Selective Hiring and Welfare Analysis in Labor Market Models

Christian Merkl
Friedrich-Alexander-University Erlangen-Nuremberg, CESifo and IZA

Thijs van Rens
University of Warwick, Centre for Macroeconomics, IZA and CEPR

October 2018*

Abstract

Firms select not only how many, but also which workers to hire. Yet, in most labor market models all workers have the same probability of being hired. We argue that selective hiring crucially affects welfare analysis. We set up a model that is isomorphic to a search model under random hiring but allows for selective hiring. With selective hiring, the positive predictions of the model change very little, but implications for welfare are different for two reasons. First, a hiring externality occurs with random but not with selective hiring. Second, the welfare costs of unemployment are much larger with selective hiring, because unemployment risk is distributed unequally across workers.

Keywords: labor market models, welfare, optimal unemployment insurance
JEL codes: E24, J65

*We are grateful to Régis Barnichon, Giuseppe Bertola, Wouter den Haan, Alberto Martin, and Giacomo Ponzetto for valuable suggestions; and to the institutional network 'Ensuring Economic and Employment Stability' for financial support. Thijs van Rens gratefully acknowledges the hospitality of the UNSW School of Economics during the final stages of preparing this paper.

1 Introduction

Standard search and matching models of the labor market often assume, for tractability, that workers are homogeneous and markets are complete. Combined, these assumptions eliminate an important source of welfare costs of unemployment: the fact that unemployment is unequally distributed across workers. In studies doing welfare analysis, the focus has been on relaxing the complete markets assumption. If markets are incomplete, ex-ante identical workers cannot share unemployment risk, making these workers heterogeneous ex post. Because one needs to keep track of the entire distribution of asset holdings, this class of models is difficult to solve. Moreover, because workers are ex ante homogeneous, welfare costs of unemployment are small (Krusell and Smith 1998).

We propose a framework, in which workers are ex-ante heterogeneous, while maintaining the assumption that markets are complete. In this model, some workers are more attractive to employers than others because they have lower training costs.[1] Individual-specific training costs are fully observable to workers and firms and determine how likely it is that a worker finds a job in a given period. We analyze two polar cases in this framework. If individual training costs are transitory or fully match-specific, each worker has the same probability of being hired in future periods in expectation. We call this version of the model the case of perfectly random hiring. If individual training costs are permanent or fully worker-specific, the probability of being hired in the future depends on current training costs, which will be the same in future periods. This is the case of perfectly selective hiring.[2]

If training costs have both a transitory (match-specific) and a permanent (worker-specific) component, then hiring is partly random and partly selective. There is ample evidence that hiring decision in the real world are partly selective, and not all workers have the same probability of finding a job. We discuss some of this evidence in section 2. By analyzing the polar cases of perfectly random and perfectly selective hiring, we aim to understand how selectivity in hiring decision matters for our understanding of the labor market.

Our model is set up in such a way that it is isomorphic to a standard search and matching model (Pissarides 2000, chapter 1) in terms of its predictions for labor market dynamics. Specifically, the model can be parameterized to generate the same aggregate job-finding and unemployment rates, and the same elasticities of these variables with respect to changes in productivity. The distribution of idiosyncratic training costs plays the same role as the aggregate matching function in the standard model. Thus, we provide a framework that on the one hand maintains most of the insights from standard labor market models, and on the other hand allows us to compare the predictions of the model under selective versus random hiring. While the predictions of the model for aggregate variables are identical under selective hiring, the implications for inequality and welfare are very different.

If hiring is selective, unemployment is costly because unemployment risk is spread unequally across workers. With perfectly selective hiring, some workers are always employed, while others are always unemployed. With partially selective hiring all workers are employ-

[1] Training costs are a convenient way to introduce heterogeneity in the context of our model. However, other sources of heterogeneity, in particular heterogeneity in productivity, have very similar implications, see section 7 and appendix A.3.

[2] This concept of selectivity in hiring is similar to Berger (2016), although in that paper firms are selective in their firing rather than hiring decisions, firing bad workers and maintaining good ones.

able, but some more so than others. Thus, unemployment risk is uninsurable and the welfare costs of unemployment are much larger than under random hiring. This effect is counteracted by a hiring externality, which leads to overhiring with random but not with selective hiring. Quantitatively, the distributional effect dominates the hiring externality. As a result, there is a role for government intervention, insuring (unborn) workers against their unemployment risk.

As an application of our framework, we study the question of the optimal level of unemployment insurance. Under random hiring, the government can replicate the efficient allocation using unemployment benefits and lump-sum taxes. In this case, unemployment benefits are set to make sure the level of job creation is efficient. Under selective hiring there is an additional motive for unemployment insurance because workers cannot self-insure against their characteristics, which determine their individual-specific unemployment risk. Thus, the government faces a trade-off between efficient job creation and efficient redistribution. We solve the Ramsey problem for the government in this case and find two results. First, the maximum welfare that can be reached under selective hiring is substantially lower than under random hiring. Second, to obtain a more equitable income distribution with selective hiring, it may be optimal to set unemployment benefits substantially higher than under random hiring.

To our knowledge, our paper is one of the first to analyze the role of ex-ante heterogeneity on optimal unemployment insurance. In a related recent contribution, Lifschitz et al. (2017) propose a model with heterogeneity in productivities, separation rates and costs of recruiting for different education groups. They show that the optimal replacement rate increases substantially with ex-ante heterogeneity, which is in line with our results, although based on a completely different framework.

The basic trade-off emphasized in the literature on optimal unemployment insurance, is that unemployment benefits insure risk-averse workers against variations in their income and consumption, but discourage search effort (Baily 1978, Chetty 2006). We contribute to this literature by pointing out that with selective hiring of heterogeneous workers, the insurance motive is (much) larger than with random hiring because unemployment risk is higher for workers with low income and high marginal utility from consumption. Previous studies have pointed out other reasons why the insurance motive may be more important, for example because it allows workers to look for high-wage jobs with high unemployment risk (Acemoglu and Shimer 1999) or because credit constraints prevent workers from self-insuring against cyclical unemployment risk (Landais, Michaillat and Saez 2016, Moyen and Stähler 2014, Mitman and Rabinovich 2015). Depending on the degree of selectivity in hiring, the effect of ex-ante heterogeneity may be much stronger than these alternative mechanisms. It seems likely that ex-ante heterogeneity also has implications for how optimal unemployment insurance depends on the business cycle (Landais et al. 2016) or on unemployment duration (Hopenhayn and Nicolini 1997, 2009, Fredriksson and Holmlund 2006). However, we leave this interesting question for future research.

This paper is only tangentially related to other studies using labor market models with worker heterogeneity. A large literature, starting with Becker (1973), studies under what conditions there is positive assortative matching between heterogeneous workers and firms.

But the models in this literature are not used for welfare analysis. Directed search models, as in Moen (1997), provide a description of the coordination friction that may underlie the aggregate matching function and give rise to ex-post heterogeneity. But these models generate heterogeneity as an equilibrium outcome and maintain the assumption that workers are ex-ante homogeneous. An exception is Fernández-Blanco and Preugschat (2018), who model an economy with directed search and worker heterogeneity, which gives rise to duration dependence in job-finding probabilities as in our model. However, the focus of their paper is entirely different from ours. Similarly, Shimer's (2007) model of mismatch unemployment can be thought of as a micro-foundation for an aggregate matching friction, which does not affect the predictions of the standard model in terms of welfare.

More recently, Chugh and Merkl (2016) and Epstein (2016) analyze how heterogeneity affects labor market dynamics. The conceptual framework in Chugh and Merkl, which models heterogeneity as shocks to match quality, is similar to the model we use in this paper, whereas Epstein models heterogeneity as workers having comparative advantages in particular jobs. Compared to both of these papers, we focus on the normative predictions of the model, which can be seen in steady state, as opposed to the predictions for business cycle fluctuations. Because of this difference in focus, we emphasize the importance of selective versus random hiring, whereas the heterogeneity in both Epstein (2016) and Chugh and Merkl (2016) gives rise to what we would call random hiring.

The remainder of this paper is structured as follows. Section 2 describes some evidence from microeconomic labor market data that hiring decisions of firms in the real world are partially selective. Section 3 sets up the model. In section 4, we derive the equilibrium job creation condition and establish the equivalence of our model with random hiring to a search model with an aggregate matching function. Section 5 deals with welfare analysis. We discuss the conditions for job creation to be efficient, and show how the welfare costs of unemployment differ starkly under selective versus random hiring due to the different distribution of consumption. Section 6 establishes the latter point quantitatively in an application to optimal unemployment insurance. Section 7 concludes.

2 Selective Hiring: Motivating Evidence

It is probably uncontroversial that hiring decisions are at least partially selective, but it may nevertheless be useful to start with a brief review of the evidence for this fact. This evidence consists of facts that have been documented in other contexts, but that have not always been interpreted as evidence for selective hiring.

The most direct evidence comes from the distribution of job-finding rates. If hiring is perfectly random, then all workers have the same probability of finding a job. If hiring is perfectly selective, then some, 'good' workers find jobs immediately, whereas other, 'bad' workers never find jobs. In the data, the job-finding rate decreases with unemployment duration, both in the US (Abraham and Shimer 2002) and in Europe (Wilke 2005).

A similar picture emerges when we compare the aggregate job-finding rate to the average unemployment duration. If all workers have the same job-finding rate, then the average unemployment duration D must simply be the inverse of the aggregate job-finding rate,

$D = 1/f$. If hiring is selective, then bad workers (with low job-finding probabilities) are over-represented in the average unemployment duration and under-represented in the average job-finding rate, so we would expect $D > 1/f$. In the data, unemployment duration is indeed much longer than expected based on the aggregate job-finding rate (Shimer 2012). By a similar argument, selective hiring may explain why the net job-finding rate, which excludes workers with unemployment duration shorter than the period of observation, is smaller than the gross job-finding rate (Shimer 2012).

The evidence for the duration dependence of individual job-finding rates is consistent with an endogenous scarring effect or loss of skill from unemployment spells as well as with ex-ante heterogeneity. However, both Hornstein (2012) and Barnichon and Figura (2015) argue that the data favor a selection story, in which workers with intrinsically lower job-finding rates are overrepresented in the unemployment pool.

