Rational Inattention in Hiring Decisions *

Sushant Acharya　　　　　　Shu Lin Wee

FRB New York　　　　Carnegie Mellon University

August 20, 2016

Abstract

This paper provides an information-based theory of match efficiency. Rationally inattentive hiring firms must expend resources to determine the viability of hiring an applicant. In a recession, firms are more selective and seek to hire and retain more productive workers to compensate for lower aggregate productivity. Stricter standards increase the variability in the pool of unemployed job-seekers, making it harder and costlier for the firm to ascertain the suitability of an applicant. These higher screening costs limit the firms' ability to effectively screen applicants, leading them to accept fewer applicants in order to avoid hiring unsuitable workers. These pro-cyclical acceptance rates form a wedge between meeting and hiring rates and corresponds to changes in match efficiency. Unlike the standard search model which generates counterfactual predictions, our model with rationally inattentive firms can account for fluctuations in measured match efficiency in the data.

Keywords: Rational Inattention, Hiring Behavior, Screening Costs, Match Efficiency, Composition of Unemployed
JEL Codes: D8, E32, J63, J64

1 Introduction

The Great Recession was marked by a severe spike in unemployment rates as well as a tripling in the ratio of unemployed job-seekers for each job opening. Despite this sharp increase in the number of job-seekers per vacancy, employers frequently complained that they were unable to find suitable workers to fill their vacancies.[1] This has led many commentators to argue that match efficiency declined during the Great Recession. In this paper, we provide an information-based theory of match efficiency. In particular, we show how the changing composition of job-seekers over the business cycle affects the cost of screening applicants and the hiring decisions of firms which in turn drives movements in match efficiency over the business cycle.

We consider a standard search and matching model in which workers permanently differ in their ability. A firm's profitability is affected by an aggregate productivity shock, a worker's ability and a match-specific component. A worker's ability is perfectly observable to the worker and to her current employer, but not to a new firm which may want to hire the worker. These new firms can conduct costly interviews to learn about the suitability of the worker and given their acquired information about the job-seeker, reject applicants who are below the bar. Match efficiency is defined as the firm's acceptance rate of a worker and is distinct from the rate at which a firm contacts a worker. The acceptance rate of firms depends on how much information firms acquire as well as the costliness of making a mistake in hiring the wrong worker. In our model, firms are *rationally inattentive* and must pay an entropy-based cost to acquire information about an applicant. More informative interviews are costlier. In equilibrium, information is more expensive when the distribution of job-seeker quality is more uniform and when the firm has less certainty about the type of job-seekers it would encounter.

The distribution of job-seeker quality is itself endogenously changing over the business cycle. When aggregate productivity is high, firms are willing to hire almost all workers except those who are deemed to be very poor matches. As a result, the pool of unemployed job-seekers is largely made up of individuals with very low productivity. This in turn implies that there is very little uncertainty about the type of unemployed job-seeker a firm would meet in a boom. In contrast, recessions are periods where the pool of unemployed job-seekers is more disparate as both luck and selection cause the inflow into unemployment to rise. The decline in aggregate productivity causes firms to release not just lower productivity workers into the unemployment pool but also high productivity workers who drew poor match quality shocks. This increased variation in the pool of unemployed job-seekers implies that a firm has more uncertainty over the type of job-seeker she would meet in a downturn and must expend more resources to learn if the worker is suitable to use for production. Thus, information is more expensive in a recession.

At the same time, firms would like to acquire more information about new hires in a recession.

Given lower aggregate productivity, firms are unwilling to hire an applicant unless they possess high ability or draw a high match-specific shock. As such, they would like to conduct more informative interviews to ensure that they hire a suitable worker. However, since information is more expensive, firms choose less informative signals about job-applicants than they otherwise would in a scenario where there were no increases in the cost of information. Given their less informative signals about the worker, firms err on the side of caution and reject workers more often in the downturn to avoid the costly mistake of hiring the wrong worker. This increase in the rejection rate reduces match efficiency. Moreover, the high rejection rates keeps the firm's uncertainty about the pool of unemployed job-seekers elevated, reinforcing selective hiring and further weighing on match efficiency. Overall, our model can generate a substantial fall in match efficiency in recessions.

The increased selectivity in hiring and retaining workers during recessions is well supported in data. A large literature has argued that the quality of the unemployment pool improves during a recession as better quality workers enter or remain in unemployment during downturns (See for example, Kosovich (2010), Lockwood (1991), Nakamura (2008) and Mueller (2015) among others.[2]). Our paper contributes to the existing literature on selective hiring and argues that selective hiring standards in a recession not only induce a rise in the average quality of the unemployment pool but also cause a corresponding increase in the firm's uncertainty over which type of job-seeker she would meet. It is precisely this increase in the variation of job-seekers during a recession that hampers firms' recruitment efforts during a downturn despite an improvement in the average quality of the unemployment pool.

It is well known that in standard full information search and matching models, hiring does not fall much in recessions resulting in a muted response of unemployment. An improvement in the job-seeker quality makes firms even less inclined to reduce job creation, exacerbating the employment volatility puzzle (Shimer, 2005). Moreover, with costless information, the notion that firms find it difficult to fill a vacancy during a downturn is hard to reconcile with the facts that both the number of unemployed job-seekers, and the quality of the average job-seeker, improve in recessions. Our paper resolves these issues. At the heart of our mechanism is a tight link between the uncertainty in the pool of unemployed job-seekers and the cost of recruiting. Although the average quality of the job-seeker pool increases during recessions, so does the cost of screening a worker. As in Pissarides (2009), these counter-cyclical costs of job creation generate more labor market volatility. Unlike Pissarides (2009) who assumes an exogenous fixed matching cost which renders the effective cost of job creation countercyclical, our model generates these countercyclical costs endogenously by linking the cost of information to the distribution of unemployed job-seekers. In our numerical exercise, a 3% drop in aggregate productivity

---

[2]Other studies such as Mirkin (2016) also study how the composition of the unemployment pool during downturns can lead to jobless recoveries.

causes the unemployment rate to rise by about 5% in the model where firms are rationally inattentive while the full information model with heterogeneous workers only predicts a rise in the unemployment by less than 0.5%. More importantly, our model predicts a decline in match efficiency of about 4.5% in response to the 3% drop in aggregate productivity. In contrast, the full information model observes little to no change in match efficiency as the average quality of the unemployment pool rises during a downturn and counteracts part of the fall in aggregate productivity. All this contributes to a more muted increase in the vacancy yield in the model with rationally inattentive firms relative to what the standard search model would predict, which is akin to what we would observe in the actual data.

The idea that recruiting strategies may change over the business cycle is not a new one. Using data from JOLTS and CPS, Figure 1 replicates the findings as in Davis et al. (2013) and shows how the implied job-filling rate from a standard constant returns to scale matching function with match efficiency assumed to be constant at 1, diverged significantly from its empirical counterpart, the vacancy yield.[3] We add to this graph the computed match efficiency which is the residual variation in hires not accounted for by unemployed job-seekers and vacancies posted. The divergence between the implied job-filling rate from the standard matching function and the vacancy yield coincides with the fall in computed matching efficiency. An influential paper by Davis et al. (2012) suggests that the divergence between the two rates and the resulting fall in computed match efficiency is due to changes in recruiting intensity by firms. Recruiting intensity - defined as a catch-all term for the other instruments and screening methods firms use to increase their rate of hires - fell dramatically during the Great Recession and remained depressed long after GDP recovered. This decline in recruiting intensity has been cited as a factor behind the drag on hiring rates. In this paper, we offer a theory of recruiting intensity. Firms can expend resources to reduce the uncertainty about an applicant's suitability. Their optimal choice of this expenditure varies over the business cycle generating cyclical movements in recruiting intensity.

Several recent papers have also tried to examine and decompose the forces driving the decline in match efficiency. Gavazza et al. (2014) consider how financial frictions, firm entry and exit together with the firm's choice of recruiting intensity can account for the drop in match efficiency. Closely related to our paper, Sedlacek (2014) considers a full-information model in which firms are differentially selective over the business cycle due to the presence of firing costs. Barnichon and Figura (2015) focus on how the composition of job-seekers (in terms of short and long term unemployed) and dispersion in local labor market conditions can help explain the variation in matching efficiency over time. In contrast, our proposed mechanism offers insight as to how the changes in the dispersion of the unemployment pool caused hiring rates to stall and vacancy yields to falter despite the large number of job-seekers available for each vacancy.

Our paper also speaks to a large literature which argues that firms use unemployment duration

---

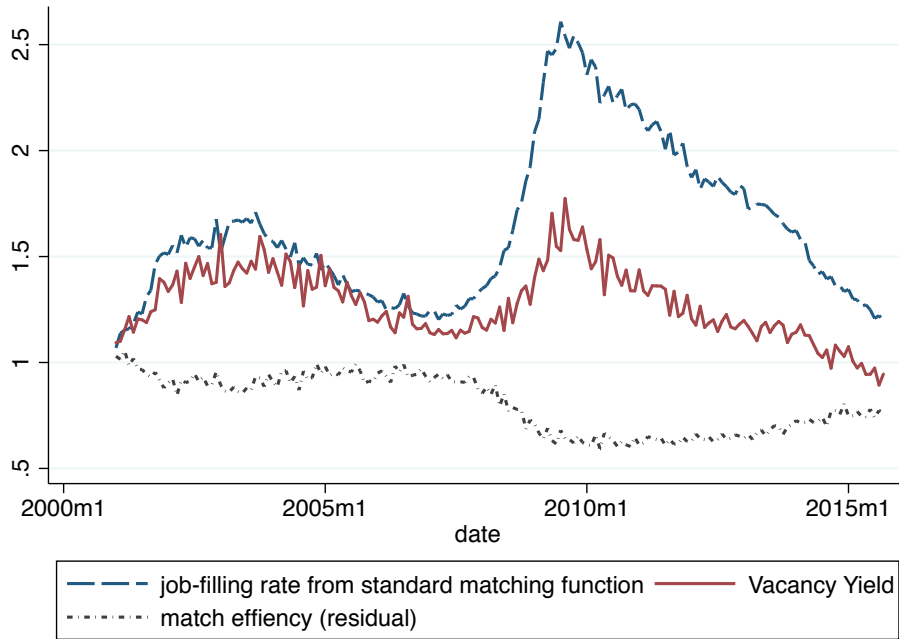[3] The vacancy yield is defined as the ratio of hires to vacancies.

Figure 1: Vacancy Yield and Match Efficiency as computed from JOLTS and CPS

as an additional tool to evaluate the suitability of a worker. Some recent papers such as Kroft et al. (2013), Eriksson and Rooth (2014) and Oberholzer-Gee (2008) among others use resume audit studies and find evidence of negative duration dependence. Firms are less likely to call back workers who have been unemployed for longer. Our model is able to generate the same behavior - in equilibrium, a worker has a lower probability of being employed the longer her unemployment spell. Under some level of screening, unemployment duration conveys some information about the ability of a worker. Given the information they've acquired from screening, firms hire workers who they perceive to be high productivity and reject workers who they perceive to be of lower productivity. As a result, workers who observe high levels of unemployment duration are likely to be those who have been rejected repeatedly.

