linear recruitment technology ($\gamma = 0.1$) and with a much higher elasticity ($\gamma = 8$). In all versions, the scale parameter c is recalibrated to match our target for the weekly job-filling rate. Unemployment income b is set at the same value (relative to earnings) as in Hagedorn and Manovskii (2008) to ensure that reasonably small productivity shocks have quantitatively significant labor market responses. Robustness regarding this parameter, as well as regarding the returns-to-scale parameter, is explored in Appendix D.

We first simulate the model for a stationary cross-section of firms. Besides match-

[20]We calibrate the model to match the size distribution of firms (rather than establishments) in the Business Employment Dynamics program of the Bureau of Labor Statistics (BLS). We note that those results relating to establishment-level statistics (e.g. Figure 2) are robust when we restrict the model sample to the first two size classes (less than 500 employees) which largely represent one-establishment businesses.

[21]To avoid division by zero at entrant firms, we assume that actual labor input $\tilde{L} = 1 + L$ is the sum of the labor inputs of the (single) owner and of the employed workers.

[22]As mentioned before (footnote 8), their estimation results are not applicable to our model.

ing the calibration targets, our model generates negative relationships between firm size and quarterly job creation and job destruction rates in different size classes (see Table 1). This is despite the fact that we do not calibrate idiosyncratic productivity processes separately for each size class. A similar negative relationship between firm size and job flows obtains at entrant and exiting firms. Our model also performs reasonably well in matching the cross-sectional dispersion of quarterly employment growth rates across firms (see Table 2).

Table 1: Firm size, employment shares and quarterly job flows

	Data			Model ($\gamma = 2$)		
Size class	1–49	50–499	≥ 500	1–49	50–499	≥ 500
Firm shares	94.9	4.6	0.4	95.1	4.5	0.4
Employment shares	29.9	25.9	44.1	31.5	26.0	42.5
Job creation	10.6	5.7	3.1	9.2	6.1	5.4
Job destruction	10.4	5.4	2.9	9.3	6.1	5.4
Job creation (openings)	3.0	0.27	0.02	1.9	0.15	0.04
Job destruction (closings)	2.9	0.32	0.04	2.9	0.31	0.1

Notes: The first two rows report firm and employment shares in the three size classes 1-49, 50-499, and ≥ 500 (calibrated). The last four rows are quarterly job creation and destruction rates in the three size classes, expressed as shares of employment. Data statistics are from the Business Employment Dynamics (1992-2011) of the BLS.

One dimension of particular interest is the relationship between employment growth, the vacancy rate and the vacancy yield, which are positively related for growing firms in the Job Openings and Labor Turnover Survey (JOLTS), see Davis et al. (2013). This indicates that the matching rate varies across firms, a feature that is not present in most standard models. To see whether our model can trace this relationship quantitatively, we calculate monthly model statistics for hires, vacancies, layoffs and employment growth rates.[23] Figure 2 shows the

[23] When L_{t-1} and V_{t-1} denotes employment and vacancies at the end of month $t-1$ and H_t are hires during month t, the hires rate is $h_t = H_t/L_{t-1}$, the vacancy rate is $v^r_{t-1} = V_{t-1}/L_{t-1}$ and the vacancy yield is $v^y_t = H_t/V_{t-1}$, so that $h_t = v^r_{t-1}v^y_t$. We use this definition, which is

Table 2: Distribution of employment growth

Growth rate interval	Data	Model ($\gamma = 2$)
-2 (exit)	0.7	0.5
$(-2, -0.2]$	7.5	11.7
$(-0.2, -0.05]$	16.5	12.7
$(-0.05, -0.02]$	9.6	9.1
$(-0.02, 0.02)$	30.9	32.9
$[0.02, 0.05)$	9.9	6.3
$[0.05, 0.2)$	16.7	14.6
$[0.2, 2)$	7.5	11.9
2 (entry)	0.7	0.4

Notes: The table reports employment shares for intervals of quarterly employment growth rates. The empirical distribution is taken from Table 2 of Davis et al. (2010). Model statistics are calculated for the benchmark calibration from a cross-section of $4.7 \cdot 10^6$ firms.

cross-sectional relationships from the data and for the three parameterizations of our model.[24] In the data, firms grow larger both by posting more vacancies and by filling vacancies faster, with the vacancy yield accounting for most of the variation. The benchmark calibration with a cubic recruitment cost function can account for around two thirds of the observed variation in vacancy yields (see the blue (solid) curve in the upper right graph). Employers that expand more rapidly offer more attractive contracts and fill these vacancies faster. There can be many different reasons why vacancy yields are higher in faster-growing firms. For example, strongly expanding firms may search more intensively or they may use alternative recruitment channels. Time aggregation can also account for part of this variation; see Davis et al. (2013) for a discussion. Our benchmark results suggest that competitive search can be one important, but perhaps not the only mechanism responsible for the observed heterogeneity in vacancy yields and

slightly different from Davis et al. (2013), for the model and data statistics. We are grateful to Jason Faberman for providing these data.

[24]To smooth the relationships, all figures in the graphs are calculated as five-bin centered moving averages, as in Davis et al. (2013).

vacancy rates.

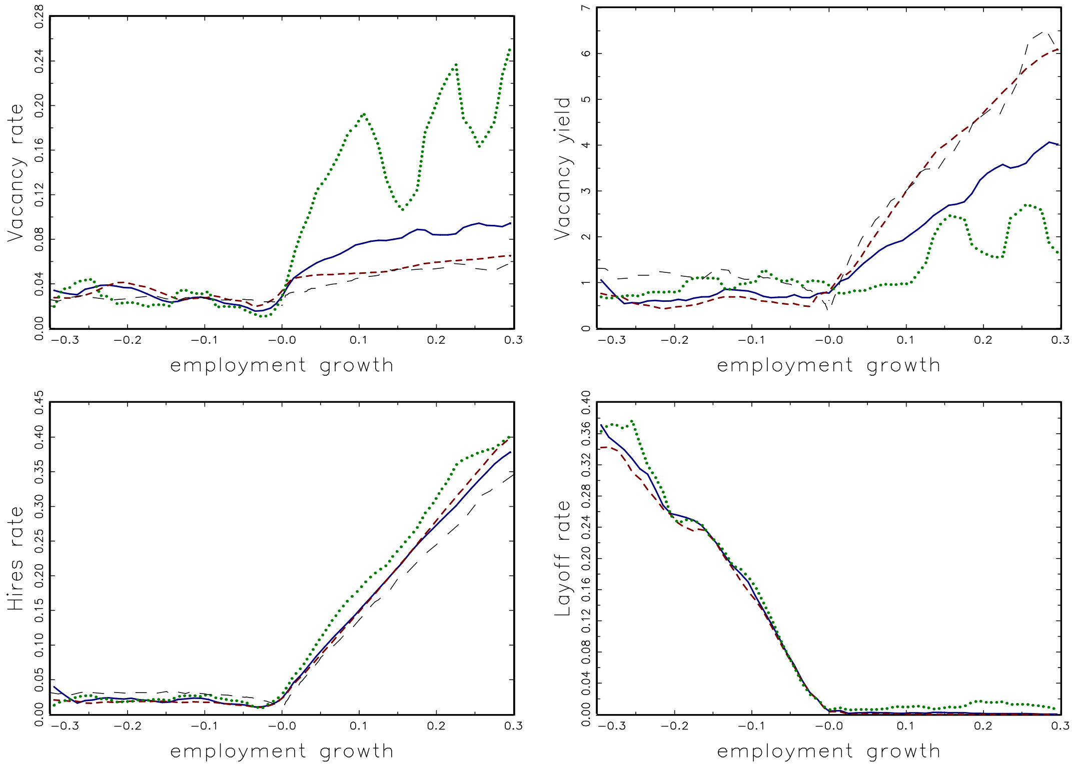

Figure 2: Cross-sectional relationships between monthly employment growth and the vacancy rate, the vacancy yield, the hires rate and the layoff rate. The dashed curves (in the first three graphs) are from the data used in Davis et al. (2013), the blue (solid) curves are for the model with cubic recruitment costs ($\gamma = 2$), the green (dotted) curves are for $\gamma = 0.1$ and the red (closely dashed) curves are for $\gamma = 8$. Model statistics are calculated from a cross-section of $4.7 \cdot 10^6$ firms.

Figure 2 further shows the results for the nearly linear recruitment technology ($\gamma = 0.1$) and for the one with higher curvature ($\gamma = 8$). With linear vacancy costs, *weekly* vacancy yields m are constant and hence do not vary with employment growth. Variations in the *monthly* vacancy yield are solely explained by time aggregation. The green (dotted) curve in the upper right graph of Figure 2 shows that the vacancy yield is indeed nearly flat for employment growth below 10 percent. Time aggregation (i.e., firms post and fill unrecorded vacancies during the month) accounts for the variation in vacancy yields beyond that point. On the other hand, as indicated by the red (closely dashed) curves in the figure, our model can principally account for the full variation in vacancy yields and vacancy rates if the curvature of the recruitment technology is sufficiently large. On a related note, Davis et al. (2013) show that vacancy yields (and vacancy rates) vary substantially by industry and employer size groups. While we have not introduced industry-specific parameters into our model, we can study the effect of size and find that smaller employers indeed have higher vacancy yields, albeit the variation is smaller than in the data. Specifically, in our benchmark calibration the vacancy yield at firms with less than 50 workers exceeds the one at firms with more than 500 workers by 10 percent, while in the data the difference is almost a factor of two.[25]

The bottom graphs in Figure 2 show that our model largely accounts for the relationships between employment growth, hires rates and layoff rates, both for growing and for shrinking firms, and regardless of the curvature parameter in the recruitment technology.[26]

To explore the impact of aggregate shocks, we first consider the model response to a permanent increase in the aggregate productivity parameter by one percent. In response to the shock, we let entry costs increase by the same factor.[27] The

[25]We expect that more flexible forms of the recruitment technology should give larger variation by employer size: for instance, if C had decreasing returns in (V, L), vacancy postings in larger firms would be less costly so that these firms prefer to recruit less intensively, reducing job-filling rates further.

[26]For the empirical relationship between employment growth and layoffs, see Davis et al. (2010) who find that layoffs dominate quits for large employment contractions. In our model, the quit rate is exogenous at s_0 so that variations in layoffs necessarily capture all variations in separations.

[27]Without the proportional increase in entry costs, firm entry would exhibit an implausible spike at the time of the shock. There are many reasons why entry costs vary with the business cycle, e.g. procyclical rental rates, capital prices, or outside opportunities of entrepreneurs. Regarding the latter, endogenous entrepreneurship could be easily introduced in our framework when unemployed workers have the option to either search for jobs or to start a business. We expect that efficiency and tractability would be preserved.

new steady-state equilibrium features more firms and higher aggregate output. Since firms are on average smaller, labor productivity increases by 1.1 percent, output increases by 1.8 percent and unemployment falls by 8 percent. In Figure 3, we compare impulse responses for the three calibrations with different curvature parameters. Relative to the model with nearly linear recruitment costs, convex costs generate a pronounced labor market propagation, featuring sluggish adjustments of the job–finding rate and of the vacancy–unemployment ratio, which are similar to the responses of these variables to a permanent productivity shock in vector autoregressions (see Appendix E for details). Fujita and Ramey (2007) and Shimer (2005) argue that standard search and matching models cannot generate such patterns because market tightness and the job-finding rate are jump variables which correlate perfectly with aggregate productivity. The bottom graphs in Figure 3 show that this is also true in our model when vacancy costs are linear,[28] but not when they are convex in which case both variables lag behind aggregate productivity by 2-3 quarters.

We emphasize that the sluggish model dynamics come about for the same parameterizations of the recruitment technology which also give rise to plausible variations of vacancy yields across firms. Micro-level features are thus directly linked to the dynamics at the aggregate level. Lagged responses to productivity shocks are neither picked up by most random search models, nor by existing models with directed search, such as Shi (2009), Menzio and Shi (2010, 2011), and Schaal (2010). In our model, convexity of recruitment technologies in combination with the entry of new firms contribute to the delayed response of the labor market: the positive shock triggers a surge of entrant firms who create only few jobs when they are small but more as they grow larger. With linear recruitment costs, all firms (young and old) would directly jump to their optimal sizes.

[28]Equation (18) implies that m is a function of the aggregate state μ_i alone if marginal vacancy costs are constant.