Another piece of evidence comes from the composition of the pools of employed and unemployed workers. If hiring is selective, we would expect the quality of the employment pool to be countercyclical, because workers that are only hired in booms are relatively bad compared to workers that already had jobs in the recession. But by the same token, we would expect the quality of the unemployment pool to be countercyclical as well, because workers that are hired in booms are relatively good compared to workers that remain unemployed even in booms. It has long been known that there is a composition bias in the cyclicality of wages consistent with this story (Solon, Barsky and Parker 1994). Mueller (2017) recently documented that the average predicted wage of unemployed workers is countercyclical as well.

3 Model Environment

Our economy is populated by a continuum of worker-consumers i, characterized by ε_{it}. We model worker characteristics as training costs: a firm that hires worker i in period t needs to pay ε_{it} for this worker to become productive. Alternatively, we may think of ε_{it} as a measure of worker productivity or match-specific skills, which would lead to minor modifications to the model but would leave the results unchanged, see section 7 for a discussion and appendix A.3 for more details on the argument. Worker characteristics are fully observable to workers and firms, so that there is perfect information in the economy. In our model, training costs (or worker characteristics in general) determine how likely an individual worker is to be hired in a given period.

Let G and g denote the distribution function and the probability density function of training costs, $\varepsilon_{it} \sim G$. The distribution G is assumed to be constant across individuals and time-invariant.[3] This modelling framework is inspired by Brown, Merkl and Snower (2015), although both the focus of the analysis and the details of the model are very different from that paper.

Whether hiring is selective or random depends on the relative importance of transitory and permanent components of individual worker characteristics. If ε_{it} is fully transitory, i.e. if it is specific to a match rather than to an individual, then each worker expects to have the

[3] This does not mean, of course, that the average job finding rate is constant over time, since other factors than ε_{it} also affect the probability to be hired.

same probability of being hired in future periods, so that hiring decisions are independent of current worker characteristics and thus effectively random from today's perspective. If ε_{it} is permanent or individual-specific, i.e. $\varepsilon_{it} = \varepsilon_i$ is fixed for each worker over time, then current worker characteristics fully determine how likely an individual worker is to be hired in the future. This is what we call selective hiring. If ε_{it} includes both transitory and permanent components, then hiring is partly random and partly selective.

3.1 Preferences

Worker-consumers are infinitely-lived, have time-separable utility and care about the expected net present value of utility from consumption c_{it} and leisure. They may be employed or unemployed. Employed workers earn a wage w_{it}, which may depend on worker characteristics ε_{it}, and unemployed workers receive unemployment benefits b_t.

We assume the utility derived from leisure is zero, so that the flow utility $U(.)$ depends only on consumption. Then, workers' objective function is given by,

$$E_0 \sum_{t=0}^{\infty} \beta^t \mathcal{U}(c_{it}) \tag{1}$$

where β is the discount factor and E_0 denotes rational expectations in period 0.

3.2 Production and Job Creation

Employed workers hold jobs, which produce output y_t in each period, including the period in which the worker was hired and trained. Our assumption that worker characteristics take the form of training cost implies that the output of a job does not depend on the characteristics of the worker that holds it. Given a wage w_{it}, the firm's profits from a job equal $y_t - w_{it}$. The cost of creating a job is the cost of training a worker, which has a fixed component K and an idiosyncratic component ε_{it}. It is worth emphasizing that there are no search frictions in our model, so that jobs with positive value can be created immediately.

Since all jobs are identical after the worker has been trained, jobs created for workers with low training costs generate more output in net present value than jobs created for workers with high training costs. Thus, if it is efficient to create a job for a worker with training costs ε, then it must also be efficient to create a job for a worker with lower training costs $\varepsilon' < \varepsilon$. We further assume that profits are at least weakly decreasing in training costs, so that this property carries over to the equilibrium allocation as well.[4] This implies that in the efficient as well as in the equilibrium allocation of our model, there exists a unique cutoff level $\tilde{\varepsilon}_t$, such that a worker seeking a job is hired if $\varepsilon_{it} < \tilde{\varepsilon}_t$ and not hired if $\varepsilon_{it} > \tilde{\varepsilon}_t$. Although the existence of this hiring threshold is a property of the efficient allocation or equilibrium and not part of the environment, we will impose it below in order to simplify the notation.

[4] This is an assumption on the wage setting mechanism. Since wages may be decreasing in training costs, workers with lower training costs can capture some of the higher surplus their labor creates. The assumption is that this dependence of wages on training costs will not revert the effect of training costs on profits. Most reasonable wage setting mechanisms, and in particular Nash bargaining as we use in this paper, satisfy this assumption.

3.3 Markets

Worker-consumers and firms interact with each other on three types of markets. Firms hire workers on the labor market. The goods firms produce are sold to consumers on the goods market. Both firms and workers trade on asset markets.

On the labor market, firms hire unemployed workers, generating jobs and employed workers. The surplus generated by a job match may be strictly positive in this model, because of the heterogeneity in training costs. In this case, we assume that firm and worker Nash bargain over this surplus in order to agree on a wage, see section 5.2. There is an exogenous probability λ that a job is destroyed, in which case the worker becomes unemployed again. We assume full commitment of both worker and firm, so that regardless of the worker's ε_{it} both the firm and the worker must continue the job unless it is destroyed by a $\lambda-$shock and there is no endogenous job destruction. Let $f(\tilde{\varepsilon}_t)$ denote the aggregate job-finding rate, the probability that an average job seeker finds a job in each period. The aggregate job-finding rate depends on the hiring threshold defined in section 3.2 above: the higher the threshold, the larger the probability that any given job seeker is hired, everything else equal. Then, the number of employed workers in the economy evolves according to,

$$n_t = (1-\lambda) n_{t-1} + f(\tilde{\varepsilon}_t) s_t = (1-\lambda)(1-f(\tilde{\varepsilon}_t)) n_{t-1} + f(\tilde{\varepsilon}_t) \tag{2}$$

where s_t is the number of workers that seek a job in a given period. We assume job destruction happens before job creation, so that the number of workers that are seeking jobs equals the number of workers that are unemployed since last period, $1 - n_{t-1}$, plus the number of workers that were employed last period but lost their job in this period, λn_{t-1}. This timing is consistent with our assumption that there are no hiring frictions in this model: workers that are separated may be rehired immediately and do not have to wait until the next period. Notice that the number of job seekers $s_t = 1 - (1-\lambda) n_{t-1}$ does not equal the number of unemployed $u_t = 1 - n_t$, because some of the job seekers find new jobs within the period.

On the goods market, goods produced by firms are sold to workers for consumption. Goods market clearing requires that the amount of goods produced equals the amount of goods consumed by workers plus the amount of goods used to pay the training costs to create jobs. We assume that if firms make any profits in excess of the amount they need to pay the training costs, then these profits are distributed lump-sum to workers and then consumed. Thus, the aggregate resource constraint is given by,

$$\int_{-\infty}^{\infty} c_{it} dG = y_t n_t - [1 - (1-\lambda) n_{t-1}] f(\tilde{\varepsilon}_t)(K + H(\tilde{\varepsilon}_t)) \tag{3}$$

where $K + H(\tilde{\varepsilon}_t)$ denotes the average training cost of all workers that were hired in period t. The idiosyncratic component of the average training costs of new hires $H(\tilde{\varepsilon}_t)$ depends on the hiring threshold defined in section 3.2 above.

Asset markets are complete. The complete markets assumption allows workers to fully insure against idiosyncratic variations in their income over time. However, since all assets are in zero net supply, aggregate risk is not insurable. More importantly, since the unborn do not have access to asset markets, workers cannot insure against their characteristics in period 0.

3.4 Road Map

In the next two sections of the paper, we solve for the equilibrium allocation of the model, evaluate its properties and compare it to the efficient allocation. Section 4 describes the equilibrium allocation, and in particular the equilibrium job creation condition. Section 5 describes the efficient allocation, and derives conditions under which the level of job creation and the distribution of consumption resulting in equilibrium are efficient. This section, which contains the main result of the paper in its simplest form, shows that efficiency of the equilibrium depends crucially on whether worker characteristics are transitory or permanent.

4 Equilibrium Unemployment

In this section, we derive the equilibrium job creation condition and explore what are the aggregate job-finding rate and unemployment rate implied by the job creation condition. To do this, we need to specify whether hiring decisions are random or selective. We explore both versions of the model and show that the predictions of our model for job creation are very similar (and under some conditions identical) to the predictions of a standard Diamond-Mortensen-Pissarides model with search frictions.[5]

The objective of this section is to show that the predictions of our model about the cyclical behavior of the labor market are very similar to those of standard search models of the labor market, and that it makes very little difference for those predictions whether hiring is random or selective. Welfare analysis, however, differs sharply with the selectivity of hiring. We postpone this issue to section 5.

4.1 Job Creation

In the decentralized equilibrium, jobs are created when both firms and workers at least weakly prefer participating in a job match over their outside options. Since wages are set by Nash bargaining, firms and workers share the surplus that the match generates and always agree on whether a match is worth creating.

It is profit-maximizing, as we argued in section 3.2, to employ all workers with training costs ε_{it} below a threshold $\tilde{\varepsilon}_t$ and let workers with training costs above this threshold be unemployed. Imposing this 'threshold property' of the equilibrium, we can think of a representative firm choosing the hiring threshold $\tilde{\varepsilon}_t$, which determines the total number of workers they employ.

For the marginal hire, with training costs $K + \tilde{\varepsilon}_t$, the benefits of hiring this worker must exactly equal the training costs. These benefits equal the expected net present value of profits generated from a job. Thus, we get the following equilibrium job creation condition.

$$K + \tilde{\varepsilon}_t = E_t \sum_{\tau=0}^{\infty} (1-\lambda)^{\tau} Q_{t,t+\tau} \left(y_{t+\tau} - \tilde{w}_{t,t+\tau} \right) \tag{4}$$

[5] This is consistent with the interpretation in Pissarides (2000, p.4) that "the matching function summarizes a trading technology between heterogeneous agents that is also not made explicit."

where $\tilde{w}_{t,t+\tau}$ is the wage in period $t+\tau$ of the worker that is the marginal hire in period t, and $Q_{t,t+1}$ is the stochastic discount factor between periods t and $t+1$.

$$Q_{t,t+1} = \frac{\beta \mathcal{U}'(c_{t+1})}{\mathcal{U}'(c_t)} \tag{5}$$

For future reference, we also define $Q_{t,t+\tau} = Q_{t,t+1}Q_{t+1,t+2}...Q_{t+\tau-1,t+\tau}$ for $\tau \geq 1$ and $Q_{t,t+\tau} = 1$ for $\tau = 0$.

It is worth noting that, as long as the threshold property holds, only the wage of the marginal hire matters for job creation. The wages of all infra-marginal workers are purely redistributive and do not change the allocation.

