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

January 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 the welfare costs of unemployment are much larger because unemployment risk is distributed unequally across workers. As a result, optimal unemployment insurance is higher and welfare lower if hiring is selective.

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.

1 Introduction

Standard search 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. 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, 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, 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.[1]

If training costs have both a transitory and a permanent 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 output, job-finding and unemployment rate, 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 employable, but some more so than others. Thus, unemployment risk is uninsurable and the welfare costs of unemployment are much larger than under random hiring. As a result, there is a role

[1] 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.

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, similar to the Hosios (1990) condition in search and matching models. 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.

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). 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 (2017), 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.

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 and solves for the efficient allocation. We describe the equilibrium of the model over two sections. In section 4, we derive the equilibrium job creation condition and discuss conditions, under which equilibrium job creation is efficient. We also formally establish the equivalence of our model with random hiring to a search model with an aggregate matching function in this section. Section 5 deals with welfare analysis. Here, we show analytically how the welfare costs of unemployment differ starkly under selective versus random hiring. Section 6 establishes this point more formally 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 enta 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). In addition, Mueller (2017) recently documented that the average predicted wage of unemployed workers is countercyclical.

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. 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.[2] 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 for each individual, then each worker expects to have the 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, 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_t and unemployed workers receive unemployment benefits b_t. Potentially, wages and unemployment benefits could depend on worker characteristics ε_{it}, but we assume that this is not the case. This assumption, which we will maintain throughout the paper, is not very restrictive, because only the wage of the marginal hire is allocative, see section 4.2 for a more detailed discussion.

We assume the parameters of the model are such that $b_t < w_t$ for all t, so that all unemployment is involuntary. For simplicity, we also assume the utility from leisure is included in the unemployment benefit b_t, so that the flow utility $\mathcal{U}(.)$ depends only on consumption. Then, workers' objective function is given by,

[2] 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.

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

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_t, the firm's profits from a job equal $y_t - w_t$. 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 profitable/efficient to create a job for a worker with training costs ε, then it must also be profitable/efficient to create a job for a worker with lower training costs $\varepsilon' < \varepsilon$. Thus, in the efficient as well as in the equilibrium allocation of this model, there exists a 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 nevertheless impose it below in order to simplify the notation.

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. 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}. 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 Efficiency

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

$$E_0 \sum_{t=0}^{\infty} \beta^t \int_{-\infty}^{\infty} \mathcal{U}(c_{it}) dG \tag{4}$$

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 may 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.

$$c_{it} = c_t \text{ for all } i \text{ and } t \tag{5}$$

The level of consumption in period t 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. This result will play a crucial role in the welfare analysis in section 5.

The efficient job creation equation determines the hiring threshold $\tilde{\varepsilon}_t$, i.e. the training cost of the marginal worker that is hired, which in turn pins down the efficient job-finding rate $f(\tilde{\varepsilon}_t)$,

$$K + \tilde{\varepsilon}_t = y_t + (1-\lambda) E_t \left[Q_{t,t+1} \left\{ K + \tilde{\varepsilon}_{t+1} - f(\tilde{\varepsilon}_{t+1}) (\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1})) \right\} \right] \tag{6}$$

where $Q_{t,t+1}$ is the stochastic discount factor between periods t and $t+1$.

$$Q_{t,t+1} = \frac{\beta \mathcal{U}'(c_{t+1})}{\mathcal{U}'(c_t)} \tag{7}$$

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 $\tau = 0$.

Condition (6) has the usual interpretation of a job creation condition, stating that the (social) cost of hiring the marginal worker must equal the expected net present (social) value of having that worker employed. The cost of hiring the marginal worker equals the training costs of the marginal worker, the fixed training costs K plus the idiosyncratic training costs of the marginal worker $\tilde{\varepsilon}_t$, which gives the left-hand side of the condition. The right-hand side of condition (6) represents the benefits of hiring this worker, which include the output produced by the worker y_t, plus the expected cost saving from not having to hire a worker next period, $K + \tilde{\varepsilon}_{t+1}$. The $f(\tilde{\varepsilon}_{t+1})(\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1}))$ term comes from the fact that the planner takes into account that by hiring an additional worker today, there will be fewer job seekers tomorrow, see section 4.1 for a more detailed discussion. The benefits of having an additional worker tomorrow are discounted not only by the stochastic discount factor $Q_{t,t+1}$, but also by the probability that the job continues to next period $1-\lambda$.