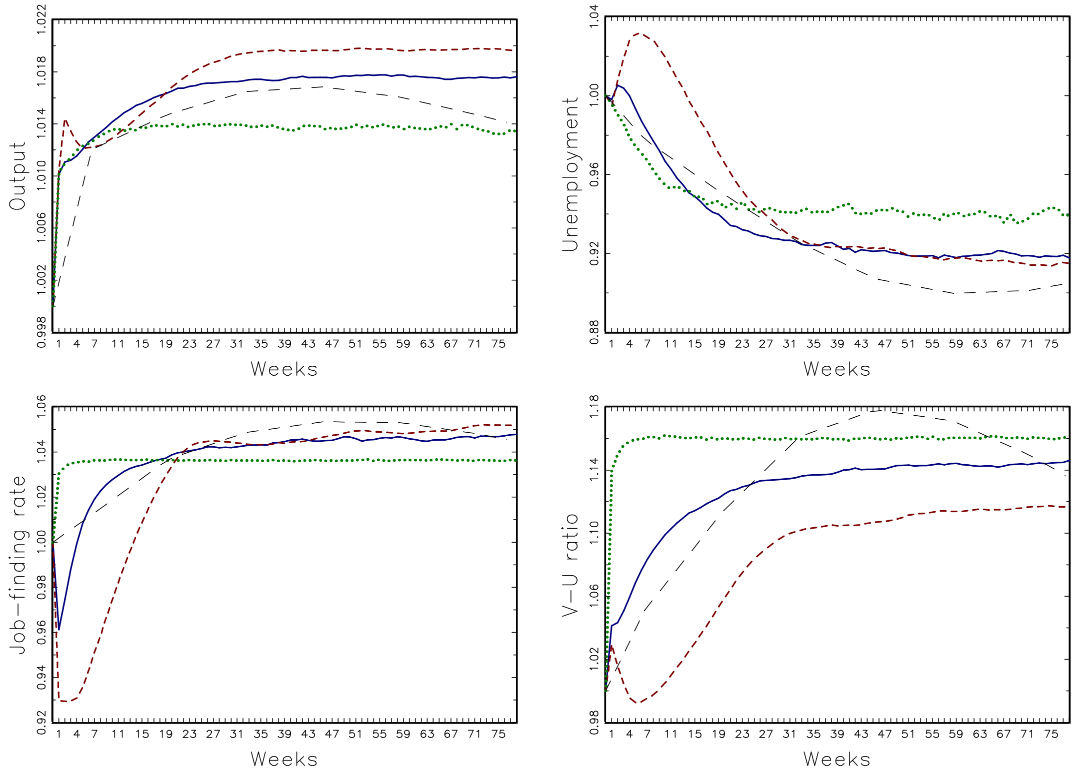

Figure 3: Impulse response to a permanent 1% increase in aggregate productivity. The dashed curves are responses from a VAR for U.S. data (see Appendix E for details), the blue (solid) curves are for the model with cubic recruitment costs ($\gamma = 2$), the green (dotted) curves are for $\gamma = 0.1$ and the red (closely dashed) curves are for $\gamma = 8$.

To study business cycle properties, we solve the model as outlined in Section 3.2. The aggregate productivity parameter attains five equally distant values in the interval $[z_{min}, 2 - z_{min}]$, and the Markov process for z is a mean–reverting process with transition probability ψ, as described in Appendix C of Shimer (2005). The two parameters $(z_{min}, \psi) = (0.95, 0.015)$ are set to match a quarterly standard deviation and autocorrelation of labor productivity around trend of 0.013 and 0.76. As before, we allow the entry cost K to vary with the aggregate state, so as to target the volatility of job creation at opening firms.[29] Table 3 shows the

[29]Specifically, we let K vary between 199.75 in the lowest productivity state and 209.67 in

outcome of this experiment for volatility and comovement with aggregate output. The key labor market variables are amplified almost as much as in the data, which is not too surprising given that we calibrate the opportunity cost of work rather close to marginal labor productivity. Relative to a model with homogeneous firms (Hagedorn and Manovskii (2008)), firm heterogeneity and decreasing returns add no more amplification.[30] Besides amplification, our model generates correlation patterns with aggregate output which are consistent with the data. Particularly, it captures procyclical job-finding rates and countercyclical separation rates. We also note that the correlation between labor productivity and the job-finding rate is positive though not perfect, in contrast to Shimer's (2005) calibration of the search and matching model with homogeneous firms.

The last two rows of Table 3 show that our business cycle model roughly captures the volatility and comovement of the aggregate vacancy yield and of the recruiting intensity as calculated by Davis et al. (2013) for JOLTS data, 2001-2011. In particular, we can decompose the aggregate vacancy yield as

$$\frac{H}{V} = m(\lambda) \sum_i \frac{m_i}{m(\lambda)} \frac{V_i}{V} \equiv m(\lambda) r \ , \tag{21}$$

where H, V are aggregate hires and vacancies, λ is the aggregate unemployment-vacancy ratio, and (m_i, V_i) are recruitment policies of firm i. Since λ is countercyclical, so is the aggregate vacancy yield, although less than a standard aggregate matching function would predict. The term r in equation (21) is a measure of the (vacancy-weighted) "recruiting intensity" which turns out to be procyclical, both in the data and in the model with $\gamma = 2$.[31] The reason why r is procyclical

the highest state for $\gamma = 2$.

[30]This is consistent with Krause and Lubik (2007), Faccini and Ortigueira (2010) and Hawkins (2011) who obtain little amplification of technology shocks in labor market models with intra-firm bargaining.

[31]Our measure of the recruiting intensity corresponds to the variable $\overline{q}_t^{1-\alpha}$ in equation (9) of Davis et al. (2013). We set $\alpha = 0.5$ as in their paper to calculate the moments in Table 3.

Table 3: Business cycle statistics

	Data		Model ($\gamma = 2$)	
	Relative volatility	Correlation with output	Relative volatility	Correlation with output
Productivity	0.79	0.659	0.45	0.872
Unemployment	7.87	-0.853	7.00	-0.921
Vacancies	8.34	0.428	5.00	0.504
Job–finding rate	4.73	0.829	4.22	0.929
Separation rate	3.28	-0.576	3.66	-0.752
Vacancy yield	7.20	-0.803	5.58	-0.940
Recruting intensity	1.22	0.852	1.56	0.861

Notes: All variables are logged and HP filtered with parameter 1600. Relative volatility measures the standard deviation of a variable divided by the standard deviation of output. Data are for the U.S. labor market (1950-2011), except the job–finding rate and separation rate series (1951-2007) which were constructed by Robert Shimer (see Shimer (2012) and his webpage http://sites.google.com/site/robertshimer/research/flows) and the vacancy yields and recruiting intensity series (2001-2011) which were constructed by Davis et al. (2013). Monthly series are converted in quarterly series by time averaging. The model statistics are obtained from a simulation of $2 \cdot 10^5$ firms over a period of 26000 weeks. Weekly series are converted into quarterly series by time averaging.

in our model is that m is concave and that the cross-sectional dispersion of λ_i (m_i) is countercyclical.[32]

Finally we briefly report the cyclicality of wages. Although our theory specifies present values rather than the time profiles of wages, we consider flat wage contracts as an empirically relevant special case. With this specification, we find that aggregate wages are procyclical and rigid, while wages for new hires are more volatile. In particular, the elasticity of wages for all workers (new hires) with respect to productivity is 0.13 (0.64), which compares with the estimates of Haefke et al. (2013) who report elasticities in the range 0.1-0.3 (0.6-1). Other implementations are clearly conceivable that would give rise to greater wage cyclicality,

[32]This seems consistent with the observation of Davis et al. (2012) that the cross-industry dispersion of job-filling rates increased during the Great Recession. We note that the procyclicality of r vanishes in our model with nearly linear recruitment costs ($\gamma = 0.1$) because all firms fill vacancies at the same matching rate.

but it is good to know that our model can account for substantial wage rigidity.

5 Conclusions

This paper investigates job reallocation in a model where firms actively compete for workers in a frictional labor market. Meaningful dynamics arise when firms cannot instantly post vacancies at constant marginal costs - for example because existing workers are required for recruitment. Firms that want to expand quickly are willing to pay higher salaries to attract more workers and hence fill vacancies faster. Matching rates are therefore not an aggregate object, as in most of the search literature, but are firm-specific as recently documented by Davis et al. (2013). Calibrated versions of the model show that it can account for this variation in vacancy yields, alongside other cross-sectional features. The same reasons that let firms vary their vacancy yields also induce delayed aggregate responses of key labor market variables to productivity shocks.

Apart from this contribution, this paper lays out a competitive search model with heterogeneous firms facing convex recruitment costs. This model provides an alternative to the current workhorse models for large firms in search markets which are based on random search and bargaining. We establish substantial differences between these environments: Competition for workers induces firm-specific matching rates, while they are identical in random-search models. Multi-worker firms in that environment always engage in inefficient hiring, whereas we show that our environment retains the efficiency properties known from competition in economies with single-worker firms. Finally, our model remains tractable both in and out of steady state, which makes it useful for applied purposes.

We conclude by noting that this framework is flexible for extensions. It is straightforward to allow for variable capital investment or for worker heterogeneity, as long as the firm-level production functions retain decreasing returns in variable inputs. A further extension is to introduce risk aversion. In constant-returns environments with exogenous separation rates, Acemoglu and Shimer (1999) and Rudanko (2011) introduce long-term contracting and analyze the implications for risk sharing, unemployment insurance and labor market dynamics. Our model could also be augmented along these lines. But different from our results, equilibrium would cease to be socially efficient, provided that the planner is allowed to redistribute income to the unemployed. Lack of unemployment insurance induces workers to search too much for low-paying but easy-to-get jobs (as in Acemoglu and Shimer (1999)), and should lead to excess employment in low-productivity firms and therefore to a misallocation of labor between heterogeneous firms.

Such extensions would also make policy analysis more relevant. Already in the current setup, we conduct a preliminary policy experiment. Since our model generates aggregate dynamics which are largely in line with the U.S. business cycle, we use it to investigate the impact of hiring credits on stabilization. Somewhat surprisingly, we find that these subsidies, especially when implemented in a counter-cyclical manner, have a destabilizing effect on the labor market since they foster more labor reallocation during recessions (see Appendix F for details).

References

Acemoglu, D. and W. Hawkins (2013), "Search with multi-worker firms." Forthcoming in *Theoretical Economics*.

Acemoglu, D. and R. Shimer (1999), "Efficient unemployment insurance." *Journal of Political Economy*, 107, 893–928.

Amir, R. (1996), "Sensitivity analysis of multisector optimal economic dynamics." *Journal of Mathematical Economics*, 25, 123–141.

Belzil, C. (2000), "Job creation and job destruction, worker reallocation, and wages." *Journal of Labor Economics*, 18, 183–203.

Bentolila, S. and G. Bertola (1990), "Firing costs and labour demand: How bad is eurosclerosis?" *Review of Economic Studies*, 57, 381–402.

Bertola, G. and R. Caballero (1994), "Cross–sectional efficiency and labour hoarding in a matching model of unemployment." *Review of Economic Studies*, 61, 435–456.

Blatter, M., S. Muehlmann, and S. Schenker (2012), "The costs of hiring skilled workers." *European Economic Review*, 56, 20–35.

Brown, C. and J. Medoff (2003), "Firm age and wages." *Journal of Labor Economics*, 21, 677–697.

Burdett, K. and D. Mortensen (1998), "Wage differentials, employer size and unemployment." *International Economic Review*, 39, 257–273.

Burdett, K., S. Shi, and R. Wright (2001), "Pricing and matching with frictions." *Journal of Political Economy*, 109, 1060–1085.

Cooper, R., J. Haltiwanger, and J. Willis (2007), "Search frictions: Matching aggregate and establishment observations." *Journal of Monetary Economics*, 54, 56–78.

Davis, S., J. Faberman, and J. Haltiwanger (2012), "Recruiting intensity during and after the great recession: National and industry evidence." *American Economic Review*, 102, 584–588.

Davis, S., J. Faberman, and J. Haltiwanger (2013), "The establishment–level behavior of vacancies and hiring." *Quarterly Journal of Economics*, 128, 581–622.

Davis, S., J. Faberman, J. Haltiwanger, and I. Rucker (2010), "Adjusted estimates of worker flows and job openings in JOLTS." In *Labor in the New Economy* (K. Abraham, J. Spletzer, and M. Harper, eds.), 187–216, University of Chicago Press, Chicago.

Ebell, M. and C. Haefke (2009), "Product market deregulation and the US employment miracle." *Review of Economic Dynamics*, 12, 479–504.

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

Faccini, R. and S. Ortigueira (2010), "Labor–market volatility in the search–and–matching model: The role of investment–specific technology shocks." *Journal of Economic Dynamics and Control*, 34, 1509–1527.

Fujita, S. (2011), "Dynamics of worker flows and vacancies: Evidence from the sign restriction approach." *Journal of Applied Econometrics*, 26, 89–121.

Fujita, S. and M. Nakajima (2013), "Worker flows and job flows: A quantitative investigation." Working Paper 13-9, Federal Reserve Bank of Philadelphia.