4.2 Job-Finding Rate

We now have a condition for the hiring threshold $\tilde{\varepsilon}_t$ in equilibrium (4). The hiring threshold determines the aggregate job-finding rate and unemployment rate. In this section, we formalize this link.

The first, and most important, observation is that in our framework, unlike in standard labor market models with search frictions, the job-finding rate is not constant across workers. Since firms hire only workers with training costs below the hiring threshold $\tilde{\varepsilon}_t$, the job-finding probability of an individual worker f_{it} is either 1 or 0, depending on her training costs ε_{it}.

$$f_{it} = \begin{cases} 1 \text{ if } \varepsilon_{it} \leq \tilde{\varepsilon}_t \\ 0 \text{ if } \varepsilon_{it} > \tilde{\varepsilon}_t \end{cases} \tag{6}$$

The aggregate job-finding rate $f(\tilde{\varepsilon}_t)$ is then given by the average of the individual job-finding probabilities of all job seekers,

$$f(\tilde{\varepsilon}_t) = \frac{\int_{-\infty}^{\infty} f_{it} s_{it} dG}{\int_{-\infty}^{\infty} s_{it} dG} \tag{7}$$

where s_{it} is the fraction of type ε_{it} workers seeking a job. Notice that $f(\tilde{\varepsilon}_t)$ is the gross job-finding rate, which includes workers who lost their job in the current period.

4.3 Random Hiring

In order to evaluate the integrals in (7), we need to know in a given period t how many workers of each type ε_{it} are looking for a job.[6] The composition of the pool of job seekers depends crucially on whether invididual training costs are transitory or permanent. If training costs are i.i.d. over time (as well as across workers), then each worker gets a new draw for ε_{it} in each period, so that in any given period, the distribution of ε_{it} in the pool of job seekers mirrors the aggregate distribution G. In this case, the number of job seekers as a fraction of workers of each type equals the total number of job seekers as a fraction of the total labor force, $s_{it} = s_t$. In this case, the aggregate job-finding rate equals the probability that training

[6] We use the phrase "looking for a job" or "job seeker" loosely. With perfectly selective hiring, there are some workers who do not have a job and who have zero probability of being offered one, because their training costs are too high. We still include those workers in the pool of unemployed workers as well as job seekers, because at the current wage rate, they would accept a job if it were offered to them. If hiring were partially but not perfectly selective, these workers would have lower but not zero probability to find jobs.

costs are below the hiring threshold.

$$f^{\mathrm{RH}}\left(\tilde{\varepsilon}_t\right)=\frac{\int_{-\infty}^{\tilde{\varepsilon}_t} 1\cdot s_t\cdot dG+\int_{\tilde{\varepsilon}_t}^{\infty} 0\cdot s_t\cdot dG}{\int_{-\infty}^{\infty} s_t\cdot dG}=G\left(\tilde{\varepsilon}_t\right) \tag{8}$$

We refer to this case as perfectly random hiring, because at the beginning of the period, each unemployed worker has the same probability of getting a good draw for ε_{it} and therefore the same probability of finding a job, regardless of her current training costs. In other words, at the beginning of the period, it is random which workers will get hired and which ones will not.

Although we focus primarily on the job-finding rate, for completeness we also calculate the steady state unemployment rate. The steady state unemployment rate equals $\bar{u}_t = 1-\bar{n}_t$, where $\bar{n}_t$ is the steady state fraction of workers that are employed implied by difference equation (2). The steady state unemployment rate for the model with random hiring equals

$$\bar{u}^{\mathrm{RH}}=\frac{\lambda\left[1-G\left(\tilde{\varepsilon}\right)\right]}{\lambda\left[1-G\left(\tilde{\varepsilon}\right)\right]+G\left(\tilde{\varepsilon}\right)} \tag{9}$$

Notice that the number of unemployed workers does *not* equal the number of workers with training costs above the hiring threshold, because many of these workers are currently still employed because they were hired in the past, when they had lower training costs.

4.4 Comparison to Models with Search Frictions

We show that the job creation equation in our model is the same as in a standard search and matching model in the tradition of Diamond (1982), Mortensen (1982) and Pissarides (1985), if we choose the distribution of training costs G appropriately. The distribution function of worker heterogeneity plays the role of an aggregate matching function in search and matching models.

The job creation condition in a model with search frictions equates the expected net present value of firms' profits $y_t - w_t$ to the expected net present value of vacancy posting costs. Vacancy posting costs may include a fixed component K, which is paid only at the start of the vacancy, but also includes a flow cost k, the expected net present value of which depends on the probability the vacancy is filled in each period q_t.

$$K+\frac{k}{q_t}=E_t\sum_{\tau=0}^{\infty}\left(1-\lambda\right)^{\tau} Q_{t,t+\tau}\left(y_{t+\tau}-w_{t+\tau}\right) \tag{10}$$

See Pissarides (2009, section 5) for the role of fixed job creation costs in this type of model.

The job creation condition in the standard search model (10) equals the job creation condition in our model (4) if $E_t w_{t+\tau} = E_t \tilde{w}_{t,t+\tau}$ and $k/q_t = \tilde{\varepsilon}_t$. The first condition is satisfied in general for some wage setting mechanisms.[7] More importantly for the purposes of this paper, if we assume wages in both models are set by Nash bargaining, the condition holds true in steady state.

[7] For example, if all workers earn the same wage, so that $\tilde{w}_{t,t+\tau} = w_{t+\tau}$, as we assumed in a previous version of this paper, the condition is trivially holds.

The vacancy filling probability q_t in the search model depends on labor market tightness θ_t, the ratio of vacancies v_t over the number of unemployed workers u_t, through the matching technology, which relates new matches m_t to the number of unemployed and the number of vacancies. With a standard constant returns to scale Cobb-Douglas matching function, $m_t = u_t^{\mu} v_t^{1-\mu}$, we get $q_t = m_t/v_t = \theta_t^{-\mu}$. The job-finding rate in this model is also related to labor market tightness through the matching function, $f_t = m_t/u_t = \theta_t^{1-\mu}$. Thus, we can write the vacancy filling probability in terms of the job-finding rate, $q_t = \theta_t^{-\mu} = f_t^{-\mu/(1-\mu)}$, so that $k/q_t = kf_t^{\mu/(1-\mu)}$. Thus, the job creation condition in our model equals the one from the standard search model if $k/q_t = kf_t^{\mu/(1-\mu)} = \tilde{\varepsilon}_t$ or

$$f_t = \left(\frac{\tilde{\varepsilon}_t}{k}\right)^{\frac{1-\mu}{\mu}} \tag{11}$$

Comparing expression (11) to (8), it is clear that we can choose a distribution function G such that the job creation condition in our model under random hiring is the same as in the standard model. The distribution that makes the job creation conditions identical is $G(\varepsilon) = (\varepsilon/k)^{(1-\mu)/\mu}$ for $0 \leq \varepsilon \leq k$, which means that $1/\varepsilon_{it}$ follows a Pareto distribution.

Since the law of motion for employment (2) is the same in both models, identical job creation conditions and job-finding rates implies that the predictions for (un)employment are identical as well. Thus, our model provides a framework to think about the selectivity of hiring, while maintaining the insights about unemployment dynamics from standard labor market models.

In our model, worker heterogeneity plays the same role as the congestion externality, modelled through the aggregate matching function, in the standard model. In a boom, when productivity is high, it becomes harder to hire in the search and matching model because the labor market gets 'congested' with vacancies. In our model, hiring is costlier in a boom because firms are forced to hire workers with larger training costs in order to increase employment.

4.5 Selective Hiring

Now consider the polar opposite case, in which individual training costs are fixed over time, $\varepsilon_{it} = \varepsilon_i$. In this case, there are two reasons why a worker may be seeking a job in period t. A worker with training costs above the hiring threshold was unemployed in period $t-1$ and is therefore a job seeker in period t. Since this worker will have the same training costs in period t as she had in period $t-1$, she will be very unlikely to be hired in period t. In fact, if the economy is in steady state, the individual job-finding probability of these workers is zero. A worker with training costs below the threshold in period $t-1$ was employed in that period. However, such a worker may have been separated from her job in period t and consequently is a job seeker as well. Again assuming the economy is in steady state, if this worker was employed in period $t-1$, she will again be offered a job in period t with probability one. The fraction of 'good' workers that are seeking jobs equals λ, the probability that any given existing job is destroyed, so that $s_i = \lambda$ if $\varepsilon_{it} \leq \tilde{\varepsilon}$. Since all 'bad' workers seek jobs, $s_i = 1$ if $\varepsilon_{it} > \tilde{\varepsilon}$. Thus, the (steady state) job-finding rate in this case is given by the following expression.

$$f^{\mathrm{SH}}(\tilde{\varepsilon}) = \frac{\int_{-\infty}^{\tilde{\varepsilon}} 1 \cdot \lambda \cdot dG + \int_{\tilde{\varepsilon}}^{\infty} 0 \cdot 1 \cdot dG}{\int_{-\infty}^{\tilde{\varepsilon}} \lambda \cdot dG + \int_{\tilde{\varepsilon}}^{\infty} 1 \cdot dG} = \frac{\lambda G(\tilde{\varepsilon})}{\lambda G(\tilde{\varepsilon}) + 1 - G(\tilde{\varepsilon})} \tag{12}$$

We refer to this case as perfectly selective hiring, because at the beginning of the period everyone knows which workers will be hired and which will remain unemployed. Firms pick out the 'good' workers, with low training costs, from the pool of job seekers and ignore the 'bad' workers.

The steady state unemployment rate for the model with selective hiring equals

$$\bar{u}^{\mathrm{SH}} = 1 - G(\tilde{\varepsilon}) \tag{13}$$

Under selective hiring, the steady state unemployment rate equals the fraction of workers with training costs above the hiring threshold, because these workers will never be hired, whereas all other workers will always be immediately rehired in case they loose their job.

The differences between the model with random and selective hiring are driven by differences in the quality of the pool of job seekers between both models. If hiring is random, the pool of job seekers is a reflection of the overall distribution of workers. If hiring is selective on the other hand, workers with low training costs are unlikely to be unemployed, so that the pool of job seekers consists largely of lemons. How large this difference is depends on the separation rate λ. If $\lambda = 0$, the job-finding rate with selective hiring is equal to zero because all job seekers have training costs that are too high to be hired. If $\lambda = 1$, the job-finding rate is the same under selective and random hiring, because in both cases job seekers are representative for the distribution of all workers.