Unemployment duration as an indicator of worker quality, however, weakens during a recession. In our model, two forces lengthen the unemployment duration of a worker during a recession. First, firms post fewer vacancies during recessions and this lowers the rate at which workers meet firms. Secondly, faced with higher screening costs, firms accept fewer applicants so as to avoid hiring unsuitable workers. Both these forces lower the job finding rate, causing unemployment duration to be a noisier indicator of the worker's type. This is in line with findings by Kroft et al. (2013) who show that call-back rates exhibit a gentler decline with unemployment duration in areas where economic activity is weaker. Since firms were using unemployment duration to defray the cost of information in our model and since unemployment duration is less well-correlated with workers' type in a downturn, our model is able to qualitatively replicate the gentler decline in the relative job-finding rates of workers across different duration in a recession

relative to a boom. Individuals who are unemployed for more than 6 months are 1 to 2 % less likely to find a job relative to those with one month of unemployment duration in a boom than in a recession. Thus, even though average acceptance rates go down during a recession, firms are less able to distinguish workers by their unemployment duration in a recession than in a boom. By construction, such dynamics are impossible in a full-information model about the workers type since the unemployment duration provides no additional information about a worker.

Our paper also relates to recent literature that examines how rational inattention can affect workers' and firms' search behavior. Cheremukhin et al. (2014) consider how the costliness of processing information can affect how targeted search is and the degree of sorting between firms and workers. While their paper demonstrates how rational inattention can lead to equilibrium outcomes that lie between random matching and directed search, we instead focus on a different question and ask how endogenously time-varying information costs affect firms' hiring behavior. Briggs et al. (2015) consider how rational inattention can rationalize the occurrence of increased labor mobility and participation amongst older workers late in their working life. Because we are focused on firms' hiring behavior, our paper instead considers the information processing problem of the firm as opposed to the worker.

Finally, our paper also relates to the literature that looks at applicant and interview strategies. Recent work by Lester and Wolthoff (2016) shows how the presence of screening costs can affect the allocation of heterogeneous workers to firms of varying productivity. Given a cost of interviewing workers, the authors consider a directed search environment and find that the optimal posted contract must specify both a wage and a hiring policy. Unlike our paper, Lester and Wolthoff (2016) treat the cost of information as given while the endogenous cost of information in our model is key to explaining the evolution of match efficiency over the business cycle.

The rest of this paper is organized as follows: Section 2 introduces the model with rational inattention in an otherwise standard random search framework. Section 3 discusses our calibration approach. Section 4 documents our results while Section 5 concludes.

2   Model

We use a standard Diamond-Mortensen-Pissarides model of labor-market frictions. The model is formulated in discrete time. We describe the economic agents that populate this economy.

Workers   The economy consists of a unit mass of workers. These workers are risk neutral and discount the future at a rate β. Each worker i has a permanent productivity-type given by $z_i \in \mathcal{Z}$. The exogenous and time-invariant distribution of worker-types is given by $\Pi_z(z)$ which has full support over $\mathcal{Z}$. Workers can either be employed or unemployed. All unemployed workers produce $b > 0$ as home-production. Unemployed workers are further distinguished by

their duration of unemployment, denoted by τ.

Firms   We define jobs as a single firm-worker pair. The per-period output of a job is given by the production function $F(a, z, e) = aze$ where a is the level of aggregate productivity and z is the type of the worker and e is a match specific shock. Aggregate productivity a is described by an exogenous mean-reverting stationary process. When a firm and worker meet, they draw match-specific shock $e \in \mathcal{E}$ which is independent of the aggregate state and the worker's type, and which stays constant throughout the duration of the match. All draws of the match-specific shock are i.i.d and drawn from a time-invariant distribution $\Pi_e(e)$. The presence of a match specific shock allows for high productivity workers, i.e. high z type workers, to be deemed as bad matches if they draw a low e shock. Likewise, low productivity workers can still be considered suitable hires so as long they draw a sufficiently high e.

Labor Market   A firm that decides to enter the market must post a vacancy at a cost $\kappa > 0$. The measure of firms in operation at any date t is determined by free-entry. Search is random and a vacancy comes into contact with a worker at a rate q_t. This contact rate depends on the total number of vacancies and job-seekers according to a constant returns to scale matching technology $m(v_t, l_t)$ where v_t is the number of vacancies posted and l_t is the number of job seekers. In our model, job-seekers consist of the unemployed and workers who are newly separated from their job at the beginning of the period. Wages are determined by Nash-Bargaining between the firm and worker. For simplicity, we assume that the firm has all the bargaining power and thus, makes each worker a take-it-or-leave-it wage offer of b every period.

So far the model is identical to a standard Diamond-Mortensen-Pissarides search model and the timing of the model is summarized in Figure 2. As can be seen in the timeline, we deviate from the standard model by assuming that a firm cannot observe the effective productivity, ze, of the applicant at the time of meeting. The firm can, however, choose to expend resources and acquire information both about the worker-productivity z and the match-specific shock e. We refer to this process as an *interview*. We assume that the firm can perfectly identify the worker's type once production has taken place. We allow a firm to fire a worker ex-post if she turns out to be unsuitable for the job. Prior to production, however, the firm has to interview the worker to reduce the uncertainty it faces about the worker's effective productivity ze. Given the information revealed in the interview, the firm decides whether or not to hire a worker. The following sections characterize the hiring strategy of a firm.

2.1   Hiring Strategy of the Firm

Consider a firm that has posted a vacancy knowing the level of aggregate productivity is given by a and the distribution of (z, e) type job-seekers for each duration length τ. The hiring
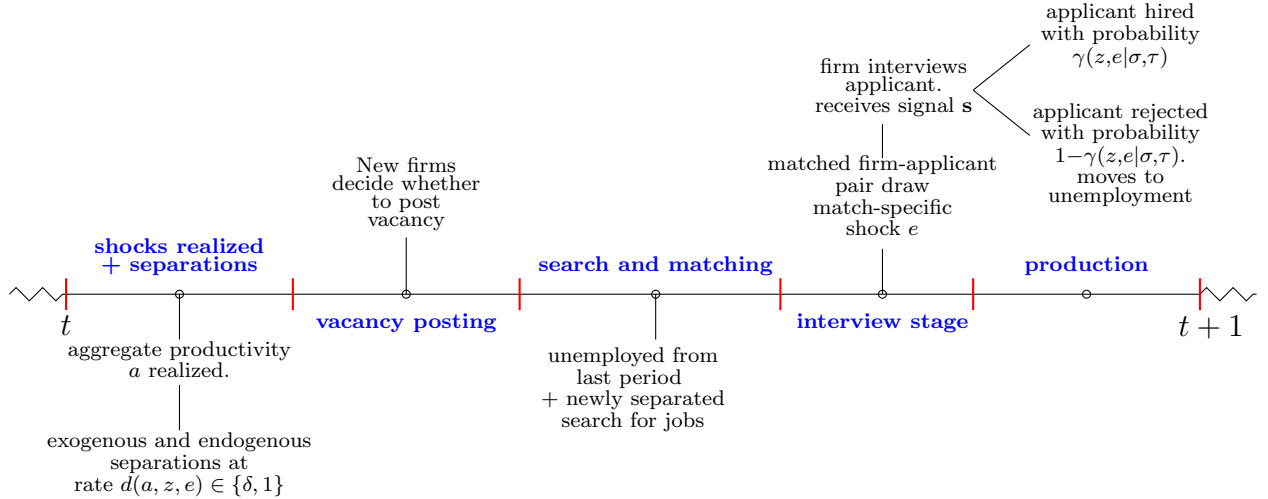
Figure 2: Timeline

strategy of a firm can be described as a two-stage process. (i) In the *first-stage*, given that the firm can observe the applicant's unemployment duration, the firm must devise an information strategy which can be roughly described as specifying how much information the firm would like to process about the worker-type z and the match-specific productivity e. This first stage ends with the firm receiving signals about the workers productivity. (ii) In the *second-stage*, based on the information elicited from the interview, the firm must then decide whether to *reject* or *hire* the applicant. Next, we characterize the firm's hiring strategy starting from the second stage problem.

2.1.1 Second-stage Problem

Let σ denote the set of aggregate state variables of the economy which will be fleshed out later. In the meantime it is sufficient to know that σ contains information about the level of aggregate productivity and the joint distribution of effective productivity (z, e) and unemployment duration τ in the pool of job-seekers. Further denote $G(z, e \mid \sigma, \tau)$ as the conditional distribution of (z, e) types given that the aggregate state is σ and the firm meets a worker of duration τ. Note that $G(z, e \mid \sigma, \tau)$ also describes the firm's prior belief over (z, e) types for each worker of unemployment duration τ.

In the second-stage, the firm has already chosen an information strategy and received signals $\mathbf{s}$ about each type (z, e) applicant who had been unemployed for τ periods prior to meeting the firm. Denote the joint-posterior belief of the firm about this applicant's ability z and match-specific shock e by by $\Gamma(z, e \mid \mathbf{s}, \sigma, \tau)$. Given this posterior belief, the firm's problem is to decide whether to *hire* or *reject* the applicant. If the firm chooses to reject the worker, she gets a payoff of zero. However, depending on the combination of (z, e), the payoff from hiring an applicant can

7

vary. Denote the payoff from hiring an applicant of type z with match-specific shock e (when the aggregate productivity is a) by $\mathbf{x}(a, z, e)$. Since the firm does not observe z or e when meeting the applicant, this payoff is a random variable. The proposition below summarizes the second stage decision problem:

Proposition 1 (Second-Stage Decision Problem of a Firm). *Given the posterior about the applicant $\Gamma(z, e \mid \mathbf{s}, \sigma, \tau)$, the firm hires the applicant iff*

$$\mathbb{E}_\Gamma[\mathbf{x}(a, z, e)] > 0$$

and rejects the applicant otherwise. Thus, the value of such a firm can be written as:

$$J\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) = \max\left\{0, \mathbb{E}_\Gamma\left[\mathbf{x}(a, z, e)\right]\right\}$$

Proof. A firm can always reject a candidate and ensure a payoff of at least 0. Thus, the firm chooses to hire only if the expected payoff from hiring a worker is larger than 0. $\qquad\square$

2.1.2 First-stage Problem

The first stage of the hiring strategy requires the firm to choose an information strategy or the set of signals a firm would like to receive about the applicant's effective productivity. We model costly information processing as an entropy-based cost function as posited in the seminal paper by Sims (2003). In other words, a firm can reduce the uncertainty about the applicant's effective productivity ze by acquiring more information about her. As is standard in the rational inattention literature, we measure uncertainty about the type in terms of entropy and the reduction of uncertainty as the mutual information.

Definition 1. *Consider a random variable $X \in \mathcal{X}$ with prior density $p(x)$. Then the entropy can be written as:*

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \ln p(x)$$

Consider a information strategy under which an agent acquires signals $\mathbf{s}$ about the realization of X. Denote the posterior density of the random variable X as $p(x \mid \mathbf{s})$. Then, the mutual information between the prior and the posterior is given by:

$$I\left(p(x), p(x \mid \mathbf{s})\right) = H(X) - \mathbb{E}_\mathbf{s} H\left(X \mid \mathbf{s}\right)$$

This can be interpreted as a measure of reduction in uncertainty about X by virtue of getting signals $\mathbf{s}$.

From this definition, a choice of the information strategy can be thought of as the firm asking

an applicant a series of questions to reduce its uncertainty about the worker's type. Every additional question provides the firm with incremental information which helps it make a more informed decision about whether to accept or reject an applicant in the second-stage. However, each additional question adds to the cost of processing information. An entropy-based cost function is natural in such a setting as the cost of information is proportional to the expected number of questions needed to implement an information strategy.[4] We are now ready to describe the firm's information strategy and thus its first stage problem.