Fujita, S. and G. Ramey (2007), "Job matching and propagation." *Journal of Economic Dynamics and Control*, 31, 3671–3698.

Galenianos, M. and P. Kircher (2009), "Directed search with multiple job applications." *Journal of Economic Theory*, 114.

Garibaldi, P. and E. Moen (2010), "Competitive on–the–job search." Unpublished Manuscript.

Guerrieri, V. (2008), "Inefficient unemployment dynamics under asymmetric information." *Journal of Political Economy*, 116, 667–708.

Haefke, C., M. Sonntag, and T. van Rens (2013), "Wage rigidity and job creation." *Journal of Monetary Economics*, 60, 887–899.

Hagedorn, M. and I. Manovskii (2008), "The cyclical behavior of equilibrium unemployment and vacancies revisited." *American Economic Review*, 98, 1692–1706.

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

Hawkins, W. (2010), "Bargaining with commitment between workers and large firms." Unpublished Manuscript.

Hawkins, W. (2011), "Do large-firm bargaining models amplify and propagate aggregate productivity shocks?" Unpublished Manuscript.

Hawkins, W. (2013), "Competitive search, efficiency, and multi-worker firms." *International Economic Review*, 54, 219–251.

Helpman, E. and O. Itskhoki (2010), "Labor market rigidities, trade and unemployment." *Review of Economic Studies*, 77, 1100–1137.

Hopenhayn, H. and R. Rogerson (1993), "Job turnover and policy evaluation: A general equilibrium analysis." *Journal of Political Economy*, 101, 915–938.

Hosios, A. J. (1990), "On the efficiency of matching and related models of search and unemployment." *Review of Economic Studies*, 57, 279–298.

Koeniger, W. and J. Prat (2007), "Employment protection, product market regulation and firm selection." *Economic Journal*, 117, F302–F332.

Krause, M. and T. Lubik (2007), "Does intra–firm bargaining matter for business cycle dynamics?" Deutsche Bundesbank Discussion Paper 17/2007.

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

Lentz, R. and D. Mortensen (2010), "Labor market friction, firm heterogeneity, and aggregate employment and productivity." University of Wisconsin Discussion Paper.

Lester, B. (2010), "Directed search with multi-vacancy firms." *Journal of Economic Theory*, 145, 2108–2132.

Menzio, G. and E. Moen (2010), "Worker replacement." *Journal of Monetary Economics*, 57, 623–636.

Menzio, G. and S. Shi (2010), "Block recursive equilibria for stochastic models of search on the job." *Journal of Economic Theory*, 145, 1453–1494.

Menzio, G. and S. Shi (2011), "Efficient search on the job and the business cycle." *Journal of Political Economy*, 119, 468–510.

Merz, M. and E. Yashiv (2007), "Labor and the market value of the firm." *American Economic Review*, 97, 1419–1431.

Moen, E. (1997), "Competitive search equilibrium." *Journal of Political Econ-*

omy, 105, 385–411.

Moscarini, G. and F. Postel-Vinay (2013), "Stochastic search equilibrium." *Review of Economic Studies*, 80, 1545–1581.

Pissarides, C. (2000), *Equilibrium Unemployment Theory*, 2nd edition. The MIT Press, Cambridge, MA.

Postel-Vinay, F. and J.-M. Robin (2002), "Equilibrium wage dispersion with worker and employer heterogeneity." *Econometrica*, 70, 2295–2350.

Rogerson, R., R. Shimer, and R. Wright (2005), "Search-theoretic models of the labor market: A survey." *Journal of Economic Literature*, 43, 959–988.

Rudanko, L. (2011), "Aggregate and idiosyncratic risk in a frictional labor market." *American Economic Review*, 101, 2823–2843.

Schaal, E. (2010), "Uncertainty, productivity and unemployment in the great recession." Unpublished Manuscript.

Schmieder, J. (2013), "What causes wage dispersion? Evidence from new firms." Unpublished Manuscript.

Shi, S. (2009), "Directed search for equilibrium wage-tenure contracts." *Econometrica*, 77, 561–584.

Shimer, R. (2005), "The cyclical behavior of equilibrium unemployment and vacancies." *American Economic Review*, 95, 25–49.

Shimer, R. (2010), *Labor Markets and Business Cycles*. Princeton University Press.

Shimer, R. (2012), "Reassessing the ins and outs of unemployment." *Review of Economic Dynamics*, 15, 127–148.

Smith, E. (1999), "Search, concave production and optimal firm size." *Review of Economic Dynamics*, 2, 456–471.

Stole, L. and J. Zwiebel (1996), "Intrafirm bargaining under non–binding contracts." *Review of Economic Studies*, 63, 375–410.

Appendix A: Proofs

Proof of Proposition 1:
Rewrite problem (5) to express the dependence of the value function on x and on the workers' search value ρ as the solution to the dynamic programming problem

$$\begin{aligned} G(L,x;\rho) &= \max_{(m,V)\geq 0} xF(L) - C(V,L,x) - D(m;\rho)V + \beta(1-\delta)G(L_+,x;\rho) \\ &\quad \text{s.t. } L_+ = L(1-s) + mV \ , \end{aligned} \tag{22}$$

where function $D(m;\rho)$ is defined in the text. It is increasing, strictly convex in m and increasing in ρ. This problem is equivalently defined on a compact state space $L \in [0,\overline{L}]$ where $\overline{L}$ is so large that it never binds. This is possible because of the Inada condition $\lim_{L\to\infty} F'(L) = 0$. The RHS in problem (22) defines an operator T which maps a continuous function $G_0(L,x;\rho)$, defined on $\mathcal{S} = [0,\overline{L}] \times [0,\overline{x}] \times [0,\overline{\rho}]$ into a continuous function $G_1(L,x;\rho) = T(G_0)(L,x;\rho)$ defined on the same domain. Here $\overline{x}$ and $\overline{\rho}$ are arbitrary upper bounds on x and ρ. Operator T is a contraction, therefore there exists a unique fixed point G^* which is a continuous function and which is the limit of any sequence G_n defined by $G_n = T(G_{n-1})$.
Starting from a continuous G_0 that is differentiable and weakly increasing in L and x and weakly decreasing in ρ, successive application of T yields a sequence G_n where each element shares these properties. Since the subset of continuous functions on $\mathcal{S}$ that are weakly increasing in L and x and weakly decreasing in ρ is closed under the sup norm, the limit G^* of sequence G_n is in this set. Because $xF(L) - C(V,L,x)$ is strictly increasing in (L,x) and since $D(m;\rho)$ is strictly decreasing in ρ, the limit G is strictly increasing in x and L and strictly decreasing in ρ.
We show in subsequent Lemmata 1 and 2 that T maps functions that are differentiable and concave in L and supermodular in L and x into functions with the same properties. Since the subset of concave and supermodular functions is closed, the same arguments as above imply that the unique fixed point G^* is concave in L and supermodular in (L,x). Since function $xF(L) - C(V,L,x)$ is strictly concave in L, G^* is also strictly concave in L. Concavity in L and differentiability of $xF(L) - C(V,L,x)$ together with the theorem of Benveniste and Scheinkman establishes differentiability of G^* in L.
Before we establish the remaining results, rewrite (22) in terms of hirings $H = mV$. Dropping argument ρ from G, we can equivalently write (22) as

$$G(L,x) = \max_H xF(L) - \mathcal{C}(H,L,x) + \beta(1-\delta)G(L(1-s)+H,x) \tag{23}$$

where

$$\mathcal{C}(H,L,x) \equiv \min_m C\Big(\frac{H}{m},L,x\Big) + D(m)\frac{H}{m} \ . \tag{24}$$

The right hand side of (23) is an equivalent expression of the fixed-point operator T. As will become clear, the per period return $xF(L) - \mathcal{C}(H,L,x)$ is supermodular in (L,H), but when $C_{13} > 0$ (which arises in first specification in (1) for $h > 0$) the per period return is strictly submodular in (H,x) and in (L_+,x) when one writes

$H = L_+ - (1-s)L$, which renders standard tools to prove supermodularity (e.g., Amir (1996)) inapplicable. To proceed, the optimality condition for problem (24) is

$$C_1\Big(\frac{H}{m}, L, x\Big) = D'(m)m - D(m) \ . \tag{25}$$

Differentiate this equation to obtain

$$\frac{dm}{dH} = \frac{C_{11}}{C_{11}\dfrac{H}{m} + D''(m)m^2} > 0 \ , \tag{26}$$

$$\frac{dm}{dL} = \frac{C_{12}m}{C_{11}\dfrac{H}{m} + D''(m)m^2} = \frac{C_{12}m}{C_{11}}\frac{dm}{dH} \leq 0 \ , \tag{27}$$

$$\frac{dm}{dx} = \frac{C_{13}m}{C_{11}\dfrac{H}{m} + D''(m)m^2} = \frac{C_{13}m}{C_{11}}\frac{dm}{dH} \geq 0 \ . \tag{28}$$

Therefore, we can express the derivatives of cost function $\mathcal{C}$ as

$$\mathcal{C}_1 = D'(m) > 0 \ ,$$

$$\mathcal{C}_2 = C_2 \ ,$$

$$\mathcal{C}_{11} = D''(m)\frac{dm}{dH} > 0 \ , \tag{29}$$

$$\mathcal{C}_{12} = D''(m)\frac{dm}{dL} \leq 0 \ , \tag{30}$$

$$\mathcal{C}_{22} = C_{22} - C_{12}\frac{H}{m^2}\frac{dm}{dL} \ , \tag{31}$$

$$\mathcal{C}_{13} = D''(m)\frac{dm}{dx} \geq 0 \ , \tag{32}$$

$$\mathcal{C}_{23} = C_{23} - C_{12}\frac{H}{m^2}\frac{dm}{dx} \ . \tag{33}$$

Lemma 1: *Suppose that G is twice differentiable and concave in L. Then $T(G)$ is twice differentiable and*

(a) concave in L if the following condition holds:

$$\mathcal{C}_{12}^2 + \mathcal{C}_{11}[xF'' - \mathcal{C}_{22}] \leq 0 \ . \tag{34}$$

(b) concave in L and supermodular in (L, x) if G is supermodular in (L, x) and if (34) and the following condition hold:

$$\mathcal{C}_{12}\mathcal{C}_{13} + \mathcal{C}_{11}[F' - \mathcal{C}_{23}] \geq 0 \ . \tag{35}$$

Lemma 2:

(a) Condition (34) holds under the following condition on the original cost function

C:

$$\mathcal{C}_{12}^2 + \mathcal{C}_{11}[xF'' - \mathcal{C}_{22}] \leq 0 \ . \tag{36}$$

(b) Condition (35) holds under the following condition on the original cost function C:

$$\mathcal{C}_{12}\mathcal{C}_{13} + \mathcal{C}_{11}[F' - \mathcal{C}_{23}] \geq 0 \ . \tag{37}$$

Proof of Lemma 1: Consider $T(G)$ defined by the RHS of (23).
Part (a). Since G is a concave and twice differentiable function of L, $T(G)$ is also twice differentiable, and a policy function exists and is differentiable. Differentiate $T(G)$ twice with respect to L to obtain

$$\frac{d^2(TG)}{dL^2} = xF'' - \mathcal{C}_{22} + \beta\varphi(1-s)G_{11} + \Big[-\mathcal{C}_{12} + \beta\varphi G_{11}\Big]\frac{dH}{dL} \ . \tag{38}$$

Differentiate the FOC $\mathcal{C}_1 = \beta(1-\delta)G_1$ with respect to L to obtain

$$\frac{dH}{dL} = \frac{\beta\varphi G_{11} - \mathcal{C}_{12}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ . \tag{39}$$

Substitute this into (38) to obtain

$$\frac{d^2(TG)}{dL^2} = xF'' - \mathcal{C}_{22} + \frac{\beta\varphi(1-s)G_{11}\mathcal{C}_{11} + \mathcal{C}_{12}^2 - 2\beta\varphi G_{11}\mathcal{C}_{12}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ .$$

In the last term, the denominator is positive and larger than $\mathcal{C}_{11}$. In the numerator, all terms involving G_{11} are negative (due to (29) and (30)); hence the numerator is smaller than $\mathcal{C}_{12}^2$. Therefore,

$$\frac{d^2(TG)}{dL^2} \leq xF'' - \mathcal{C}_{22} + \frac{\mathcal{C}_{12}^2}{\mathcal{C}_{11}} \ ,$$

which is non-positive under (34). Hence, T maps a concave and twice differentiable function into a function with the same properties.
Part (b). Since G is a concave, supermodular and twice differentiable function of (L, x), $T(G)$ is twice differentiable and a differentiable policy function exists. Differentiate $T(G)$ twice with respect to L and x to obtain

$$\frac{d^2(TG)}{dLdx} = F' - \mathcal{C}_{23} + \beta\varphi G_{12} + \Big[-\mathcal{C}_{12} + \beta\varphi G_{11}\Big]\frac{dH}{dx} \ . \tag{40}$$