3.5 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 deals with job creation and derives conditions under which efficient job creation can be supported as an equilibrium. Section 5 deals with the consumption allocation. 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 compare it to the efficient job creation condition. We show that, under some conditions, equilibrium job creation is efficient. Then, we 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.[3]

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, firms decide whether or not to create jobs.[4] Like the social planner, firms must choose how many workers as well as which workers to employ. As we argued in section 3.2, it is profit-maximizing 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 equilibrium, firms directly choose 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.[5]

$$K + \tilde{\varepsilon}_t = E_t \sum_{\tau=0}^{\infty} (1-\lambda)^{\tau} Q_{t,t+\tau} \left(y_{t+\tau} - w_{t+\tau} \right) \tag{8}$$

In recursive form, this condition can be written as,

$$K + \tilde{\varepsilon}_t = y_t - w_t + (1-\lambda) E_t \left[Q_{t,t+1} \left(K + \tilde{\varepsilon}_{t+1} \right) \right] \tag{9}$$

Appendix A.2 provides an alternative derivation of equilibrium job creation condition (9) from the profit maximization problem of a representative firm.

Comparing the equilibrium job creation condition (9) to the efficient job creation condition (6), there are two differences. First, the flow payoff from having an additional employed worker to firms does not equal the full social flow surplus y_t that this worker generates, but only the part of this surplus that accrues to the firm, $y_t - w_t$, because the worker needs to be paid a wage w_t. Second, unlike the social planner, firms do not take into account

[3] 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."

[4] We assume that workers always accept all job offers and there is full commitment of both firm and worker once a job has been created and a wage been agreed upon, see section 3.1.

[5] Given the timing assumptions in our model, the wage does not depend on idiosyncratic training costs and is the same for all workers, see section 4.2.

the effect of hiring an additional worker today on the amount of job seekers tomorrow. This externality induces firms to hire more workers than is efficient.[6] A positive wage, which makes it less profitable for firms to hire, may counteract this externality and restore efficiency of job creation.

4.2 Wage Setting

The equilibrium job creation condition (9) can be made identical to the the efficient job creation condition (6) by choosing the appropriate wage setting rule. Thus, the only reason why equilibrium job creation is inefficient in this model, is that the wage may deviate from its efficient level. Comparing equations (9) and (6), we obtain the following efficient wage rule.

$$w_t = (1-\lambda) E_t \left[Q_{t,t+1} f(\tilde{\varepsilon}_{t+1}) (\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1})) \right] \tag{10}$$

In general, the efficient wage depends on the distribution of training costs of job seekers, which affect the average cost function $H(\tilde{\varepsilon})$.

In order for job creation to be efficient, it is not necessary that wages depend on individual worker characteristics ε_{it}, as we anticipated in section 3.1. The reason is that only the wage of the marginal worker, with training costs $\tilde{\varepsilon}_t$, is allocative, as long as it is more profitable to hire workers with lower training costs in equilibrium. Therefore, we limit our attention to wages that do not depend on ε_{it} for simplicity.

The assumption that in our model all employed workers earn the same wage can be justified with appropriate timing assumptions. First, the values of the aggregate productivity y_t and of the random training cost component ε_{it} are revealed. Second, firms make their hiring decisions, taking y_t and ε_{it} into account and anticipating the result of the wage formation process. Third, after workers are hired and trained, so that the training costs are sunk, their wage is determined. These timing assumptions correspond to those in random search models (vacancies are posted, the match occurs and only afterwards the wage is determined, usually by Nash bargaining). In our model, the timing assumptions are without loss of generality because even if the intramarginal workers earned a higher wage, the allocation would be unchanged. However, these assumptions allow us to generate tractable analytical results and to show that the search and matching model and the random hiring model are isomorphic, see section 4.5.