Comparing expressions (8) and (12) for the job-finding rate and (9) and (13) for the unemployment rate, it seems that the models with random and selective hiring have very different predictions for labor market dynamics. This is not true. The difference between the job-finding and unemployment rates under selective versus random hiring is mostly a level shift. If we were to use these models to generate a standard set of business cycle statistics for the volatility, persistence and comovement of labor market variables, we would calibrate the model parameters to match the steady state job-finding or unemployment rate. The differences in calibration would offset the differences in the expressions, and the predictions of the models would be quite similar.[8]

What about the equivalence of the model with a standard search and matching model? Since the expression for the job-finding rate under selective hiring (12) is different from the one under random hiring (8), it is clear that condition (11) will not make the model with selective hiring equivalent to a standard search model with a Cobb-Douglas matching function. However, we could choose a different distribution for G that would guarantee that $f^{\mathrm{SH}}(\varepsilon) = (\varepsilon/k)^{(1-\mu)/\mu}$ for all ε. Under this modified condition, our model with selective hiring would again be equivalent to a standard search model. In words, when we change the

[8] To see this, note that the elasticity of the job finding rate with respect to productivity y_t from equations (8) and (12) equals a constant times the elasticity of the hiring threshold $\tilde{\varepsilon}_t$ with respect to y_t, which is the same in both models. The proportionality factor is different in the two models, but depends only on the separation rate λ and the shape of the training costs distribution G.

assumptions on the time series properties of training costs ε_{it}, we need to recalibrate the distribution of these costs G, but we can always find a distribution that makes our model equivalent to a standard search and matching model in terms of its predictions for aggregate labor market variables.

The fact that our model with both random and selective hiring can be made equivalent to a standard search model in terms of the job creation equation does not mean, of course, that all predictions of the model are the same for random and selective hiring. In section 2, we discussed observable predictions that allow us to distinguish one model from the other in the data. In addition, the two models have very different implications for welfare analysis, to which we turn in the next section.

5 Welfare Analysis

In this section, we derive the efficient allocation in our model and compare it to the equilibrium. In order to obtain simple, easily interpretable expressions, we evaluate the model without aggregate shocks. We show that job creation can be efficient with both random and selective hiring, under slightly different conditions. We also show that the equilibrium consumption allocation with random hiring equals the allocation chosen by the social planner, but the equilibrium consumption allocation under selective hiring is far from efficient. The reason is that under selective hiring, unemployment risk is highly unequally distributed across workers. The objective of this section is to make these points in the simplest possible setting. Section 6 presents a numerical analysis to explore how important the differences are quantitatively.

5.1 Efficient Allocation

The social welfare function aggregates the utility (1) of all workers in the economy. We assume a utilitarian welfare function, which weighs the utility of all individuals equally. Thus, the social planner maximizes,

$$\sum_{t=0}^{\infty} \beta^t \int_{-\infty}^{\infty} \mathcal{U}\left(c_{it}\right) dG \tag{14}$$

subject to the law of motion for employment (2) and the aggregate resource constraint (3). The planner chooses how many workers to employ, which workers to employ, and how to distribute consumption over all employed and unemployed workers. As we argued in section 3.2, it is efficient to employ all workers with training costs ε_{it} below a threshold $\tilde{\varepsilon}_t$ and let workers with training costs above this threshold be unemployed. Imposing this property of the efficient allocation, the planner chooses the hiring threshold $\tilde{\varepsilon}_t$ and the consumption level of each worker $\{c_{it}\}_{i=-\infty}^{\infty}$ in each period t.

The solution to the social planner problem is straightforward. Details may be found in appendix A.1. The results can be summarized in two efficiency conditions, one about the efficient consumption allocation and the second one about efficient job creation.

In the efficient allocation, consumption is equal for all workers.

$$c_{it} = c_t = c \text{ for all } i \text{ and } t \tag{15}$$

The level of consumption can be found by substituting this result into the aggregate resource constraint, but is not of interest here. The important observation is that the social planner awards the same level of consumption to all workers, whether employed or unemployed and independent of their training costs ε_{it}. Of course this result depends to some degree on specific assumptions, in particular the additive separability of utility in consumption and leisure. The intuition for the result, however, is quite general. It is also important to note that any reasonable welfare function would deliver the same result. By equalizing consumption across workers, the planner minimizes the welfare loss from poor workers, who would have very steep marginal utility of consumption.

The efficient hiring threshold $\tilde{\varepsilon}_t$ depends on whether hiring is random or selective. The (steady state version of the) efficient job creation condition under random hiring is

$$K + \tilde{\varepsilon} = \frac{1+r}{r+\lambda} y - \frac{1-\lambda}{r+\lambda} G(\tilde{\varepsilon}) (\tilde{\varepsilon} - H(\tilde{\varepsilon})) \tag{16}$$

Under selective hiring, the corresponding condition is simply

$$K + \tilde{\varepsilon} = \frac{1+r}{r+\lambda} y \tag{17}$$

where $r = 1/(1-\beta)$ is the discount rate. See appendix A.1 for the derivation of these conditions.

Efficient job creation equates the net present value of output generated by the marginal job, an annuity of output y in all future periods, discounted by the rate of time preference r and the rate of job destruction λ, to the cost of creating that job. The cost of creating the marginal job includes the cost of training the marginal worker, $K + \tilde{\varepsilon}$, and under selective hiring that is the only cost of job creation. If hiring is random, however, there is an additional cost of hiring the marginal worker, which is a worker with relatively high training costs: by hiring this worker, the planner loses the option value of next period's draw of training costs for that workers being lower. This option value realizes if this period's job, if created, still exists next period, preventing the worker from obtaining a new draw, which happens with probability $(1-\lambda)$, and next period's draw for the training costs are below the hiring threshold, which happens with probability $G(\tilde{\varepsilon})$. In this case, the expected gain equals the difference between the marginal worker's training costs $\tilde{\varepsilon}$ and the average training costs of tomorrow's hires $H(\tilde{\varepsilon})$.

To compare efficient job creation conditions (16) and (17) to equilibrium job creation condition (4), we need to be more specific about wage determination. We discuss this in the next subsection.

5.2 Wage Setting and Job Creation

Wages are set by Nash bargaining between an individual worker and the firm that employs her. We assume that bargaining happens at the start of a job, and the wage stays constant thereafter until the match is destroyed. Let ϕ denote worker's bargaining power, so that the wage is given by

$$w_{it} = \phi w_R^F(\varepsilon_{it}) + (1-\phi) w_R^W \tag{18}$$

where $w_R^F(\varepsilon_{it})$ and w_R^W are the reservation wages of the firm and the worker, respectively.

The reservation wage of the firm $w_R^F(\varepsilon_{it})$ is the wage, at which profits are zero

$$w_R^F(\varepsilon_{it}) = y - \frac{r+\lambda}{1+r}(K+\varepsilon_{it}) \tag{19}$$

It is profitable for the firm to participate in a match if it can at least recoup the training costs of the worker. Thus, workers must take a pay cut with respect to the output they produce in order to pay back (in expectation) their own training costs. As a result, 'better' workers, with lower training costs, earn higher wages.

The reservation wage of the worker w_R^W depends on whether hiring is random or selective. The expressions for these reservation wage of a worker below are derived formally in appendix A.2, but they are intuitively simple to understand.

5.2.1 Random Hiring

First, consider the case of random hiring. In this case, the reservation wage of the worker,

$$w_R^W = (1-\Phi)\,b + \Phi\left[y - \frac{r+\lambda}{1+r}(K + H(\tilde{\varepsilon}))\right] \tag{20}$$

is a weighted average of unemployment benefits and the surplus (which also equals the reservation wage of the firm) of an alternative job with average training costs $H(\tilde{\varepsilon})$. The weight

$$\Phi = \frac{\phi(1-\lambda)\,G(\tilde{\varepsilon})}{r+\lambda+\phi(1-\lambda)\,G(\tilde{\varepsilon})} \tag{21}$$

depends on the worker's bargaining power as well as on the job finding rate, the separation rate and the discount rate.

For the marginal worker that is hired, the reservation wages of firm and worker equal each other. Setting $w_R^F(\tilde{\varepsilon}) = w_R^W$, using (19), evaluated in $\tilde{\varepsilon}$, and (20), and rearranging, we get the following equilibrium job creation condition.

$$K + \tilde{\varepsilon} = \frac{1+r}{r+\lambda}(y-b) - \frac{\phi(1-\lambda)}{r+\lambda}G(\tilde{\varepsilon})(\tilde{\varepsilon} - H(\tilde{\varepsilon})) \tag{22}$$

Comparing this condition to the efficient job creation condition (16), we can calculate the efficient level of unemployment benefits.

$$b^{\text{eff}} = (1-\phi)\frac{1-\lambda}{1+r}G(\tilde{\varepsilon})(\tilde{\varepsilon} - H(\tilde{\varepsilon})) \tag{23}$$

Of interest is the special case, in which workers have all the bargaining power, $\phi = 1$. In this case, efficient unemployment benefits equal zero, $b = 0$, as in a standard search and matching model.[9] More generally, if $\phi < 1$, the efficient benefit level is strictly positive. Without these positive benefits, workers accept jobs too often, because they do not fully take into account the potential gains from a new draw for their training costs.

[9] Here, we interpret b as unemployment benefits. If b represents home production and/or utility from leisure, then any b is efficient in the standard model.

5.2.2 Selective Hiring

The case of selective hiring is considerably simpler. Since in this case individual training costs are fixed over time, and there are no frictions on the labor market, a worker would be able to immediately find an identical job if bargaining were to break down. Therefore, with selective hiring, workers capture the full surplus of a job match: the worker's reservation wage equals the equilibrium wage, which must then equal the firm's reservation wage, $w_{it} = w_R^F(\varepsilon_{it})$ for all employed workers i. An unemployed worker will be unemployed in all future periods as well. Thus, the marginal worker's wage equals unemployment benefits. Setting $w_R^F(\tilde{\varepsilon}) = b$, and using (19), evaluated in $\tilde{\varepsilon}$, we get the following equilibrium job creation condition under selective hiring.

$$K + \tilde{\varepsilon} = \frac{1+r}{r+\lambda}(y-b) \tag{24}$$

Comparing this condition to the efficient job creation condition (17) under selective hiring, we immediately see that the efficient level of unemployment benefits is always zero in this case, $b^{\text{eff}} = 0$.