Differentiate the FOC $\mathcal{C}_1 = \beta(1-\delta)G_1$ with respect to x to obtain

$$\frac{dH}{dx} = \frac{\beta(1-\delta)G_{12} - \mathcal{C}_{13}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ . \tag{41}$$

Substitute this into (40) to obtain

$$\frac{d^2(TG)}{dLdx} = F' - \mathcal{C}_{23} + \frac{\beta\varphi G_{12}\mathcal{C}_{11} + \mathcal{C}_{12}\mathcal{C}_{13} - \beta(1-\delta)G_{12}\mathcal{C}_{12} - \beta\varphi G_{11}\mathcal{C}_{13}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ .$$

In the last term, the denominator is positive and larger than $\mathcal{C}_{11}$. In the numerator, all terms involving G_{11} and G_{12} are non-negative (due to (29), (30) and (32)); hence the numerator is greater than $\mathcal{C}_{12}\mathcal{C}_{13} \leq 0$. Therefore,

$$\frac{d^2(TG)}{dLdx} \geq F' - \mathcal{C}_{23} + \frac{\mathcal{C}_{12}\mathcal{C}_{13}}{\mathcal{C}_{11}} \ ,$$

which is non-negative under (35). Hence, $T(G)$ is supermodular. □

Proof of Lemma 2: From (27), (28), (29), (30) and (32) follows that

$$\mathcal{C}_{12} = \frac{\mathcal{C}_{11}C_{12}m}{C_{11}} \ , \tag{42}$$

$$\mathcal{C}_{13} = \frac{\mathcal{C}_{11}C_{13}m}{C_{11}} \ . \tag{43}$$

Furthermore, substituting (30) into (27), and substituting (32) into (28) to eliminate $D''(m)$ imply that

$$\mathcal{C}_{22} = C_{22} - \frac{C_{12}^2}{C_{11}} + \frac{mC_{12}}{C_{11}}\mathcal{C}_{12} \ , \tag{44}$$

$$\mathcal{C}_{23} = C_{23} - \frac{C_{12}}{C_{11}}\Big[C_{13} - m\mathcal{C}_{13}\Big] \ . \tag{45}$$

Part (a): Rewrite (34) using (42) and (44) to obtain the equivalent condition

$$xF'' - C_{22} + \frac{C_{12}^2}{C_{11}} \leq 0 \ .$$

Because of $C_{11} > 0$, this condition is equivalent to (36).
Part (b): Rewrite (35) using (42), (43) and (45) to obtain the equivalent condition

$$F' - C_{23} + \frac{C_{12}C_{13}}{C_{11}} \geq 0 \ .$$

Because of $C_{11} > 0$, this condition is equivalent to (37). □

It follows from Lemma 1 and 2 that the value function $G(L, x)$ is concave in L and supermodular in (L, x) because property (C) together with the assumption that $xF(.) - C(.)$ is concave in (L, V) guarantee both (36) and (37).
Because of strict concavity of problem (22), policy functions $m^x(L)$ and $V^x(m^x(L), L)$ exist. To derive first-order conditions (6) and (7) is straightforward: The first condition directly follows from (25); the second follows from the intertemporal optimality condition $\mathcal{C}_1(H, L, x) = \beta(1-\delta)G_1(L(1-s) + H, x)$ and from using the envelope theorem and (6).

The properties of V^x stated in Proposition 1 were already established in the main text. To see how $m^x(L)$ depends on L, use (27) and (39) to get

$$\frac{dm^x(L)}{dL} = \frac{dm(H,L,x)}{dL} + \frac{dm(H,L,x)}{dH}\frac{dH}{dL} = \frac{dm}{dH}\Big[\frac{C_{12}m}{C_{11}} + \frac{\beta\varphi G_{11} - \mathcal{C}_{12}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}}\Big] \ .$$

Because of

$$\frac{C_{12}m}{C_{11}} = \frac{\mathcal{C}_{12}}{\mathcal{C}_{11}} \leq \frac{\mathcal{C}_{12}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ ,$$

the term in [.] is negative, and so is $dm^x/(dL)$.

To verify that m is increasing in x, use (28) and (41) to get

$$\frac{dm^x(L)}{dx} = \frac{dm(H,L,x)}{dx} + \frac{dm(H,L,x)}{dH}\frac{dH}{dx} = \frac{dm}{dH}\Big[\frac{C_{13}m}{C_{11}} + \frac{\beta(1-\delta)G_{12} - \mathcal{C}_{13}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}}\Big] \ .$$

Because of

$$\frac{C_{13}m}{C_{11}} = \frac{\mathcal{C}_{13}}{\mathcal{C}_{11}} \geq \frac{\mathcal{C}_{13}}{\mathcal{C}_{11} - \beta(1-\delta)G_{11}} \ ,$$

the term in [.] is positive, and so is $dm^x/(dx)$. □

Proof of Corollary 2: Because of exogenous separations, the growth rate of a firm, $[mV - sL]/L$ is perfectly correlated with the job-creation rate,

$$\mathrm{JCR}(x,L) = m^x(L)\frac{V^x(m^x(L),L)}{L} \ .$$

Differentiation of the job-creation rate with respect to x implies

$$\frac{d\mathrm{JCR}}{dx} = \frac{dm^x}{dx}\frac{V^x}{L} + \frac{m^x}{L}\frac{dV^x}{dx} + \frac{m^x}{L}\frac{dV^x}{dm}\frac{dm^x}{dx} \ .$$

In this expression, the first and the third term are strictly positive. Under the second cost function in (1), the second term is zero. Under the first cost function in (1), the second term is zero when $h = 0$, and negative but small if h is small. Thus, $d\mathrm{JCR}/(dx)$ is positive if h is sufficiently small.

Differentiation of the job-creation rate with respect to L implies

$$\frac{d\mathrm{JCR}}{dL} = \frac{dm^x}{dL}\frac{V^x}{L} + \frac{m^x}{L}\frac{dV^x}{dL} + \frac{m^x}{L}\frac{dV^x}{dm}\frac{dm^x}{dL} - m\frac{V^x}{L^2} \ .$$

In this expression, the first, the third and the fourth term are strictly negative. Under the second cost function in (1), $\frac{dV^x}{dL} = \frac{V^x}{L}$, and the second and forth terms cancel out. Under the first cost function in (1), the second term is zero when $h = 0$, and positive but small if h is small. Thus, $d\mathrm{JCR}/(dL)$ is negative if h is sufficiently small. □

Lemma 3: *In the model of Section 2 with recruitment cost $C(V,L,x) = xF(L) - xF(L - hV) + cV$, job-filling rates in the optimal firm's problem follow the dynamic* *equation*

$$\rho\Big[m_{t+1}\lambda'(m_{t+1}) - \lambda(m_{t+1})\Big] - (b+\rho)h - c = \frac{\rho h}{\beta(1-\delta)}\Big[\lambda'(m_t) - \beta\varphi\lambda'(m_{t+1})\Big]. \quad (46)$$

It has a unique steady state solution $m^* > 0$ if, and only if,

$$h < \frac{\beta(1-\delta)\overline{m}}{1-\beta\varphi}\ , \quad (47)$$

with $\overline{m} \equiv \lim_{m\to 1} m - \dfrac{\lambda(m)}{\lambda'(m)} > 0$. *Under this condition, any sequence $m_t > 0$ satisfying this dynamic equation converges to m^*.*

Proof of Lemma 3: It is straightforward to derive (46) by substitution of (6) into (7). A steady state m^* must satisfy the condition

$$\rho\Big[m - \frac{\lambda(m)}{\lambda'(m)}\Big] = \frac{\rho h(1-\beta\varphi)}{\beta(1-\delta)} + \frac{(b+\rho)h + c}{\lambda'(m)}\ . \quad (48)$$

The LHS is strictly increasing and goes from 0 to $\rho\overline{m}$ as m goes from 0 to 1. The RHS is decreasing in m with limit $\rho h(1-\beta\varphi)/[\beta(1-\delta)]$ for $m \to 1$. Hence, a unique steady state m^* exists iff (47) holds.[33] Furthermore, differentiation of (46) at m^* implies that

$$\left.\frac{dm_{t+1}}{dm_t}\right|_{m^*} = \frac{h}{\beta(1-\delta)m^* + h\beta\varphi}\ ,$$

which is positive and smaller than one iff

$$h < \frac{\beta(1-\delta)m^*}{1-\beta\varphi}\ .$$

But this inequality must be true because (48) implies

$$h = \frac{\rho[m^*\lambda'(m^*) - \lambda(m^*)] - c}{\frac{\rho[1-\beta\varphi]}{\beta(1-\delta)}\lambda'(m^*) + b + \rho} < \frac{\beta(1-\delta)m^*}{1-\beta\varphi}\ .$$

Therefore, the steady state m^* is locally stable. Moreover, the dynamic equation defines a continuous, increasing relation between m_{t+1} and m_t which has only one intersection with the 45-degree line. Hence, $m_{t+1} > m_t$ for any $m_t < m^*$ and $m_{t+1} < m_t$ for any $m_t > m^*$, which implies that m_t converges to m^* from any initial value $m_0 > 0$. □

Proof of Proposition 2:
It remains to prove existence and uniqueness. From Proposition 1 follows that the entrant's value function $J^x(0,0)$ is decreasing and continuous in ρ. Hence the expected profit prior to entry,

$$\Pi^*(\rho) \equiv \sum_{x\in X} \pi(x)J^x(0,0)$$

[33] If this condition fails, firms cannot profitably recruit workers.

is a decreasing and continuous function of ρ. Moreover, the function is strictly decreasing in ρ whenever it is positive. This also follows from the proof of Proposition 1 which shows that $G(0, x; \rho)$ is *strictly* decreasing in ρ when the new firm x recruits workers ($V^x(m^x(0), 0) > 0$). If no new firm recruits workers, expected profit of an entrant cannot be positive. Hence, equation (8) can have at most one solution for any $K > 0$. This implies uniqueness, with entry of firms if (8) can be fulfilled or without entry of firms otherwise. A solution to (8) exists provided that K is sufficiently small. To see this, $\Pi^*(0)$ is strictly positive because of $F'(0) = \infty$: some entrants will recruit workers since the marginal product $G_1(mV, x; \rho)$ is sufficiently large relative to the cost of recruitment and relative to the wage cost which are, for $\rho = 0$, equal to mVb (see equation (22)). But when $\Pi^*(0) > 0$, a sufficiently small value of K guarantees that (8) has a solution since $\lim_{\rho \to \infty} \Pi^*(\rho) = 0$. □

Proof of Proposition 3:
We will show that the first-order conditions that uniquely characterize the decentralized allocation are also first order conditions to the planner's problem. The same auxiliary problem that we employ in the proof of Lemma 4 part (b) (see the proof of Proposition 4) then establishes that the planner cannot improve upon this allocation. We denote by $S_{N,a}$ the derivative of S with respect to N_a and by $S_{L,a,x}$ the derivative of S with respect to L^x_a. The multiplier on the resource constraint is $\mu \geq 0$. First-order conditions with respect to N_0, V^x_a, and m^x_a, $a \geq 0$, are

$$\sum_{x \in X} \pi(x) \Big[xF(0) - C(V^x_0, 0, x) \Big] - K + \beta(1-\delta) S_{N,1} - \mu \sum_{x \in X} \pi(x) \lambda(m^x_0) V^x_0 = 0 \ , \quad (49)$$

$$-N_a \pi(x) \Big[C_1(V^x_a, L^x_a, x) + \mu \lambda(m^x_a) \Big] + \beta S_{L,a+1,x} m^x_a \leq 0 \ , \ V^x_a \geq 0 \ , \quad (50)$$

$$\beta S_{L,a+1,x} - \mu N_a \pi(x) \lambda'(m^x_a) = 0 \ . \quad (51)$$

Here condition (50) holds with complementary slackness. The envelope conditions are, for $a \geq 1$ and $x \in X$,

$$S_{L,a,x} = N_a \pi(x) \Big[xF'(L^x_a) - C'_2(V^x_a, L^x_a, x) - b - \mu \Big] + \beta(1-s) S_{L,a+1,x} \ , \quad (52)$$

$$S_{N,a} = \sum_{x \in X} \pi(x) \Big[xF(L^x_a) - C(V^x_a, L^x_a, x) - bL^x_a \Big] - \mu \sum_{x \in X} \pi(x) \Big(L^x_a + \lambda(m^x_a) V^x_a \Big) + \beta(1-\delta) S_{N,a+1} \ . \quad (53)$$