Recall that conditional on meeting a worker with unemployment duration τ in aggregate state σ, the firm's prior about the workers effective productivity is given by the distribution $G(z, e \mid \sigma, \tau)$. Through the interview, firms can choose to receive signals $\mathbf{s}$ in order to update her belief about worker-productivity z and match-specific shock e of the applicant. More informative signals cost more than less informative ones. The following definition characterizes an information strategy of the firm

Definition 2 (Information Strategy). *The information strategy of a firm who has met a worker with unemployment duration τ (when the aggregate state of the economy is given by σ) is given by a joint distribution of signals $\mathbf{s}$ and types, $\Gamma(z, e, \mathbf{s} \mid \sigma, \tau)$ such that:*

$$G(z, e \mid \sigma, \tau) = \int_{\mathbf{s}} d\Gamma(z, e, \mathbf{s} \mid \sigma, \tau) \tag{1}$$

Equation (1) simply requires that a firm's priors and posteriors are consistent with each other. A consequence of this consistency requirement is that the firm is only free to choose $\Gamma(\mathbf{s} \mid z, e, \sigma, \tau)$. Thus, an information strategy can be thought of as choosing what set of signals a firm chooses to observe when it has met a particular type of worker. Clearly, the most informative set of signals a firm could choose is to select a different signal $\mathbf{s}$ for each (z, e) type worker and perfectly distinguish between workers. However, choosing such a set of signals is very costly. Defining the cost in terms of how much one must pay to reduce uncertainty, we can write down the cost of screening in terms of entropy:

$$c(G, \Gamma \mid \tau) = \lambda \left[H(G(\cdot \mid \sigma, \tau)) - E_s H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) \right]$$

where λ is equal to the marginal cost of reducing uncertainty, $H(G)$ is the firm's initial uncertainty given the distribution $G(z, e \mid \sigma, \tau)$ and $E_s H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right)$ is the firm's residual uncertainty after obtaining signals about the worker. Clearly, if the firm chose signals to have zero residual uncertainty about the worker, she would pay the maximal cost of $\lambda H(G)$. The cost of information for the firm, therefore depends on the distribution of workers $G(\cdot \mid \sigma, \tau)$, and the informativeness of the signals it chooses. The Proposition below summarizes the first-stage problem of the firm.

---

[4]For details, see the coding theorem (Shannon, 1948) and Matejka and McKay (2015).

Proposition 2 (First-Stage Problem of a Firm). *Denote the joint pmf associated with G as g. Then the firm's first-stage problem involves choosing an information strategy to maximize ex-ante payoffs from the second-stage less the cost of information for each potential duration of unemployment τ:*

$$\mathbf{V}(\sigma, \tau) = \max_{\Gamma \in \Delta} \sum_z \sum_e \int_s J\left[\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right] d\Gamma(\mathbf{s} \mid z, e, \sigma, \tau) g(z, e \mid \sigma, \tau) - c(\Gamma, G \mid \tau)] \tag{2}$$

where

$$c(\Gamma, G \mid \tau) = \lambda \left[H\left(G(\cdot \mid \sigma, \tau)\right) - \mathbb{E}_s H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) \right] \tag{3}$$

The firm's first stage problem consists of her ex-ante payoff for each (z, e) worker given signals $\mathbf{s}$, this is given by $J\left[\Gamma\left(\cdot \mid \mathbf{s}, \sigma, \tau\right)\right]$. Since the firm does not know which worker she would meet and therefore which signal she would receive, the firm's payoff is a weighted sum over the signals $d\Gamma\left(\mathbf{s} \mid z, e, \sigma, \tau\right)$ and job-seekers, $g\left(z, e \mid \sigma, \tau\right)$ she encounters.

2.1.3 A Simple Static Model

To understand the mechanism and how the cost of information depends on the distribution of job-seekers, consider the following simple static example. For ease of exposition, we suppress the dependence of our firm's payoffs on e, the match quality shock. By construction, all workers have the same unemployment duration: 0 months in a static model. Suppose for now, the population of job-seekers is exogenously given and is made up of two productivity types $\{\bar{z}, \underline{z}\} \in \mathcal{Z}$. Denote $g(z)$ for $z \in \{\bar{z}, \underline{z}\}$ as the probability mass of type i where $g(\underline{z}) = \alpha$ proportion of job-seekers are type $\underline{z}$ and $g(\bar{z}) = 1 - \alpha$ are of type $\bar{z}$. Further assume that the payoff from hiring a $\bar{z}$ type for a firm yields a payoff of $\mathbf{x}(\bar{z}) > 0 > \mathbf{x}(\underline{z})$.

Consider the problem of a firm that randomly meets a job-seeker. The firm's initial uncertainty about the type of the worker can be quantified in terms of the entropy of her prior which is given by:

$$H(G) = -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha) \tag{4}$$

where G is the distribution associated with the probability mass functions of $g(\cdot)$. It is easy to see that this uncertainty is the greatest when the types are equally represented in population, i.e. where $\alpha = 0.5$. In other words, the flatter the distribution of types, the larger the uncertainty.

Given this distribution of job-seekers in the economy, the firm wants to choose signals such as to maximize her expected payoff from hiring a worker. Consider the following information strategy of the firm. As in the general problem, the firm must choose a conditional distribution

of signals for every given type of worker $\Gamma(s \mid z)$. Suppose that the firm chooses the following information strategy:

$$
\begin{aligned}
\Gamma(s = 1 \mid z = \overline{z}) &= q \in [0.5, 1] \\
\Gamma(s = 0 \mid z = \underline{z}) &= p \in [0.5, 1] \\
\Gamma(s = 1 \mid z = \underline{z}) &= 1 - p \\
\Gamma(s = 0 \mid z = \overline{z}) &= 1 - q
\end{aligned}
$$

A simple interpretation of the above strategy is that the firm would like to see a signal $s = 1$ with probability q and a signal of $s = 0$ with probability $1 - q$ whenever she meets a $\overline{z}$ type applicant. The firm can also choose to receive signals for an applicant of type $z = \underline{z}$. In particular, the firm allows for a signal of $s = 1$ with probability $1 - p$ and a signal of $s = 0$ with probability p whenever type $z = \underline{z}$ applicant matches with her. Notice that if the firm chooses to set $(p, q) = (1, 1)$, the firm can perfectly identify the type of worker on the basis of signals. In contrast, a signal choice of $(p, q) = (0.5, 0.5)$ gives the firm no information since no matter which type she meets, the firm has equal probability of observing a signal of $s = 1$ and a signal of $s = 0$. Thus, the firm can *reduce uncertainty* about the applicant whenever she increases both p and q together.[5] However, choosing p and q away from the combination of $(0.5, 0.5)$ comes at a cost. Under the assumption of an entropy-based cost of reducing uncertainty, one can show that this cost can be expressed as:

$$
\begin{aligned}
c(g, \Gamma) = \lambda \Bigg[&-\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha) \\
&+ q(1 - \alpha) \log \left(\frac{q(1 - \alpha)}{(1 - \alpha)q + \alpha(1 - p)} \right) + \alpha(1 - p) \log \left(\frac{\alpha(1 - p)}{(1 - \alpha)q + \alpha(1 - p)} \right) \\
&+ (1 - \alpha)(1 - q) \log \left(\frac{(1 - q)(1 - \alpha)}{(1 - \alpha)(1 - q) + \alpha p} \right) + \alpha p \log \left(\frac{\alpha p}{(1 - \alpha)(1 - q) + \alpha p} \right) \Bigg]
\end{aligned}
$$

The first line on the RHS of the cost equation represents the firm's initial uncertainty as measured in terms of entropy. The second line on the RHS of the cost equation represents the residual uncertainty the firm has about a worker given that she observes a signal of $s = 1$, multiplied by the probability of observing a signal of $s = 1$. In the same vein, the third line of the cost equation represents the firm's residual uncertainty conditional on seeing a signal of $s = 0$, weighted by the

---

[5] Note that $p = q = 0.5$ implies that a firm would receive a signal of 0 or 1 in a purely random fashion and provide no information about worker ability. A signal of $s = 0$ with $p > 0.5$ instead implies a firm is more likely to get a signal $s = 0$ when the firm meets a $\underline{z}$ worker relative to when she meets a type $\overline{z}$ worker. Note that the symmetry of the problem implies that considering p and q in the range of $[0, 0.5]$ gives us that no information is acquired for $(p, q) = (0.5, 0.5)$ and the firm discerns types perfectly for $(p, q) = (0, 0)$. Thus, decreasing (p, q) together towards $(0, 0)$ also allows the firm to reduce uncertainty with the opposite convention.

probability that she sees a signal of zero. To see how the cost of information changes with the informativeness of the firm's information strategy $\Gamma(s \mid z)$, Figure 3 illustrates how the cost of information varies over p and q for $\alpha = 0.7$, i.e. for an economy where 70% of the population is type $\underline{z}$. As aforementioned, the highest cost of information is attained at $(p, q) = (1, 1)$ where the firm perfectly discerns between the two types of individuals. In contrast, the cost of information is lowest at the points where $(p, q) = (0.5, 0.5)$. As this is the case where the firm essentially gets no information, her cost at this point is zero. In contrast, getting more informative signals about the worker, in the form of increasing (p, q) is coincident with the firm having to incur higher costs to reduce her uncertainty.
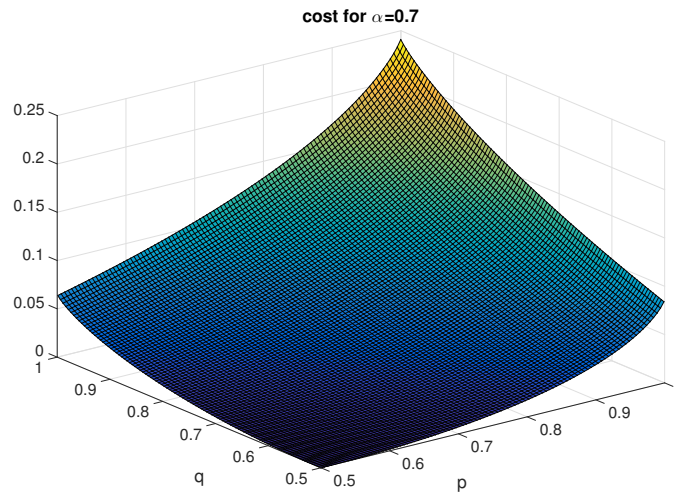

Figure 3: Cost of Information

The cost of information, however, is not just affected by the informativeness of the signals chosen by the firm but is also affected by the distribution of job-seekers. As the distribution of job-seekers becomes more uniform and less skewed towards one type, the cost of information rises as the firm has more initial uncertainty about which type of job-seeker she would encounter. Figure 4 shows how the cost of information rises as the distribution of workers becomes more uniform. Observe that $\alpha \to 0.5$ is the case of maximal uncertainty since prior to getting signals the firm has equal chance of meeting a $\overline{z}$ or a $\underline{z}$ type. Holding fixed $p = 1$, Figure 4 shows that the cost of information is everywhere higher for any choice of $q \geq 0.5$ whenever the distribution of workers becomes more uniform. This is shown by how the black-dashed line for $\alpha = 0.6$ lies everywhere above the blue solid line for $\alpha = 0.7$ for $q \geq 0.5$. Identical results are attained when we hold q fixed at 1 and allow p to vary for $\alpha = 0.7$ and $\alpha = 0.6$.