We argue the efficient amount of job creation can always be supported as an equilibrium with the right wage setting mechanism. Therefore, we must verify that there exist parameter values, for which the efficient wage rule (10) satisfies the full participation condition, $w_t > b_t$ for all t. To see that this is the case, notice that the average training costs of employed workers $H(\tilde{\varepsilon}_t)$ must be strictly lower than the training costs of the marginal worker $\tilde{\varepsilon}_t$. Therefore,

[6] Precisely, the externality is

$$E_t \left[Q_{t,t+1} \frac{ds_{t+1}}{dn_t} f(\tilde{\varepsilon}_{t+1}) (\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1})) \right] = -(1-\lambda) E_t \left[Q_{t,t+1} f(\tilde{\varepsilon}_{t+1}) (\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1})) \right]$$

Having one more employed worker today reduces the number of hires tomorrow by $(ds_{t+1}/dn_t) f(\tilde{\varepsilon}_{t+1}) = -(1-\lambda) f(\tilde{\varepsilon}_{t+1})$. Each hire less tomorrow means less output in the amount of $K + \tilde{\varepsilon}_{t+1}$ (the opportunity cost of training a worker tomorrow) but also less training costs, which equal $K + H(\tilde{\varepsilon}_{t+1})$ on average. Since the marginal hire has higher training costs than the average hire, $\tilde{\varepsilon}_{t+1} - H(\tilde{\varepsilon}_{t+1}) \geq 0$, the hiring externality is negative.

the efficient wage is always positive. Then, we can set unemployment benefits $0 < b_t < w_t$ such that full participation is satisfied.

What kind of wage determination mechanism would give rise to efficient wage rule (10)? Clearly, Nash bargaining would not.[7] While potentially interesting, the relation between wage setting and labor market efficiency is not the purpose of the present study. Therefore, in the remainder of this section as well as in the next section we assume wages are set such that job creation is efficient, without explicitly specifying the wage determination mechanism. In the numerical analysis in section 6 we use a simple wage rule of the form $w_t = w(y_t, b_t)$ and verify that under this assumption, although the equilibrium is no longer efficient, our results change very little.

4.3 Job-Finding Rate

We now have a condition for the hiring threshold $\tilde{\varepsilon}_t$ in equilibrium (9), which under the efficient wage setting rule (10) reduces to the efficient job creation condition (6). 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}.

$$f_{it} = \begin{cases} 1 \text{ if } \varepsilon_{it} \leq \tilde{\varepsilon}_t \\ 0 \text{ if } \varepsilon_{it} > \tilde{\varepsilon}_t \end{cases} \tag{11}$$

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,

$$f(\tilde{\varepsilon}_t) = \frac{\int_{-\infty}^{\infty} f_{it} s_{it} dG}{\int_{-\infty}^{\infty} s_{it} dG} \tag{12}$$

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.

4.4 Random Hiring

In order to evaluate the integrals in (12), we need to know in a given period t how many workers of each type ε_{it} are looking for a job.[8] 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 get 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

[7] For a similar setup where Nash bargaining does not achieve efficiency, see Chugh and Merkl (2016).

[8] 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.

force, $s_{it} = s_t$. In this case, the aggregate job-finding rate equals the probability that training costs are below the hiring threshold.

$$f^{\mathrm{RH}}\left(\tilde{\varepsilon}_t\right) = \frac{\int_{-\infty}^{\tilde{\varepsilon}_t} 1 \cdot s_t \cdot dG + \int_{\tilde{\varepsilon}_t}^{\infty} 0 \cdot s_t \cdot dG}{\int_{-\infty}^{\infty} s_t \cdot dG} = G\left(\tilde{\varepsilon}_t\right) \tag{13}$$

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

$$\bar{u}^{\mathrm{RH}} = \frac{\lambda\left[1 - G\left(\tilde{\varepsilon}\right)\right]}{\lambda\left[1 - G\left(\tilde{\varepsilon}\right)\right] + G\left(\tilde{\varepsilon}\right)} \tag{14}$$

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 worker had lower training costs in the past and are currently still employed because they were hired then.

4.5 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.

$$K + \frac{k}{q_t} = y_t - w_t + (1-\lambda)\, E_t\left[Q_{t,t+1}\left(K + \frac{k}{q_{t+1}}\right)\right] \tag{15}$$

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 (15) equals the job creation condition in our model (9) if $k/q_t = \tilde{\varepsilon}_t$. The vacancy filling probability q_t in this 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

$$f_t = \left(\frac{\tilde{\varepsilon}_t}{k}\right)^{\frac{1-\mu}{\mu}} \tag{16}$$

Comparing expression (16) to (13), 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 also the same in both models, 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.