5.2.3 Hiring Externality

The difference in job creation between random and selective hiring is an externality. Firms do not take into account the effect of hiring an additional worker today on the amount of job seekers tomorrow. With random hiring, i.e. if training costs are fully match-specific, then a worker that was not hired today might be very employable (have low training costs) tomorrow. Since firms fail to take this effect into account, they hire more workers than is efficient. Positive unemployment benefits push up wages and counteract this overhiring, thus restoring efficiency of job creation.

With selective hiring, i.e. with fully worker-specific training costs, each worker will be just as employable tomorrow as they are today. If it is not profitable to hire a particular worker today, then this worker will not be hired tomorrow either, so that there is no external effect of today's hiring decisions on tomorrow's. Moreover, if workers have all the bargaining power, $\phi = 1$, then workers effectively pay for their own training, so that they capture the full (social) costs and benefits of hiring and thus internalize the effect of today's hiring decisions on tomorrow's. In both cases, efficient job creation is achieved by setting unemployment benefits to zero.

The hiring externality is the first difference for welfare analysis if hiring is selective. The second difference has to do with the consumption distribution, to which we now turn.

5.3 Consumption Distribution

In equilibrium, each worker i chooses her consumption in each period t in order to maximize the net present value of her utility (1), subject to a budget constraint. In order to smooth their consumption over time workers trade assets, which are in zero net supply. Since asset markets are complete, the consumption of all individual workers moves in lock-step with aggregate consumption.

$$\frac{c_{it}}{c_{it+1}} = \frac{c_t}{c_{t+1}} \text{ for all } i \text{ and } t \tag{25}$$

Without aggregate shocks, aggregate consumption is constant over time, so that individual consumption is constant over time for all individuals as well.

$$c_{it} = c_i \text{ for all } i \text{ and } t \tag{26}$$

Complete asset markets, in the absence of aggregate shocks, allow consumers to insure against variations in their income and fully smooth their consumption over time.

The level of consumption of each individual is determined by her life-time budget constraint. Assuming individuals are born with zero assets, life-time income equals the expected value of income at birth,

$$c_i = E_0 m_{it} = u_i b + (1 - u_i) E_0 w_{it} + \pi \tag{27}$$

where u_i is the unemployment rate of type ε_{it} workers and where π denotes profits, which we assume to be redistributed lump-sum from firms to workers.

Expected future income, and therefore consumption, depends exclusively on unconditional unemployment risk. By assuming asset markets are complete, we rule out any welfare costs due to bad luck. We do this on purpose, in order to focus on the welfare loss deriving from the fact that unemployment risk is distributed unequally across workers. Comparing the equilibrium condition (27) to efficiency condition (15), we see that the consumption allocation is efficient, if and only if unemployment risk is distributed evenly across workers.

5.3.1 Random versus Selective Hiring

The worker-specific unemployment risk depends crucially on whether training costs ε_{it} are transitory (match-specific) or permanent (worker-specific). In the case of perfectly random hiring, with ε_{it} uncorrelated over time, each worker gets a new draw for ε_{it} in each period, so that the unconditional unemployment risk of each worker equals the aggregate unemployment rate, $u_i^{\mathrm{RH}} = u$ for all i, and the expected wage equals the average wage $E_0 w_{it} = w$, so that consumption is equal across individuals, as in the efficient allocation (15).

In the case of perfectly selective hiring, with $\varepsilon_{it} = \varepsilon_i$ fixed over time for each worker, some workers, with low training costs, are always employed, whereas other, with training costs above the hiring threshold, are always unemployed. In this case, individual unemployment risk is highly unequally distributed.

$$u_i^{\mathrm{SH}} = \begin{cases} 0 \text{ if } \varepsilon_i \leq \tilde{\varepsilon} \\ 1 \text{ if } \varepsilon_i > \tilde{\varepsilon} \end{cases} \tag{28}$$

In addition, in this case each worker expects a different wage, consistent with her training costs, $E_0 w_{it} = w_i$. Substituting (27) and individual unemployment risk (28) and wages into the welfare function (14), we get welfare under selective and random hiring.

$$\begin{aligned} \mathcal{W}^{\mathrm{SH}} &= u\, \mathcal{U}(b + \pi) + (1 - u) \int_{-\infty}^{\tilde{\varepsilon}} \mathcal{U}(w_{it} + \pi)\, dG \\ &\leq \mathcal{U}\left(u (b + \pi) + (1 - u) \left(\int_{-\infty}^{\tilde{\varepsilon}} w_{it} dG + \pi \right) \right) = \mathcal{W}^{\mathrm{RH}} \end{aligned} \tag{29}$$

Since average wages and unemployment benefits are assumed to be the same under random and selective hiring, the inequality follows directly from the concavity of utility function $\mathcal{U}$ by Jensen's inequality.

By assuming asset markets are complete and there are no aggregate shocks, we have assumed that individual workers can completely self-insurance against unemployment risk due to bad luck. However, the differences in unemployment risk between 'good' workers with low training costs and 'bad' workers with high training costs in the model with selective hiring, are uninsurable. Once a worker is born and enters the labor market, her type ε_{it} is observable to all market participants. At that point, for workers with high training costs the bad shock has already realized and they can no longer buy insurance against it. It is this unemployment risk across workers, rather than the unemployment risk over the life-time of a worker, that drives the difference in efficiency between the models with selective and random hiring. A different way to see the same point, is that while the two models are equally efficient in creating jobs, the distribution of job opportunities is more equitable with random hiring.

In the model with selective hiring, there is in some sense a missing asset market for insurance against individual training costs. Therefore, there is a role for government intervention, insuring unborn workers against a bad draw for their training costs. This is the second difference for welfare analysis if hiring is selective. We analyze the differences between selective and random hiring for welfare quantitatively in the next section, using unemployment insurance policy as a specific example.

6 Application: Optimal Unemployment Insurance

In the previous sections, we showed that although the predictions of our model for unemployment fluctuations are very similar under random and selective hiring, welfare analysis is different in the two versions of our model. As a concrete application of this general result, in this section we explore how optimal unemployment insurance differs under (perfectly) random and (perfectly) selective hiring. We assume the government does not observe individual training costs ε_{it} and can only redistribute income based on employment status as a proxy for individual characteristics. By providing unemployment benefits, the government tries to insure workers against a bad draw for their training costs. This is a different motive for unemployment insurance than the intertemporal insurance motive usually considered in the literature. The government faces a trade-off because unemployment benefits discourage job creation.

The objectives of this section are to illustrate the main result in a concrete application and to explore how important the difference are quantitatively. In this application, we maintain the assumption from section 5 that there are no aggregate shocks.

6.1 Ramsey Problem

To derive the optimal unemployment insurance policy, we specify the Ramsey problem for a government that sets its policy instruments, unemployment benefits and lump-sum taxes, subject to its budget constraint, such that the resulting competitive equilibrium is the best possible, in the sense that it maximizes social welfare. Thus, the government chooses b and

τ to maximize welfare (14). We assume the government needs to run a balanced budget, so that the government budget constraint is given by

$$(1-n)\,b=\tau \tag{30}$$

Notice that we focus on the implications of the model for the level of unemployment insurance and therefore do not allow the government to set time-varying unemployment benefits or taxes.

In addition to its budget constraint, the government also takes the optimality conditions for job creation (4) and consumption allocation (25), the market clearing conditions for the labor market (2) and goods market (3), and wage setting rule (18) as constraints on its optimization problem.

Without aggregate shocks, i.e. $y_t = y$ for all t, the economy converges to a steady state. Assuming we start off the economy in steady state (or wait sufficiently long so that convergence has been reached), the equilibrium conditions become static. In steady state, the Ramsey planner chooses b and τ to maximize

$$\int_{-\infty}^{\infty} \mathcal{U}\left(c_{it}\right) dG \tag{31}$$

subject to the government budget constraint (30), the steady state job creation condition (22) or (24), the steady state labor market clearing condition,

$$n=\frac{f\left(\tilde{\varepsilon}\right)}{\lambda+\left(1-\lambda\right)f\left(\tilde{\varepsilon}\right)} \tag{32}$$

and the optimal consumption rule (27). Aggregate consumption must satisfy the aggregate resource constraint (3)

$$c=yn-\left[1-\left(1-\lambda\right)n\right]f\left(\tilde{\varepsilon}\right)\left(K+H\left(\tilde{\varepsilon}\right)\right) \tag{33}$$

6.2 Calibration

We use an inverse-Pareto distribution for the idiosyncratic component of training costs under random hiring G, which makes the model isomorphic to a standard search and matching model in terms of the steady state of all labor market variables as explained in section 4.4 above, see equation (11). We assume $\mu = 0.6$ and set k to match the job finding probability, as is standard in the literature (see. e.g. Mortensen and Nagypal 2007). The distribution function under selective hiring is recalibrated such that the job finding rate under selective hiring as in equation (12) equals the job finding rate under random hiring, as in equation (8), see the last paragraph of section 4.5 for details.

The targets for the job finding probability of 0.8 per quarter, and the separation rate, which we set to $\lambda = 0.07$ per quarter, are consistent with the timing of the model, which allows workers to find a job within the period after being separated, see Galí and van Rens (2017). We set the mean training costs, which would correspond to the fixed costs of vacancy creation in a search and matching model, to zero, $K = 0$.

Finally, we use logarithmic utility over consumption $U(c_i) = \log c_i$, normalize productivity to $y = 1$, and set the discount rate to 4% per year, $\beta = 0.99$. All calibration targets are matched for the model with zero unemployment benefits, $b = 0$, and full worker bargaining power, $\phi = 1$, and are kept constant across model comparisons.

6.3 Results

As a benchmark, we start with the case of random hiring and full bargaining power for workers, $\phi = 1$. Figure 1 shows employment, aggregate production, and welfare for various levels of unemployment benefits in this model. In this case, it is optimal to set unemployment benefits to zero. The reason is that there is no trade-off. With $\phi = 1$, workers effectively pay for their own training costs, so that they full internalize the hiring externality. And since unemployment risk is equally distributed across all workers, no redistribution is necessary. Nor are unemployment benefits an effective instrument for redistribution, because all workers are unemployed an equal amount of time and thus receive the same amount of benefits. Thus, production and welfare are maximized at $b = 0$.

In Figure 2, we show the same set of results, still with random hiring, but for $\phi = 0.5$. In this case, employment at $b = 0$ is inefficiently high. Workers no longer fully internalize the hiring externality because they bear only part of the costs and benefits of hiring. Therefore, workers with high training costs are employed, whereas society would be better off if they remained unemployed for one period in order to receive a new, possibly lower, draw for their training costs. Profit-maximizing firms do not take this effect into account, because the benefits of the lower training costs next period are likely to accrue to a different firm. Higher unemployment benefits counteract the externality, and in our calibration, production and welfare are maximized for a benefit level of $b = 0.20$.