Use (51) to substitute $S_{L,a,x}$ into (52) to obtain

$$xF'(L^x_{a+1}) - C_2(V^x_{a+1}, L^x_{a+1}, x) - b - \mu = \frac{\mu}{\beta(1-\delta)} [\lambda'(m^x_a) - \beta \varphi \lambda'(m^x_{a+1})] \ .$$

This equation is the planner's intertemporal optimality condition; it coincides with equation (7) for $\mu = \rho$. This is intuitive: when the social value of an unemployed worker μ coincides with the surplus value that an unemployed worker obtains in search equilibrium, the firm's recruitment policy is efficient. Next substitute (51) into (50) to obtain, for $a \geq 0$ and $x \in X$,

$$C_1(V_a^x, L_a^x, x) \geq \mu[m_a^x \lambda'(m_a^x) - \lambda(m_a^x)] \ , \ V_a^x \geq 0 \ . \tag{54}$$

Again for $\mu = \rho$, this condition coincides with the firm's intratemporal optimality condition in competitive search equilibrium, equation (6). Lastly, it remains to verify that entry is socially efficient when the value of a jobless worker is $\mu = \rho$. The planner's choice of firm entry, condition (49), together with the recursive equation for the marginal firm surplus $S_{N,a}$, equation (53), shows that

$$K = \sum_{a \geq 0} [\beta(1-\delta)]^a \sum_{x \in X} \pi(x) \Big[xF(L_a^x) - bL_a^x - C(V_a^x, L_a^x, x) - \mu(L_a^x + \lambda(m_a^x)V_a^x) \Big] \ . \tag{55}$$

On the other hand, the expected profit value of a new firm is

$$\sum_{x \in X} \pi(x) J^x(0,0) = \sum_{a \geq 0} [\beta(1-\delta)]^a \sum_{x \in X} \pi(x) \Big[xF(L_a^x) - W_a^x - C(V_a^x, L_a^x, x) \Big] \ .$$

Hence, the free-entry condition in search equilibrium, equation (8), coincides with condition (55) for $\mu = \rho$ if, for all $x \in X$,

$$\sum_{a \geq 0} [\beta(1-\delta)]^a \Big[(b+\mu)L_a^x + \mu\lambda(m_a^x)V_a^x - W_a^x \Big] = 0 \ . \tag{56}$$

Now after substitution of

$$\begin{aligned} L_a^x &= \sum_{k=0}^{a-1} (1-s)^{a-1-k} m_k^x V_k^x \ , \text{ and} \\ W_a^x &= \sum_{k=0}^{a-1} (1-s)^{a-1-k} V_k^x \Big[\frac{\rho\lambda(m_k^x)(1-\beta\varphi)}{\beta(1-\delta)} + m_k^x(b+\rho) \Big] \end{aligned}$$

into (56), it is straightforward to see that the equation is satisfied for $\mu = \rho$. □

Proof of Proposition 4:
Part (a):
Let $\beta^t \psi(z^t)\mu(z^t) \geq 0$ be the multiplier on the resource constraint (15) in history node z^t. That is, $\mu(z^t)$ is the social value of a worker in history z^t. Write $\boldsymbol{\mu} = (\mu(z^t))$ for the vector of multipliers. Let $G_t(L, x, z^t)$ denote the social value of an existing firm with employment stock L, idiosyncratic productivity x and aggregate productivity history z^t. The sequence G_t obeys the recursive equations

$$\begin{aligned} G_t(L, x, z^t) = \max_{\delta, s, V, m} \ & xz_t F(L) - bL - \mu(z^t)[L + \lambda(m)V] - C(V, L, xz_t) - f \\ & + \beta(1-\delta) E_{x, z^t} G_{t+1}(L_+, x_+, z^{t+1}) \end{aligned} \tag{57}$$

$$\text{s.t.} \quad L_+ = (1-s)L + mV \ ,$$
$$\delta \in [\delta_0, 1], \ s \in [s_0, 1], \ m \in [0,1], \ V \geq 0 \ .$$

We first prove the equivalence between problem (57) and the planner's problem (14) (Lemma 4). Then we show that the reduced problem (16) solves (57) if entry is positive in all states.

Lemma 4:

(a) For given multipliers $\mu(z^t)$, there exist value functions $G_t : \mathbb{R}_+ \times X \times Z^{t+1} \to \mathbb{R}$, $t \geq 0$, satisfying the system of recursive equations (57).

(b) If $\mathbf{X} = (\mathbf{N}, \mathbf{L}, \mathbf{V}, \mathbf{m}, \mathbf{s}, \boldsymbol{\delta})$ is a solution of the planning problem (14) with multipliers $\boldsymbol{\mu} = (\mu(z^t))$, then the corresponding firm policies also solve problem (57) and the complementary-slackness condition

$$\sum_{x \in X} \pi_0(x) G_t(0, x, z^t) \leq K(z_t) \ , \ N_0(z^t) \geq 0 \ , \tag{58}$$

is satisfied for all z^t. Conversely, if $\mathbf{X}$ solves for every firm problem (57) with multipliers $\boldsymbol{\mu}$, and if condition (58) and the resource constraint (15) hold for all z^t, then $\mathbf{X}$ is a solution of the planning problem (14).

Proof of Lemma 4:
Part (a): The RHS in the system of equations in (57) defines an operator T which maps a sequence of bounded functions $G = (G_t)_{t \geq 0}$, with $G_t : [0, \overline{L}] \times X \times Z^t \to \mathbb{R}$ such that $\|G\| \equiv \sup_t \|G_t\| < \infty$, into another sequence of bounded functions $\tilde{G} = (\tilde{G}_t)_{t \geq 0}$ with $\|\tilde{G}\| = \sup_t \|\tilde{G}_t\| < \infty$. Here $\overline{L}$ is sufficiently large such that the bound $L_+ \leq \overline{L}$ does not bind for any $L \in [0, \overline{L}]$. The existence of $\overline{L}$ follows from the Inada condition for F: the marginal product of an additional worker $xzF'(L_+) - b$ must be negative for any $x \in X$, $z \in Z$, for all $L_+ \geq \overline{L}$ with sufficiently large $\overline{L}$; hence no hiring will occur beyond $\overline{L}$. Because the operator satisfies Blackwell's sufficient conditions, it is a contraction in the space of bounded function sequences G. Hence, the operator T has a unique fixed point which is a sequence of bounded functions.
Part (b): Take first a solution $\mathbf{X}$ of the planning problem, and write $\beta^t \psi(z^t) \mu(z^t) \geq 0$ for the multipliers on constraints (15). Then $\mathbf{X}$ maximizes the Lagrange function

$$\begin{aligned} \mathcal{L} \ = \ \max \sum_{t \geq 0, z^t} \beta^t \psi(z^t) \Bigg\{ & - K(z_t) N_0(z^t) + \sum_{a \geq 0, x^a} N(x^a, z^t) \Bigg[x_a z_t F(L(x^a, z^t)) - bL(x^a, z^t) \\ & - f - C(V(x^a, z^t), L(x^a, z^t), x_a z_t) - \mu(z^t) \Big[L(x^a, z^t) + \lambda(m(x^a, z^t)) V(x^a, z^t) \Big] \Bigg] \Bigg\} \end{aligned}$$

For each individual firm, this problem is the sequential formulation of the recursive problem (57) with multipliers $\mu(z^t)$. Hence, firm policies also solve the recursive problem; furthermore, the maximum of the Lagrange function is the same as the sum of the social values of entrant firms plus the social values of firms which already exist at $t = 0$, namely,

$$\begin{aligned}\mathcal{L} &= \max_{N_0(.)} \sum_{t,z^t} \beta^t \psi(z^t) N_0(z^t) \Big[-K(z_t) + \sum_x \pi_0(x) G_t(0, x, z^t) \Big] \\ &\quad + \sum_{z \in Z} \psi(z^0) \sum_{a \geq 1, x^a} N(x^a, z^0) G_0(L(x^a, z^0), x_a, z^0) \ .\end{aligned}$$

This also proves that the complementary-slackness condition (58) describes optimal entry.
To prove the converse, suppose that $\mathbf{X}$ solves for every firm the recursive problem (57) with given multipliers $\mu(z^t)$, and that (58) and the resource constraints (15) are satisfied. Define an auxiliary problem (AP) as an extension of the original planning problem (14) which allows the planner to rent additional workers (or to rent out existing workers) at rental rate $\mu(z^t)$ in period t. Formally, the (AP) differs from the original problem in that the resource constraint (15) is replaced by

$$\sum_{a \geq 0, x^a} N(x^a, z^t) \Big[L(x^a, z^t) + \lambda(m(x^a, z^t)) V(x^a, z^t) \Big] \leq \mathcal{M}(z^t) \ , \tag{59}$$

with $\mathcal{M}(z^t) - 1 > 0$ workers hired or $\mathcal{M}(z^t) - 1 < 0$ workers hired out. Further, the rental cost (rental income) term $-\mu(z^t)[\mathcal{M}(z^t) - 1]$ is added into the braces in the objective function (14). Then it follows immediately that the multiplier on constraint (59) is equal to $\mu(z^t)$. We further claim that allocation $\mathbf{X}$ solves problem (AP), and hence also solves the original planning problem. To see this, suppose that there is an allocation $(\mathbf{X}', \mathcal{M})$ which is feasible for problem (AP) and which strictly dominates $\mathbf{X}$. Write

$$O(x^a, z^t) \equiv x_a z_t F(L(x^a, z^t)) - bL(x^a, z^t) - f - C(V(x^a, z^t), L(x^a, z^t), x_a z_t)$$

for the net output created by firm (x^a, z^t) in allocation $\mathbf{X}$ and write $O'(x^a, z^t)$ for the same object in allocation $\mathbf{X}'$. Further, write S for the total surplus value in allocation $(\mathbf{X}, \mathbf{1})$ and write $S' > S$ for the surplus value in allocation $(\mathbf{X}', \mathcal{M})$. Then

$$\begin{aligned}S' &= \sum_{t \geq 0, z^t} \beta^t \psi(z^t) \Bigg\{ -K(z_t) N_0'(z^t) + \sum_{a \geq 0, x^a} N'(x^a, z^t) O'(x^a, z^t) - \mu(z^t)[\mathcal{M}(z^t) - 1] \Bigg\} \\ &\leq \sum_{t \geq 0, z^t} \beta^t \psi(z^t) \Bigg\{ -K(z_t) N_0'(z^t) + \mu(z^t) \\ &\quad + \sum_{a \geq 0, x^a} N'(x^a, z^t) \Big[O'(x^a, z^t) - \mu(z^t) \Big(L'(x^a, z^t) + \lambda(m'(x^a, z^t)) V'(x^a, z^t) \Big) \Big] \Bigg\} \\ &\leq \sum_{t \geq 0, z^t} \beta^t \psi(z^t) N_0'(z^t) \Big[-K(z_t) + \sum_x \pi_0(x) G_t(0, x, z^t) \Big]\end{aligned}$$

$$
\begin{aligned}
&+\sum_{z\in Z}\psi(z^0)\sum_{a\geq 1,x^a}N(x^a,z^0)G_0(L(x^a,z^0),x_a,z^0)+\sum_{t,z^t}\beta^t\psi(z^t)\mu(z^t)\\
\leq\ &\sum_{t\geq 0,z^t}\beta^t\psi(z^t)N_0(z^t)\Big[-K(z_t)+\sum_x\pi_0(x)G_t(0,x,z^t)\Big]\\
&+\sum_{z\in Z}\psi(z^0)\sum_{a\geq 1,x^a}N(x^a,z^0)G_0(L(x^a,z^0),x_a,z^0)+\sum_{t,z^t}\beta^t\psi(z^t)\mu(z^t)=S\ .
\end{aligned}
$$

Here the first inequality follows from resource constraint (59). The second inequality follows since the discounted sum of surplus values for an individual firm which is of age a at time t, namely

$$
\sum_{\tau\geq t}\beta^{\tau-t}\sum_{x^{a+\tau-t}z^\tau}\psi(z^\tau|z^t)\pi(x^{a+\tau-t}|x^a)\prod_{k=t}^{\tau-1}[1-\delta(x^{a+k-t},z^k)]
$$

$$
\Big[O'(x^{a+\tau-t},z^\tau)-\mu(z^\tau)[L'(x^{a+\tau-t},z^\tau)+\lambda(m'(x^{a+\tau-t},z^\tau))V'(x^{a+\tau-t},z^\tau)]\Big]\ ,
$$

is bounded above $G_t(0,x_0,z_t)$ (for new firms, $a=0$) or by $G_0(L(x^a,z^0),x_a,z^0)$ (for firms of age $a>0$ existing at $t=0$) by definition of G_t. The third inequality follows from the complementary-slackness condition (58): either the term $-K(z_t)+\sum_x\pi_0(x)G_t(0,x,z^t)$ is zero in which case the first summand is zero on both sides of the inequality; or it is strictly negative in which case $N_0(z^t)=0$ and $N_0'(z^t)\geq 0$. The last equality follows from the definition of surplus value S and the assumption that allocation $\mathbf{X}$ solves problem (57) at the level of each individual firm. This proves $S'\leq S$ and hence contradicts the hypothesis $S'>S$. This completes the proof of Lemma 4.□