4.6 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.

$$f^{\mathrm{SH}}(\tilde{\varepsilon}) = \frac{\int_{-\infty}^{\tilde{\varepsilon}} 1 \cdot \lambda \cdot dG + \int_{\tilde{\varepsilon}}^{\infty} 0 \cdot 1 \cdot dG}{\int_{-\infty}^{\tilde{\varepsilon}} \lambda \cdot dG + \int_{\tilde{\varepsilon}}^{\infty} 1 \cdot dG} = \frac{\lambda G(\tilde{\varepsilon})}{\lambda G(\tilde{\varepsilon}) + 1 - G(\tilde{\varepsilon})} \tag{17}$$

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

$$\bar{u}^{\mathrm{SH}} = 1 - G(\tilde{\varepsilon}) \tag{18}$$

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 (13) and (17) for the job-finding rate and (14) and (18) 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.[9]

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 (17) is different from the one under random hiring (13), it is clear that condition (16) 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 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

[9] To see this, note that the elasticity of the job finding rate with respect to productivity y_t from equations (13) and (17) 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.

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 equilibrium consumption choices of workers and compare the resulting consumption allocation to the efficient allocation. In order to obtain simple, easily interpretable expressions, we evaluate the model without aggregate shocks. We show that, under these assumptions, 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 this point in the simplest possible setting. Section 6 presents a numerical analysis to support the results of this section in a more general version of the model.

5.1 Equilibrium Consumption

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.

$$\frac{c_{it}}{c_{it+1}} = \frac{c_t}{c_{t+1}} \text{ for all } i \text{ and } t \tag{19}$$

If we further assume that there are no aggregate shocks, aggregate consumption is constant over time, so that individual consumption is constant over time for all individuals as well.

$$c_{it} = c_i \text{ for all } i \text{ and } t \tag{20}$$

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,

$$c_i = E_0 m_{it} = u_i b + (1 - u_i)\, w + \pi \tag{21}$$

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.

Without aggregate shocks, 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 (21) to efficiency condition (5), we see that the consumption allocation is efficient, if and only if unemployment risk is distributed evenly across workers.

5.2 Random versus Selective Hiring

The worker-specific unemployment risk depends crucially on whether training costs ε_{it} are transitory or permanent. 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. 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.

$$u_i^{\mathrm{SH}} = \begin{cases} 0 \text{ if } \varepsilon_i \leq \tilde{\varepsilon} \\ 1 \text{ if } \varepsilon_i > \tilde{\varepsilon} \end{cases} \tag{22}$$

Substituting (21) and individual unemployment risk (22) into the welfare function (4), and dropping the expectation operators because there is no aggregate risk, we get welfare under selective and random hiring.

$$\mathcal{W}^{\mathrm{SH}} = u\,\mathcal{U}\,(b+\pi) + (1-u)\mathcal{U}\,(w+\pi) \leq \mathcal{U}\,(u\,(b+\pi) + (1-u)\,(w+\pi)) = \mathcal{W}^{\mathrm{RH}} \tag{23}$$

Since 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. We analyze this issue formally in the next section, using unemployment insurance policy as an 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 very 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 from the intertemporal insurance motive typically 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 whether the result holds in a fully specified setup with conventional assumptions on the wage determination mechanism. 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_t to maximize welfare (4). We assume the government needs to run a balanced budget, so that the government budget constraint is given by

$$(1 - n_t)\, b = \tau_t \tag{24}$$

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.

In addition to its budget constraint, the government also takes the optimality conditions for job creation (9) and consumption allocation (19), the market clearing conditions for the labor market (2) and goods market (3), and an equilibrium wage setting rule 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. Appendix A.3 proves convergence and derives the steady state equilibrium conditions.

In steady state, the Ramsey planner chooses b and τ to maximize

$$\int_{-\infty}^{\infty} \mathcal{U}\left(c_{it}\right) dG \tag{25}$$

subject to the government budget constraint (24) and the steady state equilibrium conditions of the model. These steady state equilibrium conditions include the steady state job creation equation,

$$K + \tilde{\varepsilon} = \frac{y - w}{1 - \beta(1 - \lambda)} \tag{26}$$

the steady state labor market clearing condition

$$n = \frac{f(\tilde{\varepsilon})}{\lambda + (1 - \lambda) f(\tilde{\varepsilon})} \tag{27}$$

and the consumption rule, $c_i = c = (1 - n) b + nw + \pi - \tau$ under random hiring and $c_i = c_u = b + \pi - \tau$ if i is unemployed and $c_i = c_n = w + \pi - \tau$ if i is employed under selective hiring. Aggregate consumption equals c under random hiring and $c = (1 - n) c_u + nc_n$ under selective hiring and must satisfy the aggregate resource constraint (3),

$$c = yn - [1 - (1 - \lambda) n] f(\tilde{\varepsilon}) (K + H(\tilde{\varepsilon})) \tag{28}$$

which pins down the level of profits π.