The cost of information, however, is only one part of the firm's problem. To fully characterize the firm's first-stage problem of choosing signals to identify which worker to hire, we must also consider the benefits associated with such signals. Suppose the payoffs associated with hiring
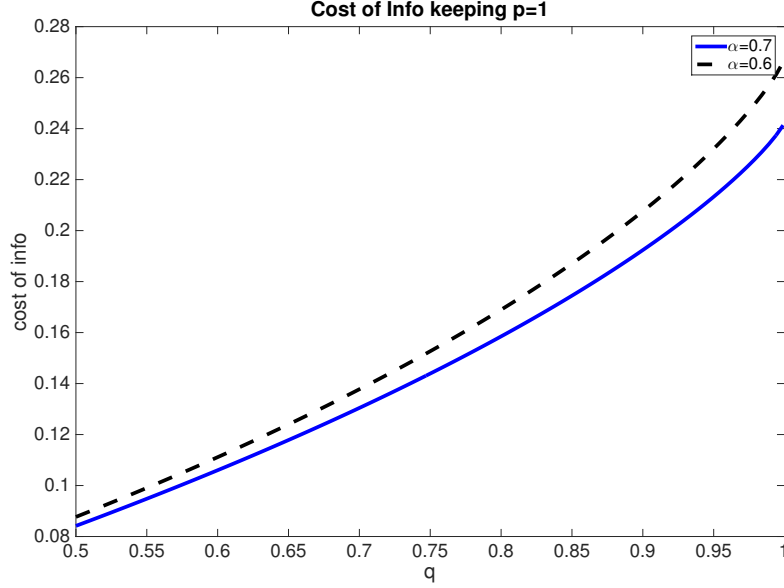
Figure 4: Cost of Information Increases as Distribution becomes more uniform

a $\bar{z}$ type worker gives $x(\bar{z}) = 2$ and the payoffs associated with hiring a $\underline{z}$ type worker gives $x(\underline{z}) = -2$.

Recall that in the second stage, the firm only hires if conditional on the signal she receives, the expected payoffs from the worker are non-negative, i.e. $\mathbb{E}_\Gamma(\mathbf{x}(z) \geq 0)$. The firm's problem in the first stage consists of choosing signals to maximize her ex-ante payoff from the second stage given the distribution of types in the economy and the aggregate state. This firm's value is given by:

$$
\begin{aligned}
V(g) \;=\; & \max_{\Gamma \in \Delta} \underbrace{[(1-\alpha)q + \alpha(1-p)]}_{\text{probability see } s=1} \underbrace{\max\left\{0, (1-\alpha)qx(\bar{z}) + \alpha(1-p)x(\underline{z})\right\}}_{\text{ex-ante payoff from 2nd stage given } s=1} \\
& + \underbrace{[(1-\alpha)(1-q) + \alpha p]}_{\text{probability see } s=0} \underbrace{\max\left\{0, (1-\alpha)(1-q)x(\bar{z}) + \alpha p x(\underline{z})\right\}}_{\text{ex-ante payoff from 2nd stage given } s=0} \\
& - \; c(g, \Gamma)
\end{aligned}
$$

where $p \in [0.5, 1]$ and $q \in [0.5, 1]$. The first line in the firm's value is given by her expected payoff from the second stage given that she observes $s = 1$ multiplied by the probability that she receives a signal $s = 1$. The second line corresponds to the firm's corresponding weighted payoff if she observes a signal of $s = 0$. The last line in the firm's value captures the cost of information as already defined.

Figure 5 shows the optimal decision rules from the firm's problem as α varies (depicted on the horizontal axis). The left panel indicates the optimal choice of p while the right panel indicates the optimal q. Notice that for $\alpha > 0.6$, the cases where more than 60% of the job-seekers are

type $\underline{z}$, the firm optimally chooses to get no information and the cost of information is zero. Under the optimal decision rules and for $\alpha \geq 0.6$, the corresponding expected payoffs if the firm sees a signal of 0 or 1 is negative. Thus, under the optimal decision rules, hiring shuts down for $\alpha \geq 0.6$ as the firm gathers no information and the expected payoffs associated with any signal at this point are negative. In other words, for a large enough mass of low types, firms find it too costly to process information relative to the benefits and instead chooses to not hire anyone.

In contrast, for $\alpha \in [0.5, 0.6)$, the firm optimally chooses to get information about workers. Because the firm chooses $q \to 1$ for $\alpha \in [0.5, 0.6)$, the expected payoffs associated with seeing a signal of $s = 0$ is negative since the firm knows only a type $\underline{z}$ is present when she sees a signal of $s = 0$. In contrast, the expected payoffs associated with seeing a signal of $s = 1$ are strictly positive under these optimal decision rules for $\alpha \in [0.5, 0.6)$. Notice however, that as the distribution flattens and α goes towards 0.5, the firm chooses coarser signals. This is shown by the decline in p from about 0.77 when $\alpha = 0.6$ to $p \approx 0.7$ when $\alpha = 0.5$. Recall that increasing both p and q together allows the firm to better distinguish between worker types. The increase in costs as the distribution flattens, however, causes the firm to optimally choose coarser signals to maximize expected profits.
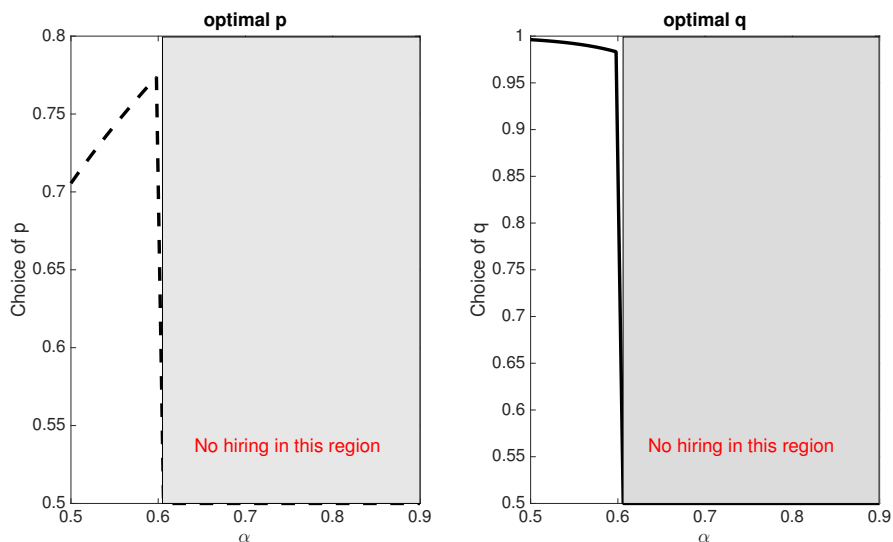

Figure 5: Optimal Decision Rules

2.1.4 Reformulated problem

While the simple model shows how the firm may choose signals under a static environment with two types of workers, the choice of the optimal signal structure, in general, is not easy to characterize. Rather than solving for the optimal signal structure, following Matejka and McKay (2015), we instead solve the identical but transformed problem in terms of choosing

14

state-contingent choice probabilities and the associated payoffs. Let $\mathcal{S}$ be the set of signals that lead the firm to take the action *hire* for an applicant of type (z, e) of duration τ in aggregate state σ. Denote the induced probability of hiring a (z, e)-type worker with unemployment duration τ in aggregate state σ by $\gamma(z, e \mid \sigma, \tau)$. This induced probability is defined as the probability of drawing a signal in $\mathcal{S}$ conditional on being a (z, e) worker with unemployment duration τ:

$$\gamma(z, e \mid \sigma, \tau) = \int_{s \in \mathcal{S}} d\Gamma(s \mid z, e, \sigma, \tau)$$

Similarly, we can define the average probability of hiring a worker of duration τ in state σ as the average induced probability of hiring over the entire pool of job-seekers of that particular unemployment duration:

$$\mathcal{P}(\sigma, \tau) = \sum_z \sum_e \gamma(z, e \mid \sigma, \tau) g(z, e \mid \sigma, \tau)$$

The following Lemma presents the reformulated problem in terms of these choice probabilities:

Lemma 1 (Reformulated First-Stage Problem)**.** *The problem in Proposition 2 is equivalent to the transformed problem below:*

$$\mathbb{V}(\sigma, \tau) = \max_{\{\gamma(z, e \mid \sigma, \tau) \in [0,1]\}} \sum_z \sum_e \gamma(z, e \mid \sigma, \tau) \mathbf{x}(a, z, e) g(z, e \mid \sigma, \tau) - c(\mathcal{P}, G \mid \tau) \qquad (5)$$

subject to:

$$0 \leq \gamma(z, e \mid \sigma, \tau) \leq 1 \ , \ \forall z, e$$

where $c(\mathcal{P}, G \mid \tau)$ denotes the cost associated with acquiring additional signals and can be written as:

$$
\begin{aligned}
c(\mathcal{P}, G \mid \tau) \ = \ \lambda \Bigg[&- \mathcal{P}(\sigma, \tau) \log \mathcal{P}(\sigma, \tau) - [1 - \mathcal{P}(\sigma, \tau)] \log [1 - \mathcal{P}(\sigma, \tau)] \\
&+ \ \sum_z \sum_e \Big(\{ \gamma(z, e \mid \sigma, \tau) \log \gamma(z, e \mid \sigma, \tau) + [1 - \gamma(z, e \mid \sigma, \tau)] \log [1 - \gamma(z, e \mid \sigma, \tau)] \} \\
\times \ \ & g(z, e \mid \sigma, \tau) \Big) \Bigg]
\end{aligned}
$$

Proof. The proof mirrors the one in Appendix A of Matejka and McKay (2015). $\qquad \square$

Intuitively, the convexity of the cost function implies that each action is associated with a particular signal. As firms seek to minimize the cost expended on acquiring information, receiving multiple signals that lead to the same action is inefficient as the additional information acquired

is not acted upon. The firm can economize by instead choosing one signal that induces one action. The transformed problem in Lemma 1 is more tractable than the original problem. The proposition below characterizes the optimal information strategy of a firm.

Proposition 3 (Optimal Information Strategy). *Under the optimal information strategy, the firm chooses a set of signals which induce the firm to hire a worker of productivity-type z and match-specific shock e with unemployment duration τ in aggregate state σ with probability $\gamma(z, e \mid \sigma, \tau)$ which can be written as:*

$$\gamma\left(z, e \mid \sigma, \tau\right) = \frac{\mathcal{P}\left(\sigma, \tau\right) e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 + \mathcal{P}\left(\sigma, \tau\right) \left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda}} - 1\right]} \tag{6}$$

Consequently, the unconditional probability that a firm hires an applicant of duration τ after meeting her is implicitly defined by:

$$1 = \sum_z \sum_e \frac{e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 + \mathcal{P}\left(\sigma, \tau\right) \left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda}} - 1\right]} g(z, e \mid \sigma, \tau) \tag{7}$$

Proof. See Appendix A.1 □

Equation (6) reveals an important feature of the information strategy. Consider two applicants with the same match-specific shock e and duration τ, but with worker-productivity z_1 and z_2 where $z_1 > z_2$. Then the optimal information strategy implies the following:

$$\log \frac{\gamma\left(z_1, e \mid \sigma, \tau\right)}{1 - \gamma\left(z_1, e \mid \sigma, \tau\right)} - \log \frac{\gamma\left(z_2, e \mid \sigma, \tau\right)}{1 - \gamma\left(z_2, e \mid \sigma, \tau\right)} = \frac{\mathbf{x}(a, z_1, e) - \mathbf{x}(a, z_2, e)}{\lambda} \tag{8}$$

The equation above implies that the firm chooses signals such that the induced odds-ratio of accepting a more-productive applicant relative to a less productive applicant is proportional to the difference in the payoffs from hiring the two types of workers. Furthermore, equation (8) implies that the higher the cost of information λ, the less likely a firm is to process information distinguishing different productivity workers. This is reflected in a smaller odds ratio. In the limit as $\lambda \to \infty$, a firm processes no information and the odds ratio tends to zero implying that no applicant is interviewed and has the same chance of getting hired (rejected).