With selective hiring there is a motive for redistribution, so that the government faces a trade-off: by raising unemployment benefits, the government redistributes income from unemployed to employed workers, but at the same time discourages job creation, which would be efficient with $b = 0$. Therefore, in this case the Ramsey planner cannot replicate the efficient allocation, and we would expect optimal unemployment benefits to be higher than under random hiring, resulting in an inefficiently low level of employment. Figures 3 shows employment, production and welfare as a function of unemployment benefits for the model with selective hiring. The maximum level of welfare that can be reached under selective hiring is lower than under random hiring, because unemployment benefits distort job creation.

In our calibration, the optimal level of unemployment benefits under selective hiring is higher than under random hiring, around $b = 0.26$. Comparing the results for selective hiring in Figure 3 to those for random hiring with $\phi = 0.5$ in Figure 2, we see two different reasons for positive levels of unemployment benefits, which are valid in two different worlds. If hiring is random, then positive unemployment benefits are a way to drive up wages, which corrects for inefficiently high job creation due to the hiring externality. If hiring is selective, then unemployment benefits provide insurance against idiosyncratic shocks that realize before birth and that are therefore uninsurable through private markets. Despite the observation that in our simulations the optimal levels of unemployment benefits in both worlds may be similar for some parameterizations, it is worth noting that the welfare losses of inefficiently low unemployment benefits are two orders of magnitude greater under selective hiring than under random hiring, and we therefore think of selective hiring as the more convincing argument for positive unemployment benefits.

7 Conclusions

In the real world, hiring decisions are selective. Firms choose not only how many, but also which workers to hire. As a result, job-finding probabilities and unemployment risk vary across workers. In standard search models of the labor market, however, hiring is random, in the sense that the job-finding probability is the same for all workers. In this paper we argue that selectivity in hiring strongly affects conclusions about welfare.

We present a model, in which hiring decisions may be random or selective. The predictions of this model for unemployment fluctuations are identical to those of a standard search and matching model. We also show, however, that the predictions of the model regarding welfare may be very different for selective versus random hiring. As an application, we analyze optimal unemployment insurance in our framework.

Under random hiring, the government can replicate the efficient allocation, using unemployment benefits and lump-sum taxes as instruments. In this case, unemployment benefits are set to make sure the level of job creation is efficient. The optimal level of unemployment benefits is usually positive, because a hiring externality leads to overhiring, but the welfare losses of inefficiently low benefit levels are small.

Under selective hiring, the hiring externality is not present, but -because unemployment risk is distributed unequally across workers- the government faces a trade-off between efficient job creation and efficient redistribution. There is an additional motive for unemployment insurance, because workers cannot self-insure against their characteristics, which determine their individual-specific unemployment risk. As a result, under selective hiring unemployment benefits are usually higher (and employment and welfare lower) than under random hiring, and the welfare gains from positive benefit levels are substantial.

We believe the point we make in this paper applies fairly generally. In our model, heterogeneity takes the form of differences in training costs across workers. However, other sources of heterogeneity are likely to have very similar implications. In particular, we show in appendix A.3 that a model with workers that differ in their productivity, in the absence of aggregate shocks, is isomorphic to our model.

A Appendices

A.1 Social Planner Problem

The value function and the Bellman equation of the social planner problem (14) are given by

$$V(n_{t-1}; y_t) = \max_{\left\{\tilde{\varepsilon}_{t+\tau}, \{c_{it+\tau}\}_{i=-\infty}^{\infty}\right\}_{\tau=t}^{\infty}} E_t \sum_{\tau=0}^{\infty} \beta^{\tau} \int_{-\infty}^{\infty} \mathcal{U}(c_{it+\tau})\, dG \tag{34}$$

$$= \max_{\tilde{\varepsilon}_t, \{c_{it}\}_{i=-\infty}^{\infty}} \left\{ \int_{-\infty}^{\infty} \mathcal{U}(c_{it})\, dG + \beta E_t V(n_t; y_{t+1}) \right\} \tag{35}$$

where y_t is an exogenous state variable and n_{t-1} an endogenous state variable, with law of motion as in equation (2),

$$n_t = (1-\lambda)\, n_{t-1} + JC(n_{t-1}, \tilde{\varepsilon}_t) \tag{36}$$

Notice that hiring is instantaneous (there are no search frictions in this economy), so that n_{t-1}, not n_t, is the state variable. The hiring threshold and consumption in period t are chosen subject to the aggregate resource constraint (3).

$$\int_{-\infty}^{\infty} c_{it} dG = y_t n_t - JC(n_{t-1}, \tilde{\varepsilon}_t)(K + H(\tilde{\varepsilon}_t)) \tag{37}$$

Let μ_t denote the multiplier associated with the aggregate resource constraint in period t.

The first set of efficiency conditions resulting from this optimization problem are the first-order conditions for c_{it}

$$\mathcal{U}'(c_{it}) = \mu_t \tag{38}$$

These first-order conditions for c_{it} immediately imply that $c_{it} = c_t$, as in equation (15) in the main text. The level of consumption in period t can be found by substituting this into the aggregate resource constraint, noting that G is a CDF, so that $\int_{-\infty}^{\infty} c_{it} dG = c_t \int_{-\infty}^{\infty} dG = c_t$.

Efficient job creation depends on whether hiring is random or selective.

A.1.1 Random hiring

Under random hiring, a fraction $G(\tilde{\varepsilon}_t)$ of all job seekers receive a relatively low draw for their training costs, making them employable. Thus, the number of workers that are hired equals

$$JC(n_{t-1}, \tilde{\varepsilon}_t) = [1 - (1-\lambda)\, n_{t-1}]\, G(\tilde{\varepsilon}_t) \tag{39}$$

a fraction $G(\tilde{\varepsilon}_t)$ of the total number of non-employed workers, including the ones that were separated in this period.

Substituting this expression for job creation into the law of motion for employment (36) and the aggregate resource constraint (37) above, we get the following first order condition for $\tilde{\varepsilon}_t$

$$y_t - K - H(\tilde{\varepsilon}_t) - \frac{H'(\tilde{\varepsilon}_t)\, G(\tilde{\varepsilon}_t)}{g(\tilde{\varepsilon}_t)} + \frac{\beta E_t [V'(n_t; y_{t+1})]}{\mu_t} = 0 \tag{40}$$

and envelope condition for n_{t-1}

$$\begin{aligned} V'(n_{t-1};y_t) &= (1-\lambda)\mu_t\left\{(1-G(\tilde{\varepsilon}_t))y_t + G(\tilde{\varepsilon}_t)(K+H(\tilde{\varepsilon}_t))\right\} \\ &\quad + (1-\lambda)(1-G(\tilde{\varepsilon}_t))\beta E_t V'(n_t;y_{t+1}) \end{aligned} \tag{41}$$

Using the first order condition for $\tilde{\varepsilon}_t$ to substitute out for $E_t[V'(n_t;y_{t+1})]$ in the envelope condition for n_{t-1}, we get an expression for $V'(n_{t-1};y_t)$.

$$V'(n_{t-1};y_t) = (1-\lambda)\mu_t\left(K + H(\tilde{\varepsilon}_t) + (1-G(\tilde{\varepsilon}_t))\frac{H'(\tilde{\varepsilon}_t)G(\tilde{\varepsilon}_t)}{g(\tilde{\varepsilon}_t)}\right) \tag{42}$$

Substituting this expression back into the envelope condition for n_{t-1}, we get an Euler equation for the hiring threshold $\tilde{\varepsilon}_t$,

$$K + M(\tilde{\varepsilon}_t) = y_t + \beta(1-\lambda)E_t\left[\frac{\mu_{t+1}}{\mu_t}\left\{K + M(\tilde{\varepsilon}_{t+1}) - G(\tilde{\varepsilon}_{t+1})(M(\tilde{\varepsilon}_{t+1}) - H(\tilde{\varepsilon}_{t+1}))\right\}\right] \tag{43}$$

where $\mu_t = \mathcal{U}'(c_t)$ and

$$M(\tilde{\varepsilon}_t) = H(\tilde{\varepsilon}_t) + \frac{H'(\tilde{\varepsilon}_t)G(\tilde{\varepsilon}_t)}{g(\tilde{\varepsilon}_t)} \tag{44}$$

To evaluate $M(\tilde{\varepsilon}_t)$, we need the distribution of training costs ε_{it} in the pool of job seekers. Under perfectly random hiring, the distribution of training costs among job seekers equals the unconditional distribution of training costs, so this distribution is just G. Then, average training costs are given by the following expression.

$$H(\tilde{\varepsilon}_t) = \frac{\int_{-\infty}^{\tilde{\varepsilon}_t}\varepsilon dG(\varepsilon)}{G(\tilde{\varepsilon}_t)} = \frac{\int_{-\infty}^{\tilde{\varepsilon}_t}\varepsilon dG(\varepsilon)}{G(\tilde{\varepsilon}_t)} \tag{45}$$

Taking a derivative with respect to the training costs of the marginal hire $\tilde{\varepsilon}_t$, we get,

$$H'(\tilde{\varepsilon}_t) = \frac{\tilde{\varepsilon}_t g(\tilde{\varepsilon}_t)}{G(\tilde{\varepsilon}_t)} - \frac{\int_{-\infty}^{\tilde{\varepsilon}_t}\varepsilon dG(\varepsilon)}{G(\tilde{\varepsilon}_t)^2}g(\tilde{\varepsilon}_t) = \frac{g(\tilde{\varepsilon}_t)}{G(\tilde{\varepsilon}_t)}\left[\tilde{\varepsilon}_t - H(\tilde{\varepsilon}_t)\right] \tag{46}$$

so that $M(\tilde{\varepsilon}_t) = \tilde{\varepsilon}_t$. Substituting into Euler equation (43) gives the efficient job creation condition,

$$K + \tilde{\varepsilon}_t = y_t + (1-\lambda)E_t\left[Q_{t,t+1}\left\{K + \tilde{\varepsilon}_{t+1} - G(\tilde{\varepsilon}_{t+1})(\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1}))\right\}\right] \tag{47}$$

and evaluating this equation in steady state yields condition (16) in the main text.