The final constraint on the Ramsey problem of the government is a wage setting rule. In sections 4 and 5, we assumed wages are set such that job creation is efficient. Here, we deviate from that assumption because such a wage setting rule is probably not realistic and we want to explore whether our results hold up to more standard wage determination mechanism. More importantly, we want to think of the government facing a trade-off between efficient consumption redistribution and efficient job creation. If wages are always set efficiently and do not depend on the level of unemployment insurance, such a trade-off does not exist. Thus, we specify an ad-hoc wage rule, which is loosely inspired by the surplus sharing rule for wages in standard search and matching models.

$$w = \phi y + (1 - \phi) b \tag{29}$$

According to this rule, higher unemployment benefits improve workers' outside option, which drives up their wage and erodes profits, thus discouraging job creation.

6.2 Random Hiring

The first order conditions for the Ramsey problem are straightforward to derive, but messy and hard to interpret. Therefore, we present the results for the the optimal unemployment insurance policy numerically for a calibrated version of the model. This exercise is meant to be illustrative rather than quantitative. For the numerical results, we use logarithmic utility over consumption $U(c) = \log c$, normalize productivity to $y = 1$, assume no discounting, $\beta = 1$, set the separation rate to $\lambda = 0.1$ per quarter and use a uniform distribution for the idiosyncratic component of training costs under random hiring, $G = U[-4, 4]$. The distribution function under selective hiring is recalibrated such that equation (17) is the same as equation (13), see the last paragraph of section 4.6 for details. We calibrate the mean training costs $K = 1$ to target a job-finding rate of 0.5, which corresponds to an unemployment rate of 9%. In our baseline calibration, we set the parameter of the wage setting rule to $\phi = 0.75$.

Under random hiring, the government can achieve the first best allocation. The reason is that there is no trade-off in this case. Since unemployment risk is equally distributed across all workers, no redistribution is necessary. Moreover, since all workers are unemployed an equal amount of time, unemployment benefits are not redistributive. Thus, unemployment benefits are irrelevant for the consumption allocation and the government can set b in order to implement the efficient level of job creation.

The solid line in figure 1 shows welfare as a function of unemployment benefits for the model with random hiring. The optimal amount of unemployment benefits, which implements the efficient amount of job creation, equals 0.6, which by equation (29) implies a wage of 0.9, so that the earnings of unemployed workers are two thirds of those of employed workers. Of course this result is sensitive to the parameterization. In figure 2 we explore an alternative calibration, setting the parameter of the wage setting rule to $\phi = 0.25$. In this case, efficient job creation is achieved by setting unemployment benefits to 0.87, which also implies a wage of 0.9 but a replacement ratio of 96%.

6.3 Selective Hiring

In the model 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. Therefore, in this case the Ramsey planner cannot replicate the first best allocation, and we would expect optimal unemployment benefits to be higher than under random hiring, resulting in an inefficiently low level of employment.

The dashed line in figures 1 and 2 shows welfare as a function of unemployment benefits for the model with selective hiring. In both calibrations, the maximum level of welfare that can be reached under selective hiring is lower than under random hiring, because unemployment benefits distort job creation. Under our baseline calibration in figure 1, the optimal level of unemployment benefits under selective hiring is much higher than under random hiring, around 0.7 compared to 0.6. Since the wage, by equation (29), is also lower under selective hiring, the difference in the replacement ratio between the two models is small.

In the calibration in figure 2, the wage is more sensitive to the level of unemployment benefits, so that unemployment insurance is more distortionary in terms of job creation. As a result, the optimal unemployment benefits in that case are only slightly higher than under random hiring. It is even possible to construct examples, in which optimal unemployment benefits are lower under selective than under random hiring. This happens if job creation is very sensitive to the level of unemployment benefits. In this case, lowering unemployment benefits slightly increases employment by a lot. And with employment close to full employment, the motive for redistribution disappears, so that the cost of lower unemployment benefits disappears as well.

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 are completely different for selective versus random hiring. With random hiring, as in the standard model, the welfare costs of unemployment are small. With selective hiring, unemployment risk is distributed unequally across workers. Therefore, the welfare costs of unemployment are much larger in this case.