Lemma 2 (Information Strategy with Costless Information). *If information is costless ($\lambda = 0$), the induced probability of hiring a particular type of worker (z, e) with duration τ under the optimal information strategy is given by:*

$$\gamma(z, e \mid \sigma, \tau) = \begin{cases} 1 & if \ \mathbf{x}(a, z, e) \geq 0 \\ 0 & else \end{cases} \tag{9}$$

16

Proof. See Appendix A.2 □

The above Lemma implies that if information is costless, the firm can ascertain the worker-productivity z and match-specific productivity e and this scenario corresponds to the full-information case. In this case, the payoff from hiring an applicant is non-random and the firm accepts an applicant only if $\mathbf{x}(a, z, e) \geq 0$. Interestingly, even with full-information, $\mathcal{P}(\sigma) < 1$ if some applicants have $\mathbf{x}(a, z, e) < 0$. Thus, relative to the standard search and matching model, worker heterogeneity can result in a wedge between the contact rate and the job-filling rate for firms.

2.2 Value of a Firm

With the hiring strategy characterized, all that remains is to close the model. The previous subsections characterized the hiring decisions of a firm conditional on meeting an applicant of type (z, e) with duration τ given $\mathbf{x}(a, z, e)$. Given our assumption that the worker's type is revealed after one period of production, the firm's payoff to hiring a worker of type (z, e), $\mathbf{x}(a, z, e)$, can be written as:

$$\mathbf{x}(a, z, e) = F(a, z, e) - b + \beta \mathbb{E}_{a'|a} \max_{d(a', z, e) \in \{\delta, 1\}} [1 - d(a', z, e)] \mathcal{J}(a', z, e) \tag{10}$$

where δ is the exogenous rate of separation. Since firms learn the worker's productivity perfectly after production, endogenous separations may also occur if after production the value of match cannot be sustained. Notice that the actual payoff to the firm does not depend on an applicant's unemployment duration but depends on her true effective productivity. Let $\mathcal{J}(a, z, e)$ denote the value of a firm that knows the type of her worker when aggregate productivity is a. This value is given by:

$$\mathcal{J}(a, z, e) = F(a, z, e) - b + \beta \mathbb{E}_{a'|a}(1 - d^*(a', z, e)) \mathcal{J}(a', z, e) \tag{11}$$

and d^* can be written as:

$$d^*(a, z, e) = \begin{cases} \delta & \text{if } \mathcal{J}(a, z, e) \geq 0 \\ 1 & \text{if else} \end{cases}$$

2.2.1 Free Entry Condition

The total number of firms that post vacancies in a particular period is determined by a free-entry condition. Each firm posting a vacancy makes zero profits in expectation. This condition pins down the equilibrium market-tightness and hence, the rate at which firms and workers meet.

Denote $g_\tau(\tau \mid \sigma)$ as the probability mass of job-seekers of duration τ given aggregate state σ, i.e., define $g_\tau(\tau \mid \sigma)$ as:

$$g_\tau(\tau \mid \sigma) = \sum_z \sum_e g(z, e, \tau \mid \sigma)$$

Then from the free-entry condition, we have:

$$\kappa \geq q(\theta) \sum_\tau \mathbb{V}(\sigma, \tau) g_\tau(\tau \mid \sigma)$$

$$\left[\kappa - q(\theta) \sum_\tau \mathbb{V}(\sigma, \tau) g_\tau(\tau \mid \sigma) \right] \theta = 0 \tag{12}$$

where $\mathbb{V}(\sigma, \tau)$ denotes the value of a firm from hiring a worker of duration τ net of interview costs and is defined in equation (5). Since we assume random search, it is clear that, θ, the labor market-tightness only depends on the aggregate state as summarized by σ. Further, we can now decompose the job-filling rate into two components. The free entry condition pins down the first component, $q(\theta) = \frac{m(v,l)}{v}$, which is the rate at which a firm meets a job-seeker. We refer to this as the *contact rate*. The second component that affects a firm's hiring rate of a worker of duration τ is given by the firm's acceptance rate, $P(\sigma, \tau)$. Formally, we can now express the aggregate job-filling rate in our model as the product of these two components:

$$\text{Job-filling rate} = \underbrace{q(\theta)}_{\text{contact rate}} \times \underbrace{\sum_\tau \{g_\tau(\tau \mid \sigma) P(\sigma, \tau)\}}_{\text{average acceptance rate}}$$

Notice that there now exists a wedge between the job-filling rate and the rate at which a firm meets a worker. This wedge arises because a firm can choose to reject an applicant after interviewing the applicant.[6] We refer to this wedge given by the firm's average acceptance rate, $\sum_\tau \{g_\tau(\tau \mid \sigma) P(\sigma, \tau)\}$, as our measure of match efficiency. Correspondingly, our measure of the job-finding rate is the product of the rate at which a worker meets a firm, $p(\theta) = \frac{m(v,l)}{l}$, and the average acceptance rate of the firm.

2.3 Composition of the pool of job seekers over the business cycle

Thus far, we have shown how the firm's hiring problem works for a given distribution $G(z, e, \tau)$. The firm's choice of information strategy, crucially depends on the prior distribution of job-seekers $G(z, e \mid \sigma, \tau)$. Thus far, we have not specified how these priors evolve over the business cycle. We address this issue next.

At this point, it is now essential to define the state variables σ for this economy. At any date

---

[6] As explained earlier, this wedge also potentially exists in the model with no information costs because of the presence of worker heterogeneity. The cost of information affects the size and cyclicality of this wedge.

t, the economy can be fully described by $\sigma_t = \{a, n_{t-1}(z, e), u_{t-1}(z, \tau)\}$ where $n_{t-1}(z, e)$ is the measure of employed (z, e) individuals at the end of last period and $u_{t-1}(z, \tau)$ is the measure of unemployed individuals of type z and of duration τ by the end of last period. The aggregate state σ is always known to the firm at the start of each period. Knowing the aggregate state, the firm can always compute the prior distribution of workers who have been unemployed for τ periods. Formally, we define $G(z, e \mid \sigma, \tau)$ as the prior distribution about the effective productivity of a job-seeker who has been unemployed for τ periods in aggregate state σ.

The aggregate laws of motion for each type of worker are known by all firms. In particular, the evolution of the mass of job-seekers of duration τ with worker productivity z in period t can be written as:

$$l_t(z, \tau) = \begin{cases} \sum_e d(a, z, e) n_{t-1}(z, e) & \text{if } \tau = 0 \\ u_{t-1}(z, \tau) & \text{if } \tau \geq 1 \end{cases} \tag{13}$$

The first part of equation (13) refers to job-seekers of type z with zero unemployment duration. These job-seekers of duration zero are the fraction of employed workers at the end of last period, $t-1$, who were either endogenously or exogenously separated from their firms at the beginning of the current period, t. The second line in equation (13) refers to all the unemployed of type z and duration τ at the end of the last period. By construction, all unemployed individuals at the end of a period have duration $\tau \geq 1$. To see this, consider the law of motion for the mass of unemployed individuals with productivity z and duration τ. This is given by:

$$u_t(z, \tau) = l_t(z, \tau - 1) \left\{ 1 - p\left[\theta(\sigma_t)\right] + p\left[\theta(\sigma_t)\right] \sum_e \pi_e(e) \left(1 - \gamma\left[z, e \mid \sigma, \tau - 1\right]\right) \right\}, \forall \tau \geq 1 \tag{14}$$

The first term on the RHS of Equation (14) refers to all job-seekers of duration $\tau - 1$ at the beginning of the period who have productivity z. With probability $1 - p\left[\theta(\sigma)\right]$, a job-seeker of type z and duration $\tau - 1$ fails to meet a firm and remains unemployed. With probability $p\left[\theta(\sigma)\right]$, the worker meets a firm, draws match productivity e with probability $\pi_e(e)$, but is rejected with probability $(1 - \gamma\left[z, e \mid \sigma, \tau - 1\right])$ and remains unemployed. Note if a job-seeker fails to find a job within a period, her duration of unemployment must increase by 1 period. As such, all $l_t(z, \tau - 1)$ job-seekers who fail to be hired by the end of period t form the mass of unemployed $u_t(z, \tau)$ at the end of period t. This is the mass of unemployed job-seekers of type z and duration τ that will carry over into the beginning of period $t+1$, $l_{t+1}(z, \tau)$.

Similarly, we can define the law of motion for the employed of each type (z, e) as:

$$n_t(z, e) = \left[1 - d(a_t, z, e)\right] n_{t-1}(z, e) + p(\theta(\sigma_t)) \pi_e \sum_{\tau=0}^{\infty} \gamma(z, e \mid \sigma, \tau) l_t(z, \tau) \tag{15}$$

The first term on the RHS of Equation (15) are the fraction of employed workers at the end of last period of (z, e) type who are not separated from the firm. Across all durations of job-seekers of type z, a fraction $p(\theta)$ meet a firm and draw match specific productivity e with probability $\pi_e(e)$. Conditional on their duration of unemployment, τ, they are then hired by the firm after the interview with probability $\gamma(z, e \mid \sigma, \tau)$. Finally, we have the accounting identity that the sum of employed and unemployed workers of type z must equal to the total number of workers of type z in the economy.

$$\sum_{\tau} u_t(z, \tau) + \sum_e n_t(z, e) = \pi_z(z)$$

Given the law of motion for the employed and unemployed of each type and duration, we can now construct the probability masses of each type in the economy. Denote $l_t(\tau)$ as the mass of job-seekers of duration τ and l_t as the total mass of job-seekers, i.e.

$$l_t(\tau) = \sum_z l_t(z, \tau)$$

$$l_t = \sum_{\tau} l_t(\tau)$$

Then we can define the probability mass of job-seekers of type z conditional on τ as:

$$g_z(z \mid \sigma, \tau) = \frac{g_{z,\tau}(z, \tau \mid \sigma)}{g_\tau(\tau \mid \sigma)} \equiv \frac{l_t(z,\tau)/l_t}{l_t(\tau)/l_t} = \frac{l_t(z, \tau)}{l_t(\tau)} , \forall \tau \geq 0 \tag{16}$$

where $g_z(z \mid \sigma, \tau)$ is defined simply the share of job-seekers of duration τ who are of type z. Since the match-specific productivity e is drawn independently of z and any past realizations each time a worker matches with a firm, the joint probability mass of drawing a worker of type (z, e) from the pool of job-seekers is simply given by $g_z(z \mid \sigma, \tau)\pi_e(e)$, i.e.

$$g(z, e \mid \sigma, \tau) = g_z(z \mid \sigma, \tau)\pi_e(e)$$

As this is an environment with random search, a firm's prior about any workers type (z, e) given τ is simply given by the joint distribution $G(z, e \mid \sigma, \tau)$. This concludes the description of the model. In the next section, we proceed to discuss the numerical exercises we perform with our model.

3 Numerical Exercise

We discipline the parameters of the model using data on the aggregate flows of workers in the US labor market. The length of a period in our model is a month. Thus, we set $\beta = 0.9967$ which is consistent with an annualized risk free rate of about 4%. We assume that the rate at

which a worker meets a firm $p(\theta)$ takes the form of $p(\theta) = \theta(1 + \theta^\iota)^{-1/\iota}$ which ensures that the probability of a worker meeting the firm is bounded between 0 and 1. We set ι to be 0.5 as standard in the literature.[7] We assume that the production function takes the following form $F(a, z, e) = a \times z \times e$, and that $\log(a)$ follows an AR(1) process:[8]

$$\log a_t = \rho_a \log a_{t-1} + \sigma_a \varepsilon_t \ , \ \varepsilon_t \sim N(0, 1) \tag{17}$$

We set the persistence $\rho_a = 0.983$. We set the standard deviation $\sigma_a = 0.0165$ as in Shimer (2005).