A.1.2 Selective hiring

Under selective hiring, the problem is asymmetric depending on whether the initial level of employment is above or below steady state. If last period's employment was below the desired steady state level, then all workers with training costs below $\tilde{\varepsilon}_t$ are hired immediately, and employment jumps to its steady state level. If we start from a level of employment that is above the desired level or the steady state, i.e. if $n_{t-1} \geq G(\tilde{\varepsilon}_t)$, then only those workers that are separated *and* have training costs below $\tilde{\varepsilon}_t$ will be hired, i.e. a fraction $G(\tilde{\varepsilon}_t)/n_{t-1}$

of the λn_{t-1} workers that are separated, so that

$$JC\left(n_{t-1}, \tilde{\varepsilon}_t\right) = \lambda G\left(\tilde{\varepsilon}_t\right) \tag{48}$$

Employment will slowly fall, as workers with training costs above $\tilde{\varepsilon}_t$ that are employed for historic reasons gradually lose their jobs.

Substituting this expression for job creation into the law of motion for employment (36) and the aggregate resource constraint (37) above, the Bellman equation can be written as,

$$V\left(n_{t-1}; y_t\right) = \max_{\tilde{\varepsilon}_t, \{c_{it}\}_{i=-\infty}^{\infty}} \left\{ \int_{-\infty}^{\infty} \mathcal{U}\left(c_{it}\right) dG + \beta E_t V\left(n_t; y_{t+1}\right) \right\} \tag{49}$$

where

$$n_t = (1-\lambda) n_{t-1} + \lambda G\left(\tilde{\varepsilon}_t\right) \tag{50}$$

and subject to the aggregate resource constraint.

$$\int_{-\infty}^{\infty} c_{it} dG = y_t n_t - \lambda G\left(\tilde{\varepsilon}_t\right)\left(K + H\left(\tilde{\varepsilon}_t\right)\right) \tag{51}$$

The first-order condition for $\tilde{\varepsilon}_t$

$$y_t - K - H\left(\tilde{\varepsilon}_t\right) - \frac{G\left(\tilde{\varepsilon}_t\right) H'\left(\tilde{\varepsilon}_t\right)}{g\left(\tilde{\varepsilon}_t\right)} + \frac{\beta E_t V'\left(n_t; y_{t+1}\right)}{\mu_t} = 0 \tag{52}$$

and the envelope condition for n_{t-1}

$$V'\left(n_{t-1}; y_t\right) = (1-\lambda) \mu_t y_t + (1-\lambda) \beta E_t V'\left(n_t; y_{t+1}\right) \tag{53}$$

can be combined to get the following Euler equation for $\tilde{\varepsilon}_t$

$$K + M\left(\tilde{\varepsilon}_t\right) = y_t + \beta (1-\lambda) E_t \left[\frac{\mu_{t+1}}{\mu_t} \left\{K + M\left(\tilde{\varepsilon}_{t+1}\right)\right\} \right] \tag{54}$$

where

$$M\left(\tilde{\varepsilon}_t\right) = H\left(\tilde{\varepsilon}_t\right) + \frac{H'\left(\tilde{\varepsilon}_t\right) G\left(\tilde{\varepsilon}_t\right)}{g\left(\tilde{\varepsilon}_t\right)} \tag{55}$$

as under random hiring.

In steady state, the distribution of training costs among new hires equals G, as for random hiring, so that $M\left(\tilde{\varepsilon}\right) = \tilde{\varepsilon}$. This gives the efficient job creation condition under selective hiring (17) in the main text.

A.2 Nash bargaining over Wages

A.2.1 Derivation of equation (20)

Let V^U and $V^E\left(w_{it}\right)$ denote the expected net present value of payoffs for an unemployed worker, and an employed worker with wage w_{it}, respectively.

Under random hiring, the values V^U and $V^E\left(W\left(\tilde{\varepsilon}\right)\right)$ are determined by the following system of Bellman equations,

$$V^U = b + G(\tilde{\varepsilon})\beta V^E(W(\tilde{\varepsilon})) + (1 - G(\tilde{\varepsilon}))\beta V^U \quad (56)$$

$$V^E(w_{it}) = w_{it} + \lambda(1 - G(\tilde{\varepsilon}))\beta V^U + (1-\lambda)\beta V^E(w_{it}) + \lambda G(\tilde{\varepsilon})\beta V^E(W(\tilde{\varepsilon})) \quad (57)$$

where $W(\tilde{\varepsilon})$ is the wage of the average employed worker.

At the reservation wage w_R^W, a worker is just indifferent between working or being unemployed, $V^E(w_R^W) - V^U = 0$, so that

$$V^E(\tilde{w}) - V^U = \tilde{w} - b + (1-\lambda)\beta\left(V^E(\tilde{w}) - V^U\right) - (1-\lambda)G(\tilde{\varepsilon})\beta\left(V^E(W(\tilde{\varepsilon})) - V^U\right) = 0 \quad (58)$$

$$w_R^W = b + (1-\lambda)G(\tilde{\varepsilon})\beta\left(V^E(W(\tilde{\varepsilon})) - V^U\right) \quad (59)$$

Averaging the Bellman equation for an employed worker over training costs, we get

$$V^E(W(\tilde{\varepsilon})) - V^U = W(\tilde{\varepsilon}) - b + (1-\lambda)(1-G(\tilde{\varepsilon}))\beta\left(V^E(W(\tilde{\varepsilon})) - V^U\right) \quad (60)$$

$$= \frac{1+r}{r+\lambda+(1-\lambda)G(\tilde{\varepsilon})}(W(\tilde{\varepsilon}) - b) \quad (61)$$

and substituting this into (59) we get the following expression for the reservation wage of workers.

$$w_R^W = b + (1-\lambda)G(\tilde{\varepsilon})\frac{W(\tilde{\varepsilon}) - b}{r+\lambda+(1-\lambda)G(\tilde{\varepsilon})} \quad (62)$$

To eliminate $W(\tilde{\varepsilon})$, we evaluate the Nash bargaining wage rule (18) and combine it with expression (19), evaluated in $H(\tilde{\varepsilon})$, and (62) to get an expression for w_R^W in terms of the hiring threshold $\tilde{\varepsilon}$ and parameters only.

$$W(\tilde{\varepsilon}) = \phi w_R^F(H(\tilde{\varepsilon})) + (1-\phi)w_R^W = \phi\left[y - \frac{r+\lambda}{1+r}(K + H(\tilde{\varepsilon}))\right] + (1-\phi)w_R^W \quad (63)$$

Substituting into (62),

$$w_R^W = b + (1-\lambda)G(\tilde{\varepsilon})\frac{\phi\left[y - \frac{r+\lambda}{1+r}(K+H(\tilde{\varepsilon}))\right] + (1-\phi)w_R^W - b}{r+\lambda+(1-\lambda)G(\tilde{\varepsilon})} \quad (64)$$

$$= \frac{r+\lambda}{r+\lambda+\phi(1-\lambda)G(\tilde{\varepsilon})}b + \frac{\phi(1-\lambda)G(\tilde{\varepsilon})}{r+\lambda+\phi(1-\lambda)G(\tilde{\varepsilon})}\left[y - \frac{r+\lambda}{1+r}(K+H(\tilde{\varepsilon}))\right] \quad (65)$$

gives equation (20) in the main text.

A.2.2 Derivation of equation (22)

Using (19), evaluated in $\tilde{\varepsilon}$

$$w_R^F(\tilde{\varepsilon}) = y - \frac{r+\lambda}{1+r}(K+\tilde{\varepsilon}) \quad (66)$$

and (20), and setting $w_R^F(\tilde{\varepsilon}) = w_R^W$, we get the following expression that pins down $\tilde{\varepsilon}$.

$$y - \frac{r+\lambda}{1+r}(K+\tilde{\varepsilon}) = (1-\Phi)b + \Phi\left[y - \frac{r+\lambda}{1+r}(K+H(\tilde{\varepsilon}))\right] \quad (67)$$

Add and subtract $\tilde{\varepsilon}$ in the right-hand-side expression,

$$y - \frac{r+\lambda}{1+r}(K+\tilde{\varepsilon}) = (1-\Phi)\,b + \Phi\left[y - \frac{r+\lambda}{1+r}\left(K+\tilde{\varepsilon}-\tilde{\varepsilon}+H(\tilde{\varepsilon})\right)\right] \tag{68}$$

take $y - \frac{r+\lambda}{1+r}(K+\tilde{\varepsilon})$ to the left-hand side,

$$(1-\Phi)\left[y - \frac{r+\lambda}{1+r}(K+\tilde{\varepsilon})\right] = (1-\Phi)\,b + \Phi\left[\frac{r+\lambda}{1+r}\left(\tilde{\varepsilon}-H(\tilde{\varepsilon})\right)\right] \tag{69}$$

take y to the right-hand side, divide by $1-\Phi$, and multiply by $-\frac{1+r}{r+\lambda}$

$$K+\tilde{\varepsilon} = \frac{1+r}{r+\lambda}(y-b) - \frac{\Phi}{1-\Phi}\left(\tilde{\varepsilon}-H(\tilde{\varepsilon})\right) \tag{70}$$

Finally, simplify the $\frac{\Phi}{1-\Phi}$ term,

$$\frac{\Phi}{1-\Phi} = \frac{\phi(1-\lambda)\,G(\tilde{\varepsilon})}{r+\lambda+\phi(1-\lambda)\,G(\tilde{\varepsilon})-\phi(1-\lambda)\,G(\tilde{\varepsilon})} = \frac{\phi(1-\lambda)\,G(\tilde{\varepsilon})}{r+\lambda} \tag{71}$$

and substitute above the get equation (22) in the main text.

A.3 Heterogeneous Productivity

Job creation condition (4) already imposes the 'threshold property' of the equilibrium, i.e. all workers with training costs below a threshold $\varepsilon_{it} < \tilde{\varepsilon}_t$ are hired and all those above are not. More generally, we can rewrite this condition as a condition for each individual worker i with training costs ε_{it} to be hired if

$$K+\varepsilon_{it} \leq E_t \sum_{\tau=0}^{\infty} (1-\lambda)^{\tau}\, Q_{i,t,t+\tau}\left(y_{t+\tau}-w_{it}\right) \tag{72}$$

Without aggregate shocks, this condition, which simply states that a worker is hired if the total hiring and training costs $K+\varepsilon_{it}$ are smaller or equal than the expected net present value of profits generated by hiring the worker, reduces to

$$K+\varepsilon_{it} \leq \frac{1+r}{r+\lambda}\left(y-w_{it}\right) \tag{73}$$

The wage w_{it} will depend on individual training costs as well as aggregate productivity in period t when the worker was hired and the Nash bargain took place, as can be seen by combining equations (18), (19) and (20) for random hiring, and equation (19) and $w_{it} = w_R^F(\varepsilon_{it})$ for selective hiring, see section 5.2. Both with random and with selective hiring, the wage rule is linear in ε_{it} and y_t.

$$w_{it} = \Omega_0 + \Omega_{\varepsilon}\varepsilon_{it} + \Omega_y y \tag{74}$$

where the coefficients Ω_0, Ω_{ε} and Ω_y will be different for random versus selective hiring and will depend on model parameters and endogenous variables, in particular $\tilde{\varepsilon}$, but will be constant across individuals.