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. Under selective hiring, 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 higher, and employment and welfare are lower than under random hiring.

A Appendices

A.1 Social Planner Problem

The value function and the Bellman equation of the social planner problem (4) are given by

$$V\left(n_{t-1}; y_t\right) = \max_{\left\{\tilde{\varepsilon}_{t+\tau}, \left\{c_{it+\tau}\right\}_{i=-\infty}^{\infty}\right\}_{\tau=t}^{\infty}} E_t \sum_{\tau=0}^{\infty} \beta^{\tau} \int_{-\infty}^{\infty} \mathcal{U}\left(c_{it+\tau}\right) dG \tag{30}$$

$$= \max_{\tilde{\varepsilon}_t, \left\{c_{it}\right\}_{i=-\infty}^{\infty}} \left\{ \int_{-\infty}^{\infty} \mathcal{U}\left(c_{it}\right) dG + \beta E_t V\left(n_t; y_{t+1}\right) \right\} \tag{31}$$

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

$$n_t = (1-\lambda)\left(1 - f\left(\tilde{\varepsilon}_t\right)\right) n_{t-1} + f\left(\tilde{\varepsilon}_t\right) \tag{32}$$

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).

$$\int_{-\infty}^{\infty} c_{it} dG = y_t n_t - \left[1 - (1-\lambda) n_{t-1}\right] f\left(\tilde{\varepsilon}_t\right)\left(K + H\left(\tilde{\varepsilon}_t\right)\right) \tag{33}$$

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

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

$$\mathcal{U}'\left(c_{it}\right) = \mu_t \tag{34}$$

a first order condition for $\tilde{\varepsilon}_t$

$$y_t - K - H\left(\tilde{\varepsilon}_t\right) - \frac{H'\left(\tilde{\varepsilon}_t\right) f\left(\tilde{\varepsilon}_t\right)}{f'\left(\tilde{\varepsilon}_t\right)} + \frac{\beta E_t\left[V'\left(n_t; y_{t+1}\right)\right]}{\mu_t} = 0 \tag{35}$$

an envelope condition for n_{t-1}

$$\begin{aligned} V'\left(n_{t-1}; y_t\right) &= (1-\lambda) \mu_t \left\{\left(1 - f\left(\tilde{\varepsilon}_t\right)\right) y_t + f\left(\tilde{\varepsilon}_t\right)\left(K + H\left(\tilde{\varepsilon}_t\right)\right)\right\} \\ &\quad + (1-\lambda)\left(1 - f\left(\tilde{\varepsilon}_t\right)\right) \beta E_t V'\left(n_t; y_{t+1}\right) \end{aligned} \tag{36}$$

and the aggregate resource constraint itself. These conditions define a system of $3+M$ first-order expectational difference equations in the variables c_{it}, $\tilde{\varepsilon}_t$, μ_t and $V'\left(n_{t-1}; y_t\right)$, where $M \to \infty$ is the number of consumers i in the economy.

The first order condition for c_{it} immediately implies that $c_{it} = c_t$. 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 $\int_{-\infty}^{\infty} c_{it} dG = c_t \int_{-\infty}^{\infty} dG = c_t$.

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

$$V'\left(n_{t-1}; y_t\right) = (1-\lambda) \mu_t \left[K + H\left(\tilde{\varepsilon}_t\right) + \left(1 - f\left(\tilde{\varepsilon}_t\right)\right) \frac{H'\left(\tilde{\varepsilon}_t\right) f\left(\tilde{\varepsilon}_t\right)}{f'\left(\tilde{\varepsilon}_t\right)}\right] \tag{37}$$

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

$$K + M(\tilde{\varepsilon}_t) = y_t + \beta(1-\lambda) E_t \left[\frac{\mu_{t+1}}{\mu_t} \left\{ K + M(\tilde{\varepsilon}_{t+1}) - f(\tilde{\varepsilon}_{t+1}) \left(M(\tilde{\varepsilon}_{t+1}) - H(\tilde{\varepsilon}_{t+1}) \right) \right\} \right] \tag{38}$$

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

$$M(\tilde{\varepsilon}_t) = H(\tilde{\varepsilon}_t) + \frac{H'(\tilde{\varepsilon}_t) f(\tilde{\varepsilon}_t)}{f'(\tilde{\varepsilon}_t)} \tag{39}$$