To complete the proof of Prop. 4, part (a), let μ_i be the multiplier in aggregate state z_i, defined by (16) and (17), and write $M=(\mu_1,\dots,\mu_n)$. With $\mu(z^t)\equiv\mu_i$ for $z_t=z_i$, the unique solution of (57) coincides with the one of (16), i.e. $G_t(L,x,z^t)=G(L,x,i;M)$ for $z_t=z_i$, and also the firm-level policies coincide. If they give rise to an allocation $\mathbf{X}$ with positive entry in all aggregate states z^t, (17) implies that (58) holds for all z^t. Hence Lemma 4(b) implies that $\mathbf{X}$ is a solution of the planning problem.
Part (b): Solving (16) in the stationary case $z=\overline{z}$ involves to find a single value function $G(L,x;M)$. Application of the contraction mapping theorem implies that such a solution exists, is unique, and is continuous and non-increasing in $\mu\in\mathbb{R}$ and strictly decreasing in μ when $G(.)>0$.
Therefore, the function $\Gamma(\mu)\equiv\sum_x\pi_0(x)G(0,x;\mu)\geq 0$ is continuous, strictly decreasing when positive, and zero for large enough μ. Furthermore, when f and b are sufficiently small, $\Gamma(0)>0$; hence when $K>0$ is sufficiently small, there exists a unique $\overline{\mu}\geq 0$ satisfying equation (17).
In the stochastic case $z\in\{z_1,\dots,z_n\}$ and for any given vector $M=(\mu_1,\dots,\mu_n)\in R^n_+$, the system of recursive equations (16) has a unique solution $G(.;M)$. Again this follows from the application of the contraction-mapping theorem. Furthermore, G is differentiable in M, and all elements of the Jacobian $(dG(L,x,i;M)/(d\mu_j))_{i,j}$ are non-positive.

The RHS of (16) defines an operator mapping a function $G(L, x, i; M)$ with a strictly diagonally dominant Jacobian matrix $(dG(L, x, i; M)/(d\mu_j))_{i,j}$ into another function $\tilde{G}$ whose Jacobian matrix $(d\tilde{G}(L, x, i; M)/(d\mu_j))_{i,j}$ is diagonally dominant. This follows since the transition matrix $\psi(z_j|z_i)$ is strictly diagonally dominant and since all elements of $(d\tilde{G}(L, x, i; M)/(d\mu_j))$ have the same (non-positive) sign. Therefore, the unique fixed point has a strictly diagonally dominant Jacobian. Now suppose that $(z_1, \ldots, z_n)$ is close to $(\overline{z}, \ldots, \overline{z})$ and consider the solution $\mu_1 = \ldots = \mu_n = \overline{\mu}$ of the stationary problem. Since the Jacobian matrix $(dG(0, x, i; M)/(d\mu_j))_{i,j}$ is strictly diagonally dominant, it is invertible. By the implicit function theorem, a unique solution M to equation (17) exits. □

For the **proof of Proposition 5**, see Appendix B.

Appendix B: Decentralization

The Workers' Search Problem
Let $U(z^t)$ be the utility value of an unemployed worker in history z^t, and let $W(\mathcal{C}_a, x^k, z^t)$ be the utility value of a worker hired by a firm of age a in contract $\mathcal{C}_a$ who is currently employed at that firm in history x^k, with $k > a$. The latter satisfies the recursive equation

$$W(\mathcal{C}_a, x^k, z^t) = w_a(x^k, z^t) + \beta\Big\{(1 - \varphi_a(x^k, z^t))E_{z^t}U(z^{t+1}) \\ +\varphi_a(x^k, z^t)E_{x^k,z^t}W(\mathcal{C}_a, x^{k+1}, z^{t+1})\Big\} \; . \tag{60}$$

An unemployed worker searches for contracts which promise the highest expected utility, considering that more attractive contracts are less likely to sign. The worker observes all contracts $\mathcal{C}_a$ and he knows that the probability to sign a contract is $m/\lambda(m)$ when m is the firm's matching probability at the offered contract. That is, potential contracts are parameterized by the tuple $(m, \mathcal{C}_a)$. Unemployed workers apply for those contracts where expected surplus is maximized:

$$\rho(z^t) = \max_{(m,\mathcal{C}_a)} \frac{m}{\lambda(m)}(1 - \delta(x^a, z^t))\beta E_{x^a,z^t}\Big[W(\mathcal{C}_a, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big] \; . \tag{61}$$

The Bellman equation for an unemployed worker reads as

$$U(z^t) = b + \rho(z^t) + \beta E_{z^t}U(z^{t+1}) \; . \tag{62}$$

The Firms' Problem
A firm of age a in history (x^a, z^t) takes as given the employment stocks of workers hired in some earlier period, $(L_\tau)_{\tau=0}^{a-1}$, as well as the contracts signed with these workers, $(\mathcal{C}_\tau)_{\tau=0}^{a-1}$. For the contracts to be consistent with the firm's constraints on exit and separations, the retention probabilities must satisfy $\varphi_\tau(x^a, z^t) \leq (1 - s_0)(1 - \delta_0)$. The firm chooses an actual exit probability $\delta \geq \delta_0$ and cohort-specific layoff probabilities s_τ. For these probabilities to be consistent with separation probabilities specified in existing contracts, it must hold that $\delta \leq 1 - \varphi_\tau(x^a, z^t)$ for all $\tau \leq a - 1$, and $s_\tau = 1 - \varphi_\tau(x^a, z^t)/(1 - \delta)$ when $\delta < 1$, with arbitrary choice of s_τ when $\delta = 1$. The firm also decides new contracts $\mathcal{C}_a$ to be posted in V vacancies with desired matching probability m. It is no restriction to presuppose that the firm offers only one type of contract. When J_a is the value function of a firm of age a, the firm's problem is written as

$$J_a\Big[(\mathcal{C}_\tau)_{\tau=0}^{a-1}, (L_\tau)_{\tau=0}^{a-1}, x^a, z^t\Big] = \max_{(\delta,m,V,\mathcal{C}_a)} x_a z_t F(L) - W - C(V, L, x_a z_t) \\ -f + \beta(1 - \delta)E_{x^a,z^t}J_{a+1}\Big[(\mathcal{C}_\tau)_{\tau=0}^{a}, (L_{\tau+})_{\tau=0}^{a}, x^{a+1}, z^{t+1}\Big] \tag{63}$$

$$\text{s.t.} \quad L_{a+} = mV, \; m \in [0, 1], \; V \geq 0, \; L_{\tau+} = L_\tau \frac{\varphi_\tau(x^a, z^t)}{1 - \delta}, \; \tau \leq a - 1 \; , \tag{64}$$

$$\delta \in [\delta_0, \min_{0\leq\tau\leq a-1} 1-\varphi_\tau(x^a,z^t)], \;\; s_0 \leq 1-\varphi_\tau(x^a,z^t)/(1-\delta) \;, \tag{65}$$

$$W = \sum_{\tau=0}^{a-1} w_\tau(x^a,z^t)L_\tau \;,\;\; L = \sum_{\tau=0}^{a-1} L_\tau, \tag{66}$$

$$\rho(z^t) = \frac{m}{\lambda(m)}(1-\delta)\beta E_{x^a,z^t}\Big[W(\mathcal{C}_a, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big] \text{ if } m>0 \;. \tag{67}$$

The last condition is the workers' participation constraint; it specifies the minimum expected utility that contract $\mathcal{C}_a$ must promise in order to attract a worker queue of length $\lambda(m)$ per vacancy.

Definition: *A competitive search equilibrium is a list*

$$\Big[U(z^t), \rho(z^t), \mathcal{C}_a(x^a,z^t), m(x^a,z^t), V(x^a,z^t), \delta(x^a,z^t), J_a(.), L_\tau(x^a,z^t), N(x^a,z^t), N_0(z^t)\Big],$$

for all $t\geq 0$, $a\geq 0$, $x^a\in X^{a+1}$, $z^t\in Z^{t+1}$, $0\leq\tau\leq a$, and for a given initial firm distribution, such that

(a) Firms' exit, hiring and layoff strategies are optimal. That is, J_a is the value function and $\mathcal{C}_a(.)$, $\delta(.)$, $m(.)$, and $V(.)$ are the policy functions for problem (63)-(67).

(b) Employment evolves according to

$$\begin{aligned} L_\tau(x^a,z^t) &= L_\tau(x^{a-1},z^{t-1})\frac{\varphi_\tau(x^a,z^t)}{1-\delta(x^a,z^t)} \;,\;\; 0\leq\tau\leq a-1 \;, \\ L_a(x^a,z^t) &= m(x^a,z^t)V(x^a,z^t) \;,\;\; a\geq 0 \;. \end{aligned}$$

(c) Firm entry is optimal. That is, the complementary slackness condition

$$\sum_x \pi_0(x)J_0(x,z^t) \leq K(z_t) \;,\;\; N_0(z^t)\geq 0 \;, \tag{68}$$

holds for all z^t, and the number of firms evolves according to (11) and (13).

(d) Workers' search strategies are optimal, i.e. (ρ, U) satisfy equations (61) and (62).

(e) Aggregate resource feasibility; for all z^t,

$$\sum_{a\geq 0, x^a} N(x^a,z^t)\Big[\lambda(m(x^a,z^t))V(x^a,z^t) + \sum_{\tau=0}^{a-1} L_\tau(x^a,z^t)\Big] = 1 \;. \tag{69}$$

Proposition 5: *A competitive search equilibrium is socially optimal.*
Proof: The proof proceeds in two steps. First, substitute the participation constraint (67) into the firm's problem and make use of the contracts' recursive equations (60) to

show that the firms' recursive profit maximization problem is identical to the maximization of the social surplus of a firm. Second, show that a competitive search equilibrium is socially optimal.

First, define the social surplus of a firm with history (x^a, z^t) and with predetermined contracts and employment levels as follows:

$$G_a\Big[(\mathcal{C}_\tau)_{\tau=0}^{a-1}, (L_\tau)_{\tau=0}^{a-1}, x^a, z^t\Big] \equiv J_a\Big[(\mathcal{C}_\tau)_{\tau=0}^{a-1}, (L_\tau)_{\tau=0}^{a-1}, x^a, z^t\Big] + \sum_{\tau=0}^{a-1} L_\tau \Big[W(\mathcal{C}_\tau, x^a, z^t) - U(z^t)\Big] . \tag{70}$$

Using (60) and (62), the worker surplus satisfies

$$W(\mathcal{C}_\tau, x^a, z^t) - U(z^t) = w_\tau(x^a, z^t) - b - \rho(z^t) + \beta\varphi_\tau(x^a, z^t) E_{x^a,z^t}\Big[W(\mathcal{C}_\tau, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big] .$$