Now suppose that in addition to heterogeneity in training costs, $\varepsilon_{it} \sim G$, workers also differ in their productivity, $y_{it} \sim F$. We assume that once a match is started, productivity is fixed for the duration of the match. Then, with heterogeneous productivity, job creation condition (73) changes to,

$$K + \varepsilon_{it} \leq \frac{1+r}{r+\lambda} (y_{it} - w_{it}) \tag{75}$$

where $w_{it} = \Omega_0 + \Omega_\varepsilon \varepsilon_{it} + \Omega_y y_{it}$ analogous to (74).

It is straightforward to see that (75) can be rewritten as

$$K + \varepsilon'_{it} \leq \frac{1+r}{r+\lambda} (y - w_{it}) \tag{76}$$

where $w_{it} = \Omega_0 + \Omega_\varepsilon \varepsilon'_{it} + \Omega_y y$ and

$$\varepsilon'_{it} = \varepsilon_{it} + \frac{(1+r)(1-\Omega_y)}{r+\lambda+(1+r)\Omega_\varepsilon} (y_{it} - y) \tag{77}$$

so that heterogeneous productivity can be rephrased as heterogeneous training costs.

We conclude that we can learn all there is to learn from a model with heterogeneous productivity also from a model with heterogeneous training costs. As a corollary, we argue that while correlation between y_{it} and ε_{it} is important, in the sense that it affects the variation in ε'_{it}, it does not matter for the predictions of the model conditional on ε'_{it}. For example, if y_{it} and ε_{it} are positively correlated, then ε'_{it} will be less dispersed than ε_{it} (and vice versa if they are negatively correlated), but once we calibrate the variance of the distribution of ε'_{it} to its true value, the predictions of this model will be the same as the predictions of a model with homogeneous productivity and $\varepsilon_{it} = \varepsilon'_{it}$.

References

[1] Abraham, K. and R. Shimer (2002). Changes in Unemployment Duration and Labor-Force Attachment, in: Krueger, Alan and Robert Solow (eds.), *The Roaring Nineties: Can Full Employment Be Sustained?* Russell Sage Foundation, 367-420.

[2] Acemoglu, D. and R. Shimer (1999), Efficient Unemployment Insurance, *Journal of Political Economy*, 107 (5), 893-928.

[3] Barnichon, R. and A. Figura (2015). Labor Market Heterogeneity and the Aggregate Matching Function, *American Economic Journal: Macroeconomics,* 7 (4), 222–249.

[4] Becker, G.S. (1973). A Theory of Marriage: Part I, *Journal of Political Economy* 81(4), 813–846.

[5] Berger, D. (2016). Countercyclical Restructuring and Jobless Recoveries, Working Paper, Northwestern University.

[6] Brown, A.J.G., C. Merkl and D.J. Snower (2015). An Incentive Theory Of Matching, *Macroeconomic Dynamics*, 19 (3), 643-668.

[7] Chetty, R. (2006). A General Formula for the Optimal Level of Social Insurance. *Journal of Public Economics* 90, 1879-1901.

[8] Chugh, S. K. and C. Merkl (2016). Efficiency and Labor Market Dynamics in a Model of Labor Selection, *International Economic Review* 57 (4), 1371-1404.

[9] Diamond, P.A. (1982). Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy* 90 (5), 881–94.

[10] Epstein, B. (2016). Heterogeneous Workers, Optimal Job Seeking, and Aggregate Labor Market Dynamics, Working Paper, University of Massachusetts Lowell

[11] Fernández-Blanco, J. and E. Preugschat (2018). On the Effects of Ranking by Unemployment Duration, *European Economic Review*, 104 (C), 92-110.

[12] Fredriksson, P. and B. Holmlund (2006). Improving Incentives in Unemployment Insurance: A Review of Recent Research, *Journal of Economic Surveys*, 20 (3), 357-386.

[13] Galí, J. and T. van Rens (2017). The Vanishing Procyclicality of Labor Productivity, Working Paper.

[14] Hall, R.E. (1978). Stochastic Implications of the Life-Cycle Permanent-Income Hypothesis, *Journal of Political Economy* 86(6).

[15] Hopenhayn, H.A. and J.P. Nicolini (1997). Optimal Unemployment Insurance. *Journal of Political Economy* 105 (2), 412-438.

[16] Hopenhayn, H.A. and J.P. Nicolini (2009). Optimal Unemployment Insurance and Employment History, *Review of Economic Studies* 76 (3), 1049-1070.

[17] Hornstein, A. (2012). Accounting for Unemployment: The Long and Short of It, Working Paper, Federal Reserve Bank of Richmond.

[18] Hosios, A.J. (1990). Factor Market Search and the Structure of Simple General Equilibrium Models, *Journal of Political Economy* 98 (2), 325-355.

[19] Krusell, P. and A. Smith (1998), Income and Wealth Heterogeneity in the Macroeconomy, *Journal of Political Economy* 5, 867-896.

[20] Landais C., P. Michaillat and E. Saez (2016). A Macroeconomic Approach to Optimal Unemployment Insurance: Theory, *American Economic Journal: Economic Policy* 10 (2), 152-181.

[21] Lifschitz, A., O. Sett and Y. Yedid-Levi (2017). On the Provision of Unemployment Insurance when Workers are Ex-Ante Heterogeneous, Working Paper, Tel Aviv University

[22] Lucas, R.E. (1987), *Models of Business Cycles.* New York: Basil Blackwell, 1987

[23] Mitman, K. and S. Rabinovich (2015). Optimal Unemployment Insurance in an Equilibrium Business-Cycle Model, *Journal of Monetary Economics* 71, 99-118.

[24] Moen, E.R. (1997). Competitive Search Equilibrium, *Journal of Political Economy* 105(2), 385-411.

[25] Mortensen, D.T. (1982). Property Rights and Efficiency in Mating, Racing, and Related games. *American Economic Review* 72 (5), 968–79.

[26] Mortensen, D.T. and E. Nagypal (2007). More on unemployment and vacancy fluctuations. *Review of Economic Dynamics* 10 (3), 327–347.

[27] Moyen, S. and N. Stähler (2014). Unemployment Insurance and the Business Cycle: Should Benefit Entitlement Duration React to the Cycle? *Macroeconomic Dynamics* 18, 497-525.

[28] Mueller, A. (2017). Separations, Sorting and Cyclical Unemployment, *American Economic Review* 107 (7), 2081-2107.

[29] Pissarides, C.A. (1985). Short-Run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages. *American Economic Review* 75 (4), 676–90.

[30] Pissarides, C.A. (2000). *Equilibrium Unemployment Theory* (2nd ed.). Cambridge: MIT Press.

[31] Pissarides, C.A. (2009). The unemployment volatility puzzle: Is wage stickiness the answer? Walras-Bowley lecture, North American summer meetings of the Econometric Society, Duke University. *Econometrica* 77 (5), 1339-1369.

[32] Shimer, R. (2007). Mismatch. *American Economic Review* 97(4), 1074–1101.

[33] Shimer, R. (2012), Reassessing the Ins and Outs of Unemployment, *Review of Economic Dynamics,* 15 (2), 127-148.

[34] Solon, G., R. Barsky and J. Parker (1994), Measuring the Cyclicaly of Real Wages: How Important is Composition Bias? *Quarterly Journal of Economics* 109, 1–25.

[35] Wilke, R. (2005), New Estimates of the Duration and Risk of Unemployment for West-Germany, *Schmollers Jahrbuch* 125, 207–237.

Figure 1. Unemployment Insurance with Random Hiring, $\phi = 1$

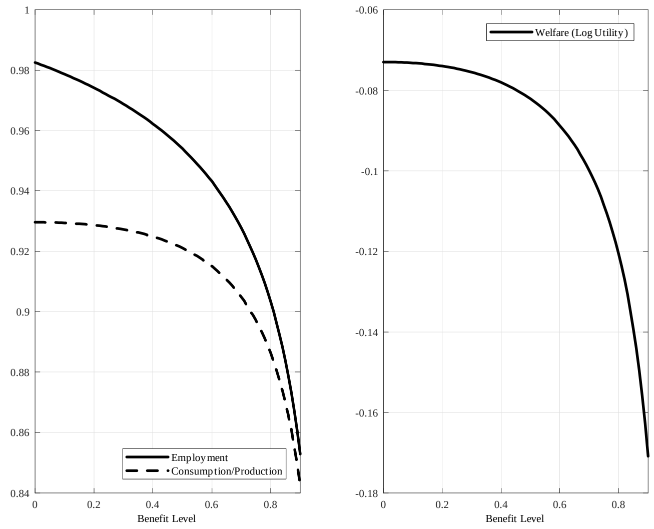

Figure 2. Unemployment Insurance with Random Hiring, $\phi = 0.5$

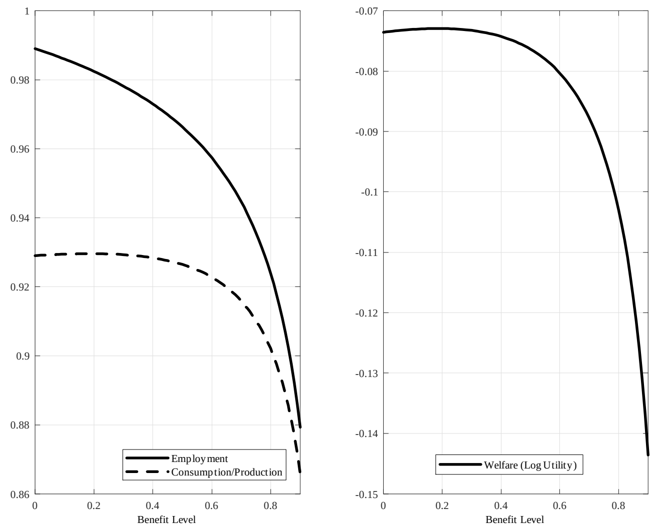

Figure 3. Unemployment Insurance with Selective Hiring