To evaluate $M(\tilde{\varepsilon}_t)$, we need the distribution of training costs ε_{it} in the pool of job seekers. Denote the cumulative distribution function of this distribution by F.[10] Then, average training costs are given by the following expression.

$$H(\tilde{\varepsilon}_t) = \frac{\int_{-\infty}^{\tilde{\varepsilon}_t} \varepsilon dF(\varepsilon)}{F(\tilde{\varepsilon}_t)} = \frac{\int_{-\infty}^{\tilde{\varepsilon}_t} \varepsilon dF(\varepsilon)}{f(\tilde{\varepsilon}_t)} \tag{40}$$

Note that job seekers with training costs below $\tilde{\varepsilon}_t$ are hired, so that $F(\tilde{\varepsilon}_t) = f(\tilde{\varepsilon}_t)$. Taking a derivative with respect to the training costs of the marginal hire $\tilde{\varepsilon}_t$, we get,

$$H'(\tilde{\varepsilon}_t) = \frac{\tilde{\varepsilon}_t f'(\tilde{\varepsilon}_t)}{f(\tilde{\varepsilon}_t)} - \frac{\int_{-\infty}^{\tilde{\varepsilon}_t} \varepsilon dF(\varepsilon)}{f(\tilde{\varepsilon}_t)^2} f'(\tilde{\varepsilon}_t) = \frac{f'(\tilde{\varepsilon}_t)}{f(\tilde{\varepsilon}_t)} \left[\tilde{\varepsilon}_t - H(\tilde{\varepsilon}_t) \right] \tag{41}$$

so that $M(\tilde{\varepsilon}_t) = \tilde{\varepsilon}_t$. Substituting into Euler equation (38) gives the efficient job creation condition (6) in the main text.

A.2 Equilibrium Job Creation

The representative firm chooses $\tilde{\varepsilon}_t$ and n_t in each period, in order to maximize the expected net present value of its profits,

$$E_0 \sum_{t=0}^{\infty} Q_{0,t} \left[(y_t - w_t) n_t - f(\tilde{\varepsilon}_t) s_t (K + H(\tilde{\varepsilon}_t)) \right] \tag{42}$$

where $Q_{0,t}$ is the stochastic discount factor as in equation (7), subject to the law of motion for its employment stocks (2),

$$n_t = (1-\lambda) n_{t-1} + f(\tilde{\varepsilon}_t) s_t \tag{43}$$

where $f(\tilde{\varepsilon}_t) s_t$ the number of new hires in period t. Since each firm is small compared to the overall size of the economy, it takes wages w_t, the stochastic discount factor $Q_{t,t+1}$ and the total number of job seekers s_t as given.

From the Bellman equation

$$V(n_{t-1}; y_t) = (y_t - w_t) n_t - f(\tilde{\varepsilon}_t) s_t (K + H(\tilde{\varepsilon}_t)) + E_t \left[Q_{t,t+1} V(n_t; y_{t+1}) \right] \tag{44}$$

[10] Under perfectly random hiring, the distribution of training costs among job seekers equals the unconditional distribution of training costs, so that $F = G$, but in general this need not be the case. In section 4.6 we explore how F relates to G under perfectly selective hiring. For now, the only thing that matters is that the distribution F exists.

where

$$n_t = (1-\lambda)\, n_{t-1} + f(\tilde{\varepsilon}_t)\, s_t \tag{45}$$

we get the first order condition for $\tilde{\varepsilon}_t$

$$K + H(\tilde{\varepsilon}_t) + \frac{f(\tilde{\varepsilon}_t)\, H'(\tilde{\varepsilon}_t)}{f'(\tilde{\varepsilon}_t)} = y_t - w_t + E_t\left[Q_{t,t+1} V'(n_t; y_{t+1})\right] \tag{46}$$

and the envelope condition for n_{t-1}

$$V'(n_{t-1}; y_t) = (1-\lambda)\left\{y_t - w_t + E_t\left[Q_{t,t+1} V'(n_t; y_{t+1})\right]\right\} \tag{47}$$

Substituting the first order condition into the envelope condition

$$V'(n_{t-1}; y_t) = (1-\lambda)\left\{K + H(\tilde{\varepsilon}_t) + \frac{f(\tilde{\varepsilon}_t)\, H'(\tilde{\varepsilon}_t)}{f'(\tilde{\varepsilon}_t)}\right\} \tag{48}$$

and substituting back into the envelope condition, gives an Euler equation for the equilibrium hiring threshold. Finally, substituting $M(\tilde{\varepsilon}_t) = \tilde{\varepsilon}_t$, see appendix A.1, gives the equilibrium job creation equation (9) in the main text.