The remaining parameters are chosen to minimize the distance between moments from the simulated data and their empirical counterparts. In particular, we use the following moments to discipline our model. To govern the amount of separations in the economy, we target an employment to unemployment transition rate (EU) of 3.2%. This is in line with the finding in Shimer (2005) where the average tenure of a worker lasts roughly 2.5 years. In the model, we define the EU rate in period t as the share of employed people at the end of $t-1$ who are unemployed at the end of period t. As in Hall (2009) and in Fujita and Moscarini (2013), we set b such that it is equal to 70% of output. Following Jarosch and Pilossoph (2016), we assume that the unobserved worker fixed effect, z, is drawn from a discretized Beta distribution, i.e. $z \sim Beta(A_z, B_z) + 1$ while the match quality shock is drawn from the Beta distribution $e \sim Beta(A_e, B_e)$.[9] Since the vacancy posting cost, κ, the marginal cost of information λ and the parameters governing heterogeneity amongst workers and matches, $\{A_z, B_z, A_e, B_e\}$ affect the rate at which workers find jobs, we use information on the aggregate unemployment rate and the relative job-finding rates across workers of different unemployment duration to govern these parameters. We target an aggregate unemployment rate of about 6.5%, which is the average unemployment rate in the data over the same coverage period as JOLTS.

Resume audit studies suggests that firms use unemployment duration to screen workers and that the observed unemployment duration across workers possesses some information about their underlying productivity. As such, we use data on unemployment duration and unemployment-to-employment transitions (UE) from the Current Population Survey (CPS) over the same period as JOLTS (2000m12 - 2016m4). As in Kroft et al. (2016) and Jarosch and Pilossoph (2016), we conduct a weighted non-linear least squares regression on the relative job-finding rate against

---

[7]See for example Menzio and Shi (2011).

[8]Wherever it is necessary, we approximate the stochastic process of a with a seven-state Markov process using the algorithm specified in Tauchen (1986). In the simulation, we use the continuous process.

[9]Specifically we set the number of worker productivity types to be $n_z = 7$ and the number of match-specific shocks to $n_e = 5$.

unemployment duration of the following form:

$$\frac{\text{UE}(\tau)}{\text{UE}(1)} = \pi_1 + (1 - \pi_1)exp(-\pi_2\tau)$$

where τ is the duration of unemployment, and $\frac{\text{UE}(\tau)}{\text{UE}(1)}$ is the average job-finding rate of an un-employed individual of duration τ relative to an unemployed individual with 1 month of unem-ployment duration. We target this relative job-finding rate in the data and cluster all those who are more than 9 months unemployed into a single bin. The dashed-curve in figure 6 depicts the fitted values of the relative job finding rates by unemployment duration.
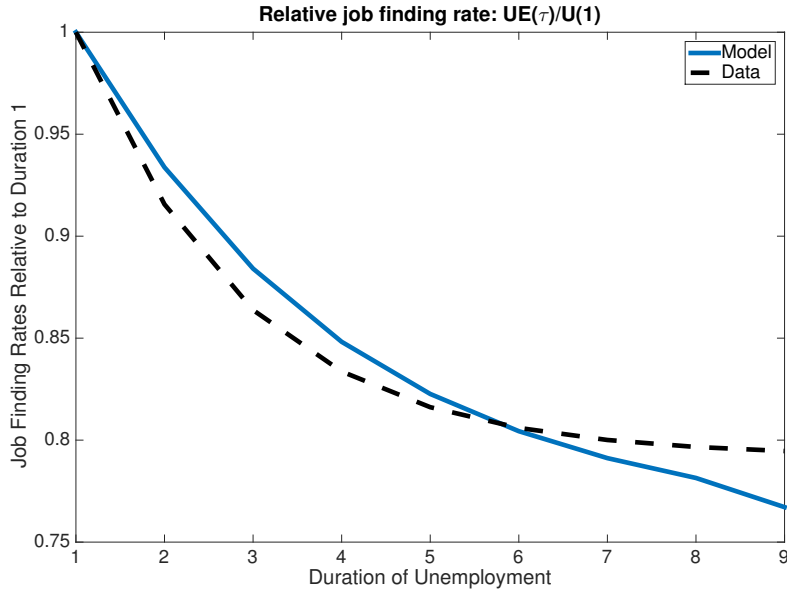

Figure 6: Relative Job finding rates by duration of unemployment

In summary, we have 8 parameters to estimate $\{\lambda, \kappa, \delta, b, A_z, B_z, A_e, B_e\}$ and we target 11 moments: the average monthly separation rate, the aggregate unemployment rate, unemployment benefits worth 70% of output and the relative job-finding rate for unemployment spells from 2 to 9 months. This is an over-identified system and we conduct simulated method of moments (SMM) to back out these parameters. Tables 1 summarizes both the fixed and inferred parameters.

The model is able to match the moments in data very well. Under the parametrization, home production is 69.41% of average output (the target was 70% of average output). The model generates an average EU rate of 3.18%[10] compared to the targeted 3.2%. Also, the model is able to match an average unemployment rate of 6.51 %. Figure 6 shows the estimated relative job-finding rates from the data and the model implied counterpart. The model does a fairly good job at replicating the relative job finding rates but under-predicts the relative job-finding rates

---

[10]This includes both exogenous and endogenous separations

Table 1: Model Parameters

Fixed Parameters

Parameter	Description	Value	Source
β	discount factor	0.9967	annualized interest rate of 4%
σ_a	std. dev. of to agg. productivity	0.0165	Shimer (2005)
ρ_a	autocorr. of agg. productivity	0.983	Shimer (2005)
ι	matching function elasticity	0.5	Menzio and Shi (2010)

Inferred Parameters

Parameter	Description	Value
b	home production	0.4964
δ	exog. separation rate	0.0315
κ	vacancy posting cost	0.0295
λ	marginal cost of information	0.2897
A_z	shape parameter - worker ability	0.105
B_z	shape parameter - worker ability	0.3497
A_e	shape parameter - match productivity	4.4421
B_e	shape parameter - match productivity	6.987

for workers who are long-term unemployed.Finally, while we did not target the aggregate job-finding rate, our model predicts an overall aggregate unemployment to employment transition rate - measured as the share of unemployed last period who transition to employment this period - of about 49%. This is in line with estimates from Shimer (2005) who reports a monthly job-finding rate of about 45%.

The cost of processing information   The estimated value of the marginal cost of information is $\lambda = 0.2897$. This cost is an order of magnitude larger than the cost of vacancy creation, revealing that the data puts relatively less importance to vacancy posting costs with respect to the total cost of job-creation. In order to understand the significance of the magnitude of the cost of information, we calculate the average screening cost (across different duration types condition on meeting a worker) as a fraction of quarterly average wage to compare with numbers established in the literature. Silva and Toledo (2009) report that the average cost from screening and interviewing workers is about 3.6% of the quarterly average wage of a fully productive worker, we calculate the cost of information in terms of quarterly wages as well. The corresponding number in our model is 3.42%.[11]

---

[11]Calculating the cost of information as a fraction of 3 months' worth of wages, $3 \times b$, in our model, this works out to be 3.42% of quarterly wages.

4   Results

4.1   What happens in recessions?

As our first exercise, we simulate a recession as a two standard deviation fall in aggregate productivity relative to steady state and compare the responses of the model with rationally inattentive firms to responses of the full information model. For the full information economy, we set the marginal cost of information to $\lambda = 0$ which allows firms to perfectly observe the applicant's suitability for free.
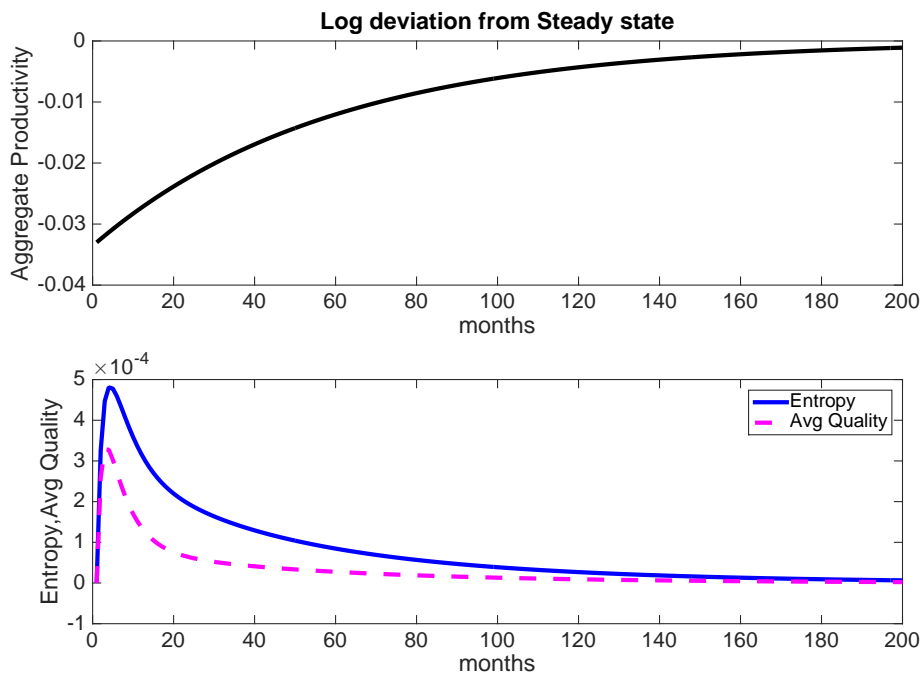

Figure 7: Response in entropy and average quality of unemployed pool to a 2 standard deviation drop in productivity

The top panel of Figure 7 plots the path of aggregate productivity in our experiment while the lower panel depicts how the recession impacts the distribution of unemployed job-seekers in the model of rational inattention.[12] The rise in average quality of unemployed job-seekers stems from firms' stricter standards regarding their retention of workers. Firms terminate some jobs since they now require a higher ability worker to compensate for the fall in aggregate productivity to stay profitable. Thus, recessions force both middle ability workers and high ability workers who were attached to jobs with low match-specific productivity into the pool of unemployed job-seekers. This increase in quality is accompanied by an increase in the uncertainty (measured as entropy) firms face regarding applicant types. In normal times, firms face very little uncertainty

---

[12]The full information model has qualitatively similar responses. However, entropy is inconsequential since firms can perfectly observe the applicant's effective productivity.

regarding the pool of seekers since it is dominated by low-ability types. In contrast, since the composition of the pool of seekers is more varied in a recession, a firm has greater uncertainty about the effective productivity of an applicant. This increase in average quality and uncertainty gradually dissipates over time (lower panel of Figure 7) as firms hire the more suitable applicants.
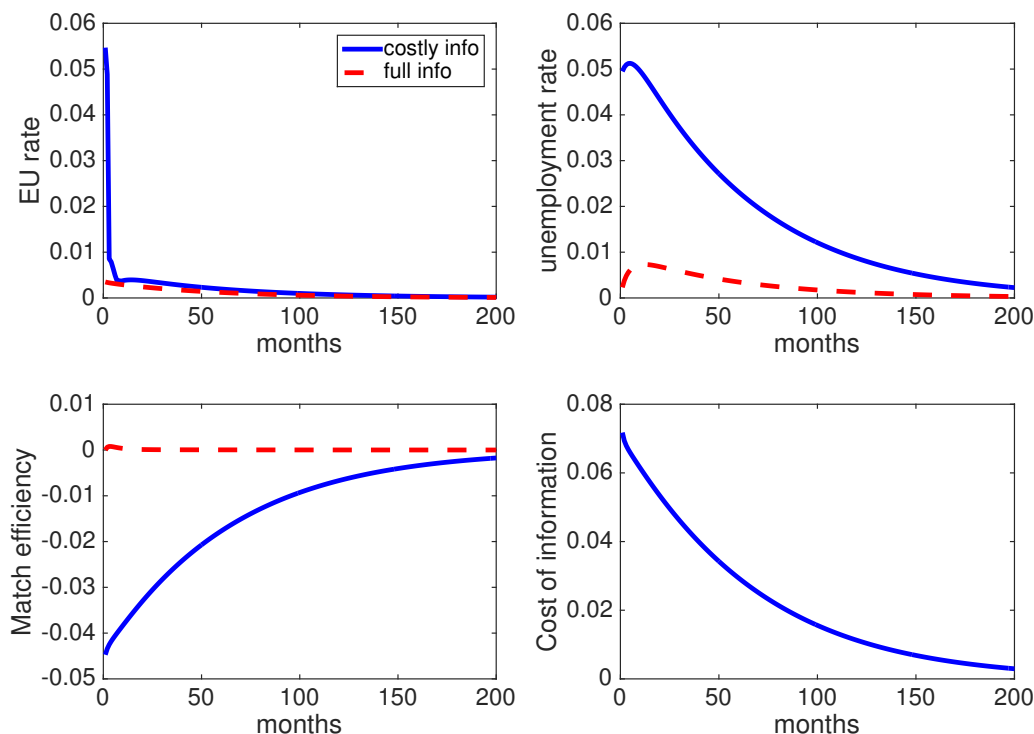