Now substitute this equation and (63) into (70), and write

$$\sigma \equiv \Big[(\mathcal{C}_\tau)_{\tau=0}^{a-1}, (L_\tau)_{\tau=0}^{a-1}, x^a, z^t\Big] \text{ and } \sigma_+ \equiv \Big[(\mathcal{C}_\tau)_{\tau=0}^{a}, (L_{\tau+})_{\tau=0}^{a}, x^{a+1}, z^{t+1}\Big] ,$$

with $L_{\tau+}$ as defined in (64) and $L = \sum_{\tau=0}^{a-1} L_\tau$, to obtain

$$\begin{aligned}
G_a(\sigma) &= \max_{\delta,m,V,\mathcal{C}_a} \Big\{ x_a z_t F(L) - C(V, L, x_a z_t) - f - \sum_{\tau=0}^{a-1} L_\tau w_\tau(x^a, z^t) \\
&\quad + \beta(1-\delta) E_{x^a,z^t} J_{a+1}(\sigma_+) \Big\} + \sum_{\tau=0}^{a-1} L_\tau \Big[w_\tau(x^a, z^t) - b - \rho(z^t) \\
&\quad + \beta\varphi_\tau(x^a, z^t) E_{x^a,z^t}\Big[W(\mathcal{C}_\tau, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big]\Big] \\
&= \max_{\delta,m,V,\mathcal{C}_a} \Big\{ x_a z_t F(L) - [b + \rho(z^t)]L - f - C(V, L, x_a z_t) + \beta(1-\delta) E_{x^a,z^t} J_{a+1}(\sigma_+) \\
&\quad + \beta \sum_{\tau=0}^{a-1} L_\tau \varphi_\tau(x^a, z^t) E_{x^a,z^t}\Big[W(\mathcal{C}_\tau, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big] \Big\} \\
&= \max_{\delta,m,V,\mathcal{C}_a} \Big\{ x_a z_t F(L) - bL - \rho(z^t)[L + \lambda(m)V] - f - C(V, L, x_a z_t) \\
&\quad + \beta(1-\delta) E_{x^a,z^t} J_{a+1}(\sigma_+) \\
&\quad + \beta(1-\delta) \sum_{\tau=0}^{a} L_{\tau_+} E_{x^a,z^t}\Big[W(\mathcal{C}_\tau, x^{a+1}, z^{t+1}) - U(z^{t+1})\Big] \Big\} \\
&= \max_{\delta,m,V,\mathcal{C}_a} \Big\{ x_a z_t F(L) - bL - \rho(z^t)[L + \lambda(m)V] - f
\end{aligned} \tag{71}$$

$$-C(V,L,x_a z_t) + \beta(1-\delta)E_{x^a,z^t}G_{a+1}(\sigma_+)\Bigg\} \ .$$

Here maximization is always subject to (64) and (65), the third equation makes use of

$$(1-\delta)L_{\tau+} = \varphi_\tau(x^a,z^t)L_\tau \ ,$$

for $\tau \leq a-1$, and

$$\rho(z^t)\lambda(m)V = \beta(1-\delta)L_{a+}E_{x^a,z^t}\Big[W(\mathcal{C}_a,x^{a+1},z^{t+1}) - U(z^{t+1})\Big] \ ,$$

and the last equation makes use of (70) for G_{a+1}. This shows that the firm solves a surplus maximization problem which is identical to the one of the planner specified in (57) provided that $\rho(z^t) = \mu(z^t)$ holds for all z^t, where μ is the social value of an unemployed worker as defined in the proof of Proposition 4. The only difference between the two problems is that the firm commits to cohort-specific separation probabilities, whereas the planner chooses in every period an identical separation probability for all workers (and he clearly has no reason to do otherwise). Nonetheless, both problems have the same solution: they are dynamic optimization problems of a single decision maker in which payoff functions are the same and the decision sets are the same. Further, time inconsistency is not an issue since there is no strategic interaction and since discounting is exponential. Hence solutions to the two problems, with respect to firm exit, layoffs and hiring strategies, are identical. In both problems the decision maker could discriminate between different cohorts in principal. Because such differential treatment does not raise social firm value, there is also no reason for competitive search to produce such an outcome. Nonetheless, there can be equilibrium allocations where different cohorts have different separation probabilities, but these equilibria must also be socially optimal because they maximize social firm value.

It remains to verify that competitive search gives indeed rise to socially efficient firm entry. When $\mu(z^t) = \rho(z^t)$, $G_0(x,z^t) = J_0(x,z^t)$ as defined in (70) coincides with $G_0(0,x,z^t)$, as defined in (57). Hence, the free-entry condition (68) coincides with the condition for socially optimal firm entry (58). Because of aggregate resource feasibility (69), the planner's resource constraint (15) is also satisfied. Since the allocation of a competitive search equilibrium satisfies all the requirements of Lemma 4(b), it is socially optimal. □

Appendix C: Calibration details

We choose the period length to be one week and set $\beta = 0.999$ so that the annual interest rate is about 5 percent. We assume a CES matching function $m(\lambda) = (1 + k\lambda^{-r})^{-1/r}$ (i.e. the inverse of the function $\lambda(m)$ used in the main text) and set the two parameters k and r to target a weekly job-finding rate of 0.129 and an elasticity of the job-finding rate with respect to the vacancy-unemployment ratio of 0.28 (Shimer (2005)).[34] Below we also target the (average) weekly job-filling rate at 0.3, which corresponds to a monthly vacancy yield of 1.3 (Davis et al. (2013)). Since in steady state the unemployment-vacancy ratio equals the ratio between the job-filling rate and the job-finding rate, we calculate the parameters k and r to attain the two targets at $\lambda = 0.3/0.129 = 2.326$.

The production technology is Cobb-Douglas with xL^{α} where the firm's idiosyncratic productivity $x = x_0 x_1$ contains a time-invariant component x_0 and a transitory component x_1 (cf. Elsby and Michaels (2013)). The time-invariant component is drawn upon firm entry from one of three values x_0^i, $i = 1, 2, 3$, with entry shares σ^i where (x_0^i, σ^i) are chosen to match the firm and employment shares within the three size classes 1-49, 50-499 and ≥ 500. The transitory component x_1 is drawn from one of five equidistant values in the range $[1 - \overline{x}, 1 + \overline{x}]$ and is redrawn every period with probability π. Parameters π and $\overline{x}$ are chosen to match a monthly separation rate of 4.2 percent and the observation that about two thirds of employment is at firms with monthly employment growth rates in the range $[-0.02, 0.02]$ (see Davis et al. (2010)). Firm exit is exogenous; that is, we set the operating cost to $f = 0$ and choose exit probabilities specific for the three firm types δ^i, $i = 1, 2, 3$, to match job losses at closing firms for the three size classes. Parameter α is set to 0.7 which gives rise to a labor share of roughly 2/3.[35]

We choose unemployment income b at 97.7 percent of mean wage earnings. This value corresponds to the parameter value chosen by Hagedorn and Manovskii (2008) which ensures that reasonably small aggregate productivity shocks have quantitatively significant labor market responses. This value of b corresponds to 68 percent of labor productivity and to 96.8 percent of the average (employment-weighted) marginal product. We explore robustness to a much lower value of b in Appendix D. The exogenous quit rate is set at $s_0 = 0.0048$ to match a monthly quit rate of 2 percent. The entry cost parameter K can be normalized arbitrarily since all firm value functions (and thus the free-entry condition) are linearly homogeneous in the vector (x, b, c, K).

As mentioned in the main text, the recruitment technology has the form $c(V) = \frac{c}{1+\gamma}(\frac{V}{L})^{\gamma}V$, where we take a cubic function ($\gamma = 2$) for the benchmark calibration. When we compare the benchmark results with those for $\gamma = 0.1$ and for $\gamma = 8$, we recalibrate parameters c and b (or K) to target the average unemployment-vacancy ratio $\lambda = 2.326$ which gives rise to an average weekly job-filling rate of 0.3 and the

[34] Note that there is no third parameter in the CES matching function since we require that $\lim_{\lambda \to \infty} m(\lambda) = 1$.

[35] Given that all capital is fixed at the level of a firm, this calculation of factor shares ignores capital income accruing from variable capital investment which would suggest a higher value of α. For a robustness analysis, see Appendix D.

same b/w ratio as in the benchmark.[36] We note that recruitment costs per hire are reasonably low for all three parameterizations (below 1% of quarterly earnings). Table 4 summarizes the parameter choices for the benchmark calibration.

Table 4: Parameter choices in the benchmark calibration ($\gamma = 2$).

Parameter	Value	Description	Explanation
β	0.999	Discount factor	Annual interest rate 5%
k	6.276	Matching fct. scale	weekly job-finding rate 0.129
r	1.057	Matching fct. elasticity	0.28 (Shimer (2005))
α	0.7	Prod. fct. elasticity	Labor share
c	0.409	Vacancy cost parameter	weekly job-filling rate 0.3
(x_0^i)	$(.274, .621, 1.488)$	permanent productivity	employment shares (3 size classes)
(σ^i)	$(99.2, .765, .035)\%$	share at entry	firm shares (3 size classes)
(δ^i)	$(2.24, .25, .03)$‰	exit rates	job losses at exiting firms
$\overline{x}$	0.11	Productivity range	monthly separation rate 4.2%
π	0.06	Adjustment prob.	2/3 of employment in firms with employment growth in [-0.02,0.02]
b	0.1	unemployment income	97.7% of wage income
K	205.0	Entry cost	Arbitrary normalization
s_0	0.48%	Quit rate	Monthly quit rate 2%

[36] Deviating from Table 4, we set $c = 0.00295$, $K = 208.64$ for $\gamma = 0.1$ and $c = 1.84 \cdot 10^7$, $K = 186.92$ for $\gamma = 8$ (fixing $b = 0.1$ throughout).

Appendix D: Robustness

We explore the robustness of the calibration exercise regarding different parameter choices for unemployment income b and for the returns-to-scale parameter α. Departing from the benchmark calibration with cubic vacancy costs we consider two variations. First, we consider the alternative of setting unemployment income to 70 percent of average wages (46% of labor productivity), instead of 97.7 percent as in the benchmark. Second, relative to the benchmark with $\alpha = 0.7$ which gives rise to a plausible labor share (with fixed capital at any individual firm) we consider the alternative of $\alpha = 0.95$ which is more in line with a model where capital can be adjusted at the firm level. In both variations, parameters c, $\overline{x}$ and (x_0^i) are readjusted so that the model hits the same calibration targets as in the benchmark calibration.
Figure 4 shows that the cross-sectional behavior of vacancy rates, vacancy yields, hires rates and layoff rates is almost unchanged relative to the benchmark calibration. That is, irrespective of the parameter values for b and α, the model with cubic vacancy costs explains more than half of the cross-sectional variation in vacancy yields for firm growth rates below 20 percent, although the curves flatten out at firm growth above 20 percent relative to the benchmark calibration (blue/solid curve).
Figure 5 shows the impulse responses to a one-percent increase in aggregate productivity. Here the two variations exhibit markedly different patterns, but this is little surprising. First, the model with a lower value of unemployment income clearly generates less amplification (red/dashed curves), which is in line with the well-known finding of Shimer (2005) that search and matching models with high match surplus generate too little labor market volatility. The propagation of the shock is similar to the benchmark, however. In the model with a higher returns-to-scale parameter (green/dotted curves), the productivity increase generates larger (and hump-shaped) responses of the job-finding rate and of the vacancy-unemployment ratio, although this does not imply that the model exhibits more labor market amplification since the output response is stronger as well.

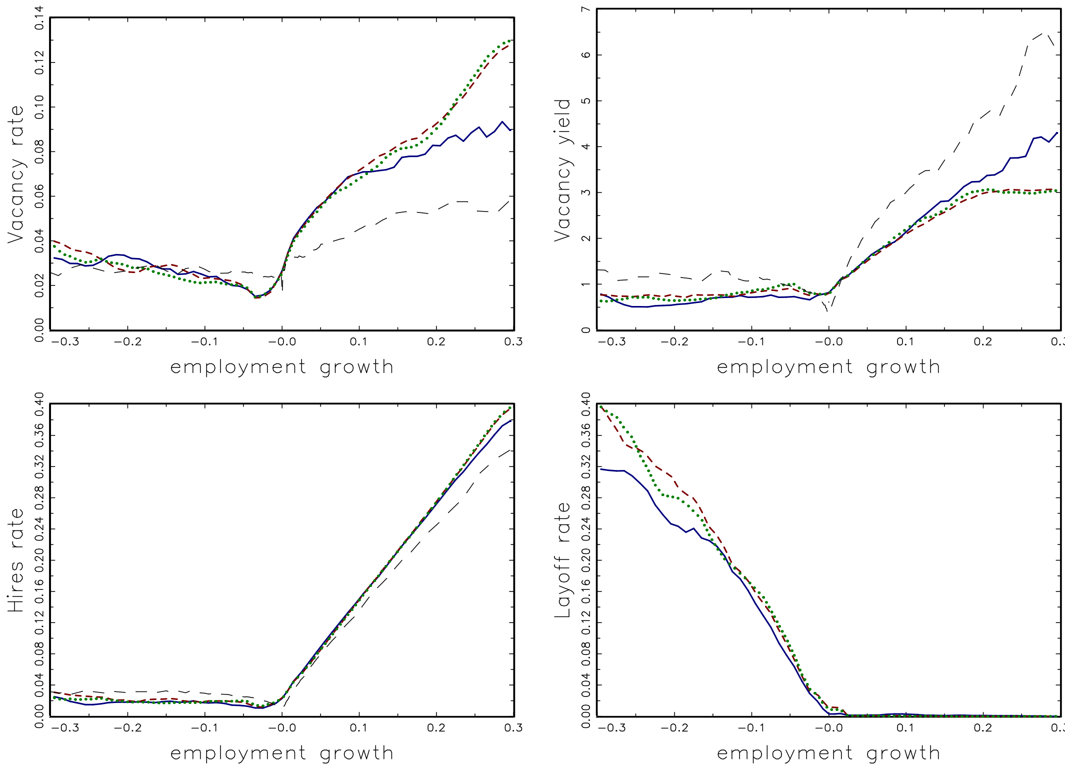

Figure 4: Cross-sectional relationships between monthly employment growth and the vacancy rate, the vacancy yield, the hires rate and the layoff rate. The dashed curves (in the first three graphs) are from the data used in Davis et al. (2013), the blue (solid) curves are for the benchmark parameterization ($b/w \approx 0.977$, $\alpha = 0.7$), the red (closely dashed) curves are for the calibration with $b/w \approx 0.7$ and the green (dotted) curves are for $\alpha = 0.95$.