A.3 Equilibrium without Aggregate Shocks

Without aggregate shocks, equilibrium consumption equals $c_i = c = (1-n_t)\, b + n_t w + \pi - \tau_t$ under random hiring and $c_i = c_u = b + \pi - \tau_t$ if i is unemployed and $c_i = c_n = w + \pi - \tau_t$ if i is employed under selective hiring, see equation (21). Notice that without aggregate shocks, the wage w is time-invariant by equation (29). Profits π, which we assume are redistributed lump-sum from firms to workers, are pinned down by the aggregate resource constraint (3),

$$c = y n_t - \left[1 - (1-\lambda)\, n_{t-1}\right] f(\tilde{\varepsilon}_t)\left(K + H(\tilde{\varepsilon}_t)\right) \tag{49}$$

where aggregate consumption $c = (1-n)\, c_u + n c_n$ under selective hiring. Notice that since aggregate consumption is constant over time, the stochastic discount factor simplifies to $Q_{t,t+1} = \beta \mathcal{U}'(c_{t+1}) / \mathcal{U}'(c_t) = \beta$.

The equilibrium job creation condition (9), in the absence of aggregate shocks, becomes simply

$$K + \tilde{\varepsilon}_t = y - w + \beta(1-\lambda)(K + \tilde{\varepsilon}_{t+1}) \tag{50}$$

and the law of motion for employment (2) is unchanged.

$$n_t = (1-\lambda)(1 - f(\tilde{\varepsilon}_t))\, n_{t-1} + f(\tilde{\varepsilon}_t) \tag{51}$$

Since $0 < \lambda < 1$ and $\beta \leq 1$, these equations are saddle path stable and admit a unique steady state for n_t and $\tilde{\varepsilon}_t$. Imposing $n_t = n$ and $\tilde{\varepsilon}_t = \tilde{\varepsilon}$ for all t produces the steady state equilibrium conditions in the main text.

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] Fernández-Blanco, J. and E. Preugschat (2017). On the Effects of Ranking by Unemployment Duration, Working Paper.

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

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

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

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

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

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

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

[18] Landais C., P. Michaillat and E. Saez (2016). A Macroeconomic Approach to Optimal Unemployment Insurance: Theory, *American Economic Journal: Economic Policy,* forthcoming.

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

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

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

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

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

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

[25] 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.

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

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

[28] 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.

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

Figure 1. Welfare as a function of unemployment benefits, $\phi = 0.75$

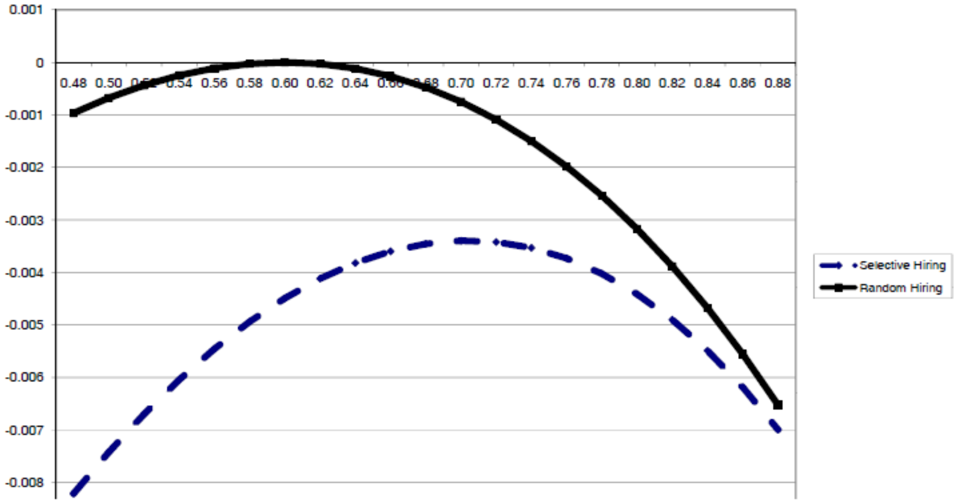

Figure 2. Welfare as a function of unemployment benefits, $\phi = 0.25$