Figure 8: Rise in Cost of Information key to Explaining Decline in Match Efficiency

Noticeably, small changes in uncertainty in the costly information model result in more amplification in terms of labor market flows relative to the full information model. Figure 8 shows the response of unemployment rates, match efficiency and inflows into unemployment from employment to a 2 standard deviation shock to aggregate productivity in both models. Compared to the full information model, the unemployment rate in the model with rationally inattentive firms jumps by about 5.5% on impact whilst the unemployment rate in the full information model increases a minuscule 0.2% on impact. Even though wages do not change with aggregate productivity, the response of unemployment to changes in aggregate productivity is muted in the full information model. As has also been pointed out by Mueller (2015), this is precisely because the improvement in the average quality of job-seekers counteracts the effect of lower aggregate productivity on job-creation in the full information model.

This rise in the unemployment rate is primarily driven by the higher cost of information firms face in ascertaining the type of workers they meet. Despite the increased likelihood of meeting a higher-ability applicant, firms find it harder to distinguish between low and high ability types

during the downturn. Given that the firm's loss from accepting a low-ability worker by mistake is greater in a recession than in a boom, the firm would prefer to screen workers more intensely if there were no change in screening costs. However, the increased uncertainty over the effective productivity of applicants causes the cost of information to rise by about 7% on impact. This increased cost of information deters firms from getting informative signals about an applicant and leads them to instead reject workers more often to avoid hiring mistakes. As such, the average acceptance rate, and correspondingly match efficiency, falls by close to 4.5% on impact.

Importantly, the relationship between the changing composition of unemployed job-seekers and screening costs is crucial in understanding the lack of hiring by firms and the consequent decline in match efficiency during a downturn. Noticeably, when information is free, match efficiency barely moves on impact as the decline in aggregate productivity is partially counteracted by a simultaneous increase in the average quality of unemployed job-seekers as shown in Figure 7. Since firms can observe the applicant's ability perfectly in the full information case, the firm does not have to worry about hiring an unsuitable worker. As such, firms do not need to reject more often.

Furthermore, the decline in match efficiency in the model with rationally inattentive firms, causes the inflow into unemployment to spike as well. The higher inflow rate is driven by two forces. Firstly, firms are more selective in terms of who they want to retain on their payroll when aggregate productivity declines. This implies more workers are fired by their employers. Secondly, lower vacancy creation along with lower acceptance rates (conditional on meeting an applicant), imply that newly separated workers have a lower probability of being re-hired within the same period. As such, the EU rate rises by 5.5% on impact. Noticeably, the lower acceptance rates we describe are absent in the full information model as can be seen by the lack of change in match efficiency. As such, the EU rate rises by only about 0.2% in the full information model on impact.

It is interesting to note that our model observes the same asymmetric features as in the data, namely recessions are periods where inflows into unemployment observe sudden spikes but outflows from unemployment are sluggish. The EU rate falls rapidly after its initial spike while the unemployment rate remains elevated above its steady state level for several periods. The slow decline of the unemployment rate reflects the slow recovery of match efficiency to its steady state level.

Crucially, there is a feedback mechanism between firms' hiring behavior today and the composition of unemployed job-seekers. Coarser signals about applicants due to higher screening costs result in the inadvertent rejection of some high ability workers, causing the amount of uncertainty in the applicant pool to dissipate slowly. This further reinforces higher screening costs and lower acceptance rates. In contrast, if firms were to screen more intensively, only lower ability workers would get left behind in the job-seeker pool on average. Going forward, this

would have reduced the firm's uncertainty regarding the ability of future applicants.
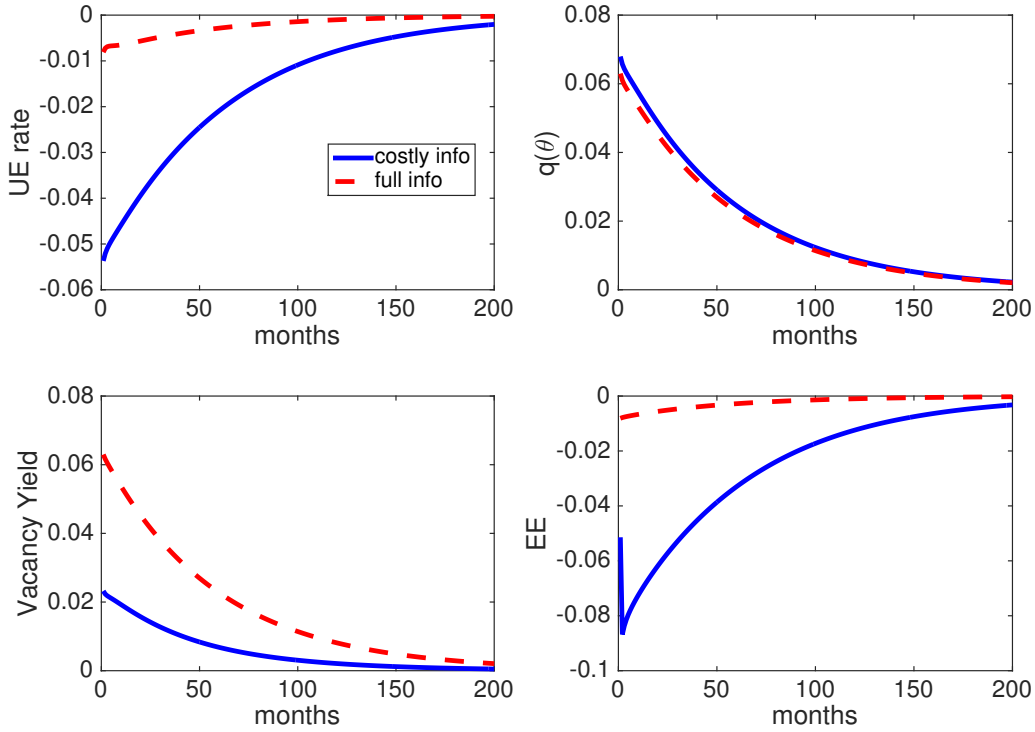

Figure 9: Response in labor flows to a 2 standard deviation drop in productivity

Figure 9 shows how it is the decline in match efficiency and not the differences in the rate at which firms contact workers, $q(\theta)$, that is critical for explaining the muted response in vacancy yield. The top right panel of Figure 9 shows that the contact rate, $q(\theta)$, rises by about 6% in both models. Recall that the hiring rate is made up of two components: the rate at which firms meet workers and the average probability of acceptance (or match efficiency). As we observed in Figure 8, match efficiency barely changed in the full information model, making changes in the job-filling rate entirely dependent on changes in the contact rate. This change in $q(\theta)$ translates one-for-one into a higher response in the vacancy yield as shown in red dashed line in the bottom left panel of Figure 9. This is akin to the rise in the implied job-filling rate with match efficiency held constant as shown in Figure 1.

In contrast, vacancy yield rises by a small amount in the model with rationally inattentive firms. Here the decline in match efficiency counteracts the rise in the meeting rate, causing a muted response in the vacancy yield. This muted response of hiring also shows up as larger declines in the unemployment-to-employment (UE) transition rate. Defining the employment-to-employment (EE) transition rate as the share of employed workers in $t-1$ who were separated at the start of period t but managed to be re-hired at the end of period t, we see that the EE rate also drops to a much larger extent in the costly information model than in the full information

27

model.

4.2 Comparisons with data

4.2.1 Response to a positive TFP shock

Before we compare how our model compares to actual match efficiency in the data, it is important to understand how the two models perform in a boom. We model a boom as a 2 standard deviation increase in aggregate productivity. Unlike a recession, match efficiency in the full information model responds strongly to a positive shock to the economy on impact. This is because the average quality of the pool of job-seekers does not change on impact when a positive shock hits the economy in the full information model. Prior to the boom, firms only hired workers who were suitable for production. With an increase in aggregate productivity, these same workers are effectively more productive and hence, are retained by firms. As such, there is no change in the firing rate when the economy experiences an increase in aggregate productivity.[13] This is shown by the red dashed line in the upper right panel of Figure 10. Since the firing rate does not change on impact, there is no worsening in the pool of unemployed job-seekers to counteract the increase in aggregate productivity. As such, match efficiency in the full information model responds strongly to the change in aggregate productivity.

In contrast, the model with rationally inattentive firms *does* observe a change in the composition of the unemployment pool on impact in both booms and recessions. This is because without perfect information, firms can make mistakes and hire workers unsuitable for production. Some workers that were mistakenly hired in the prior period are now retained due to higher aggregate productivity. As such, the firing rate in the model with rationally inattentive firms initially dips on impact. More relaxed standards for retaining workers implies only workers with very low ze are released back into the unemployment pool on impact. This causes the average quality of the unemployment pool to decline and dampens the impact of a positive aggregate shock on match efficiency. This worsening of the unemployment pool during booms is just the flip side of the findings that the pool of unemployed shifts towards higher ability workers during a recession (See for example Mueller (2015)). The differences in the cost of screening across the two models generate differential initial firing rates which in turn lead to significant differences in the way match efficiency responds in the two models as depicted in the bottom panel in Figure 10.