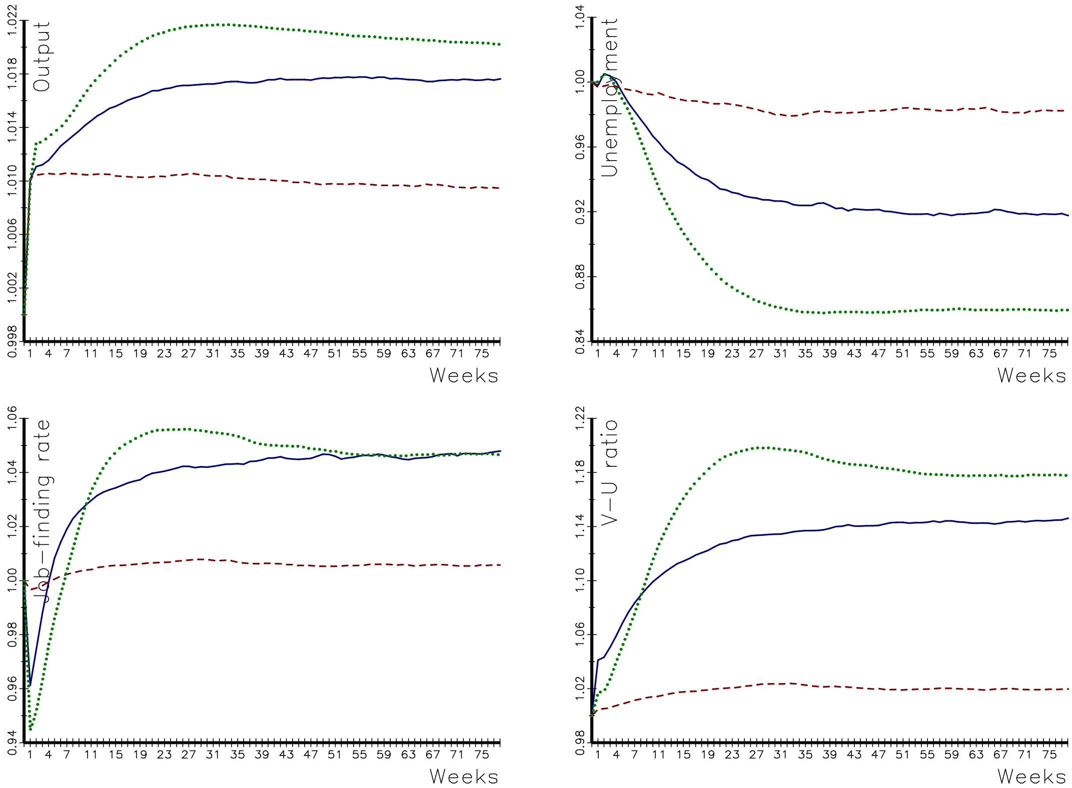

Figure 5: Impulse response to a permanent 1% increase in aggregate productivity. The blue (solid) curves are for the benchmark parameterization, the red (closely dashed) curves are for $b/w = 0.7$ and the green (dotted) curves are for $\alpha = 0.95$.

Appendix E: Impulse response in a VAR model

We borrow the methodology for constructing the impulse responses in Figure 3 straight from Fujita and Ramey (2007) - except for the details discussed below.[37] We use data from 1951:Q1 to 2011:Q4. The data is real quarterly GDP from FRED; the number of vacancies is from the Help Wanted Index from Barnichon's Composite Help-Wanted Index series;[38] the number of unemployed is from the CPS; employment is from the BLS total payroll employment; and population is from the BLS. The job-finding rate is then calculated in the same way as in Elsby, Michaels, Solon (2009).[39] All data other than GDP are averaged over their monthly (seasonally adjusted) observations to obtain quarterly series. They are then logged and detrended by regressing each on a cubic polynomial in time.

To generate impulse responses of output, employment, labor market tightness and the job-finding rate to a permanent rise in productivity, we first identify *exogenous* productivity deviations in the data series and look at how the variables of interest respond to these. Let

$$\begin{aligned} p_t &\equiv \text{observed (detrended) output per worker ,} \\ \theta_t &\equiv \text{observed (detrended) vacancy-unemployment ratio ,} \\ e_t &\equiv \text{observed (detrended) employment-population ratio ,} \\ \phi_t &\equiv \text{observed (detrended) job-finding rate ,} \end{aligned}$$

and let z_t be the unobserved exogenous productivity deviation. To identify z_t, we first estimate (by OLS) the following system:

$$\begin{aligned} \ln p_t &= \left[\begin{array}{cccc} \ln p_t & \ln \theta_t & \ln e_t & \ln \phi_t \end{array} \right] \left[\begin{array}{ccc} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \end{array} \right] \left[\begin{array}{c} L \\ L^2 \\ L^3 \end{array} \right] + \varepsilon_t^p \\ &= \left[\begin{array}{cccc} \ln p_t & \ln \theta_t & \ln e_t & \ln \phi_t \end{array} \right] A(L) + \varepsilon_t^p \ , \end{aligned}$$

where L is the lag operator. Given this estimation, $\left\{ \hat{A}_{ij} \right\}$, we follow Fujita and Ramey (2007) by assuming that the exogenous productivity deviations, $\ln z_t$, can be identified by

$$\begin{aligned} \ln p_{t<0} &= \ln \theta_{t<0} = \ln e_{t<0} = \ln \phi_{t<0} = \ln z_{t<0} = 0 \ , \\ \hat{\varepsilon}_t^p &= \ln p_t - \left[\begin{array}{cccc} \ln p_t & \ln \theta_t & \ln e_t & \ln \phi_t \end{array} \right] \hat{A}(L) \ , \\ \ln z_t &= \hat{A}_{11} \ln z_{t-1} + \hat{A}_{12} \ln z_{t-2} + \hat{A}_{13} \ln z_{t-3} + \hat{\varepsilon}_t^p \ . \end{aligned}$$

[37] We are grateful to David Ratner for providing an initial code.

[38] R. Barnichon, "Building a Composite Help-Wanted Index", *Economics Letters*, Vol. 109, 175–178, 2010.

[39] M. Elsby, R. Michaels and G. Solon, "The Ins and Outs of Cyclical Unemployment", *American Economic Journal: Macroeconomics*, Vol. 1, 84–110, 2009.

Once a series for $\ln z_t$ has been identified from the data in this way, an $AR\,(3)$ process for $\ln z_t$ can be estimated,

$$\begin{aligned}\ln z_t &= C_{01}\ln z_{t-1} + C_{02}\ln z_{t-2} + C_{03}\ln z_{t-3} + \varepsilon_t^z \\ &= C_0\,(L)\ln z_t + \varepsilon_t^z\ ,\end{aligned}$$

and the relationship between the endogenous variables $\ln e_t, \ln\theta_t,\ \ln\phi_t$ and the exogenous process $\ln z_t$ can be calculated by estimating the following relationships (by OLS):

$$\begin{aligned}\ln e_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix}\begin{bmatrix} B_{111} & B_{112} & B_{113} \\ B_{121} & B_{122} & B_{123} \\ B_{131} & B_{132} & B_{133}\end{bmatrix}\begin{bmatrix} L \\ L^2 \\ L^3 \end{bmatrix} + C_1\,(L)\ln z_t + D_1\hat{\varepsilon}_t^p + \varepsilon_t^e \\ &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix} B_1\,(L) + C_1\,(L)\ln z_t + D_1\hat{\varepsilon}_t^p + \varepsilon_t^e\ ,\end{aligned}$$

$$\begin{aligned}\ln\theta_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix} B_2\,(L) + C_2\,(L)\ln z_t + D_2\hat{\varepsilon}_t^p + \varepsilon_t^\theta\ , \\ \ln\phi_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix} B_3\,(L) + C_3\,(L)\ln z_t + D_3\hat{\varepsilon}_t^p + \varepsilon_t^\phi\ .\end{aligned}$$

The impulse-response functions to a permanent increase in exogenous productivity of 1% are then constructed by simulating these estimated relationships forward:

$$\begin{aligned}\ln z_{t<0} &= \ln e_{t<0} = \ln\theta_{t<0} = \ln\phi_{t<0} = 0\ , \\ \ln z_{t\geq 0} &= 0.01\ , \\ \hat{\varepsilon}_{t\geq 0}^p &= \hat{\varepsilon}_{t\geq 0}^z = \ln z_{t\geq 0} - \hat{C}_0\,(L)\ln z_{t\geq 0}\ , \\ \hat{\varepsilon}_t^e &= \hat{\varepsilon}_t^\theta = \hat{\varepsilon}_t^\phi = 0\ ,\end{aligned}$$

$$\begin{aligned}\ln e_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix}\hat{B}_1\,(L) + \hat{C}_1\,(L)\ln z_t + \hat{D}_1\hat{\varepsilon}_t^p\ , \\ \ln\theta_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix}\hat{B}_2\,(L) + \hat{C}_2\,(L)\ln z_t + \hat{D}_2\hat{\varepsilon}_t^p\ , \\ \ln\phi_t &= \begin{bmatrix} \ln e_t & \ln\theta_t & \ln\phi_t \end{bmatrix}\hat{B}_3\,(L) + \hat{C}_3\,(L)\ln z_t + D_3\hat{\varepsilon}_t^p\ ,\end{aligned}$$

$$\ln p_t = \begin{bmatrix} \ln p_t & \ln\theta_t & \ln e_t & \ln\phi_t \end{bmatrix}\hat{A}\,(L) + \hat{\varepsilon}_t^p\ .$$

Our construction differs from Fujita and Ramey (2007) only in that their estimations are based on data to 2005, they use three variables (p_t, θ_t, e_t) for the VAR, and they show impulse responses for a one-time rather than a permanent shock. We replicated their settings and find their results, and we checked that adding the fourth variable does not qualitatively change the outcome for the three initial variables in their methodology.

Appendix F: Policy application

In this appendix we provide some first exploration of the positive implications of policy interventions in our environment. We focus on hiring subsidies (hiring credits), as these have been extensively deployed to stimulate job growth in past recessions and have received renewed attention during the Great Recession.[40] Indeed, it is conceivable that they succeed in stabilizing business cycle fluctuations, especially when they are used in a counter-cyclical way. However, we find that the contrary is the case. We compare time-invariant and counter-cyclical subsidies, financed by lump-sum taxes. We solve the model as the solution to a quasi-planner's problem who maximizes social welfare subject to given government policy.[41] We set the subsidy per hire to 0.03 which corresponds to 8% of a monthly wage so that government expenditures on hiring subsidies are 0.3 percent of GDP. With a counter-cyclical policy, hiring firms receive the subsidy only when the aggregate productivity state is below its mean. Table 5 shows the outcome of this exercise. While both policies succeed in stabilizing the job-finding rate to some extent, they dramatically increase the volatility of separations and unemployment.[42] Perhaps surprisingly, these destabilizing forces are stronger for the counter-cyclical policy where low-productivity firms lay off even more workers during recessions and fewer workers during booms. These findings suggest that hiring subsidies are not particularly useful to stabilize the cycle, at least when they are not accompanied by additional policies aiming to dampen separations during recessions. More work on these issues will obviously be needed to explore the impact of such policies in broader environments.

[40]The Hiring Incentives to Restore Employment Act (HIRE) of 2010 includes tax exemptions from employer social security contributions and business income tax breaks for workers hired from unemployment; hiring credits were also an element of the American Jobs Act proposed by the Obama administration in 2011. For an overview, see D. Neumark, "Spurring Job Creation in Response to Severe Recessions: Reconsidering Hiring Credits", *Journal of Policy Analysis and Management*, Vol. 32, 142–171, 2013.

[41]Cf. M. Veracierto, "Firing Costs and Business Cycle Fluctuations", *International Economic Review*, Vol. 49, 1–39, 2008.

[42]In steady state, both the separation and the hiring rate increase by 50 percent in response to the (time-invariant) policy. This results in much more volatile firm dynamics with substantially more worker reallocation between employment and unemployment, which ultimately also increases the steady-state unemployment rate. We also find that the two hiring margins do not react equally to the policy: firms hire more workers by using more vacancy postings, while the aggregate vacancy yield increases only slightly.

Table 5: Business cycle effects of hiring subsidies

	Laissez faire	Constant policy	Cyclical policy
Unemployment	10.3	17.7	20.8
Vacancies	7.9	4.0	12.6
Output	1.5	2.9	2.7
Job–finding rate	6.2	4.8	4.8
Separation rate	5.6	17.5	22.7

Notes: The table reports the standard deviations of logged and HP filtered (parameter 1600) quarterly variables, where model statistics are obtained as in Table 3.