Figure 10 shows that both models show subsequent spikes in the firing rate as the positive shock to the economy dies off. While workers with a low ze can be hired when aggregate productivity is high, such matches cannot be sustained when the positive shock to the economy dissipates. The spikes in Figure 10 represent the layoffs stemming from the firm's tightening

---

[13]It is important to note that the firing rate is not equivalent to the employment-to-unemployment (EU) transition rate as the firing rate captures the share of employed workers who are laid off while the EU rate captures the share of employed workers who are laid and who could not find a job within the same period.
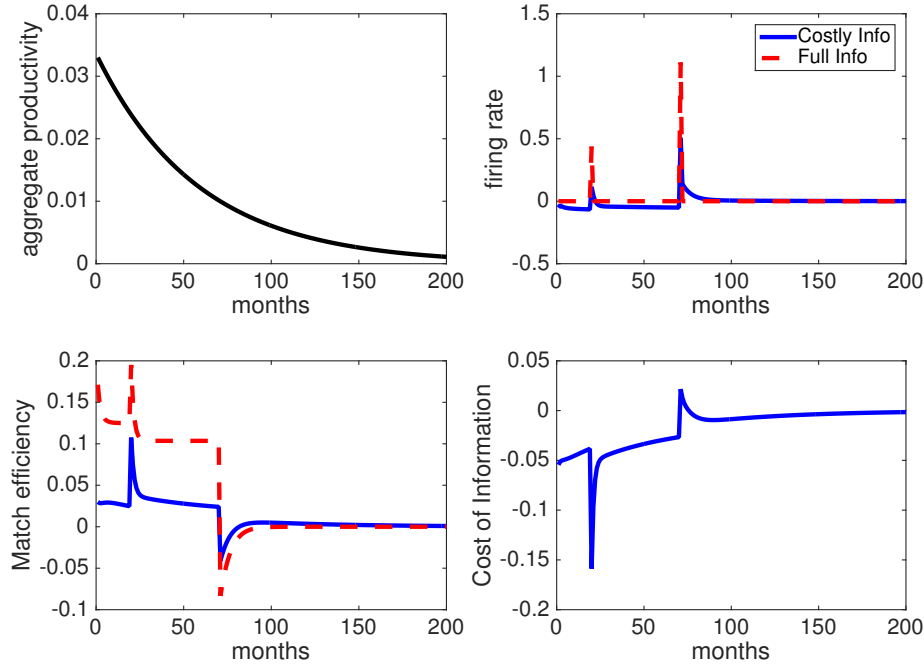
Figure 10: Firing rates from 1 time positive shock

standards for retaining workers as aggregate productivity declines. In our simulations, the second round of firing at $t = 20$ months causes the share of the second-most productive worker to increase in the pool of unemployed (See Figure 11).

The rising average quality of job-seekers and above average aggregate productivity causes firms to choose signals so as to accept workers more often on average.[14] This results in a dip in the cost of information and a spike in match efficiency at $t = 20$. As the aggregate productivity approaches its steady state, firms release workers from matches which now bring negative surplus. This corresponds to the spike in firing at $t = 70$ where aggregate productivity is less than 1 percent above its mean. However, the rising variation in the pool of unemployed job-seekers coupled with the lower aggregate productivity at this stage overwhelms the benefits of an increased average quality of job-seekers. As such, screening costs rise by 2% and match efficiency falls by close to 5%.

4.2.2 Model vs. Data Match Efficiency

Having described how the model responds in both booms and recessions, we now assess how well match efficiency in our model with rationally inattentive firms compares to actual match

---

[14] The workers with second-highest z who were released into unemployment were those who had drawn low match-specific productivity e in their previous jobs. If these workers match with a new firm, they can re-draw a better e from the unconditional distribution, making them attractive for firms to hire.
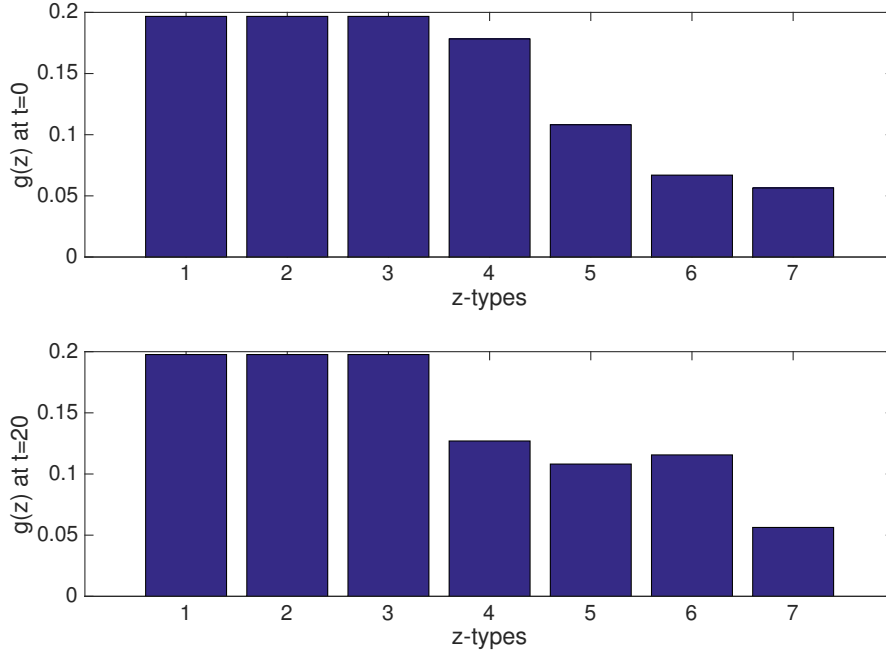
Figure 11: Probability mass of unemployed job-seekers

efficiency in the data. To back out match efficiency in the data, we assume the same matching function as in our model and measure matches as:

$$m = \frac{uv}{(u^\iota + v^\iota)^{1/\iota}}$$

We use data on total non-farm hires from JOLTS as our measure of matches, m, and data on the total non-farm job postings and total unemployed for our measures of v and u respectively. We then run a non-linear least squares regression of the above equation and back out match efficiency as the residual of that regression. We then HP-filter the monthly match efficiency data with smoothing parameter of 14400 to back out its cyclical component.

To assess how close our model generated match efficiency matches the data, we use the TFP series from Fernald (2015) and HP filter the data with smoothing parameter of 1600 to get out its cyclical component. Since the TFP data is quarterly, we apply a cubic spline on the HP filtered data to derive a monthly series. We then feed this filtered monthly TFP series into our model and assess our model predictions. Figure 12 displays the filtered TFP series in the data against the filtered match efficiency in the data. One can observe that TFP has recovered by mid 2009 while match efficiency remained far below its average during that same period. This is consistent with the notion that the labor market recovery lags the recovery in GDP.

We then turn to assess how match efficiency from both the costly information and full information models performs relative to the data. To compare it to the HP-filtered match efficiency
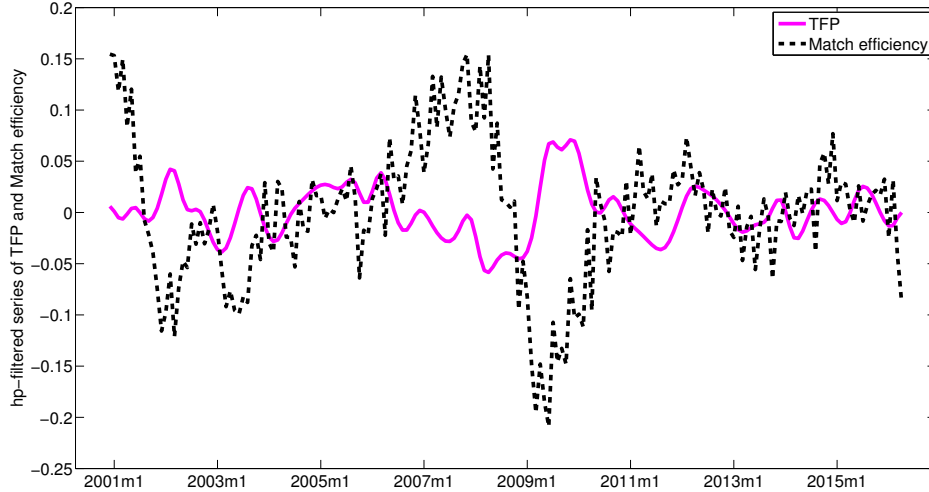
30

Figure 12: TFP vs. Match Efficiency in the Data

data, we present match efficiency generated from the models as the log deviations from its steady state value. For the full information model, we calibrate it such that moments from the full information model matches the targets we use to calibrate our benchmark model.[15] We do this so as to give the full information model the best chance of matching the match efficiency data.[16]
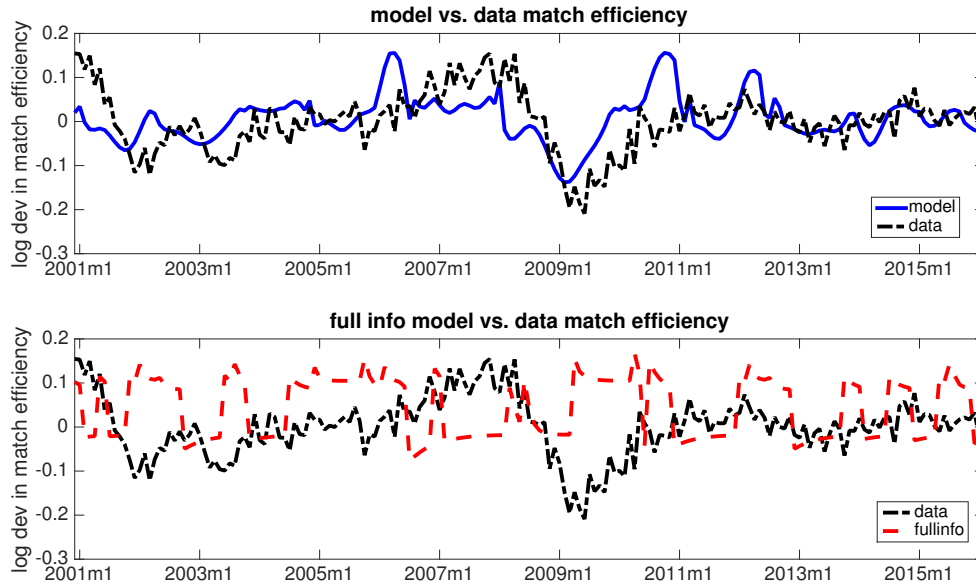

Figure 13: Model vs. Data Match Efficiency

---

[15]For the calibrated parameters used in the full information model, please see the appendix C.

[16]We refer the reader to the appendix Cfor the same graphs but where we keep parameters constant across both costly and full information models but set $\lambda = 0$ for the full information model.

Figure (13) shows the model-generated match efficiency against its data counterpart. The top panel of Figure (13) shows the match efficiency as generated from our costly information model against the data while the bottom panel displays the match efficiency generated from the full information model against the data. The full information model struggles to replicate the declines in match efficiency during a recession and far exceeds actual match efficiency whenever aggregate productivity rises above its mean. This is not surprising since the average quality of the unemployed job-seekers improves on impact during a downturn counteracting the decline in aggregate productivity, while the average of quality of the unemployed does not change on impact when aggregate productivity rises.[17] Consequently, the full information model generates small declines in match efficiency during recessions and large increases during booms. In addition, because TFP rebounds in the data before measured match efficiency recovers, this causes match efficiency to rise sharply whenever TFP recovers and diverge significantly from the behavior of actual match efficiency in the data.

In contrast, the model with rationally inattentive firms is able to generate match efficiency that is similar to its empirical counterpart. This is due to the countercyclical hiring costs that arise endogenously from the changing distribution of unemployed job-seekers over the business cycle. Lower aggregate productivity coupled with higher screening costs cause firms to reject applicants more often, leading to sharp declines in match efficiency. Because screening is not perfect and firms make mistakes in hiring, firms retain workers in booms that they otherwise would not have kept. This again causes the composition of the unemployment pool to worsen in boom times as only the worst workers get released into the unemployment pool and keeps match productivity from responding too strongly to a positive aggregate TFP shock. Overall, the correlation between the match efficiency generated from the model with costly information and the match efficiency measured from the data is about 0.436. In contrast, the correlation between match efficiency generated from the full information model and its data counterpart is -0.178.

4.3   Duration of unemployment as a signal of quality

In reality, firms elicit some information from observable worker-characteristics to defray the costs of a more rigorous interview. One such characteristic which has recently garnered a lot of attention is *unemployment duration*. We use the model to ask (i) whether a longer duration of unemployment for a worker signals low ability and (ii) whether the duration of unemployment is a less informative signal about the applicant's type in a recession?

To see why the duration of unemployment can provide the employer additional information about a worker, note that the failure of an applicant to find employment can be either because she

---

[17]See the response in Figures 8 and 10.

didn't meet a firm or conditional on meeting a firm, she did not clear the interview. Furthermore, the inability to clear an interview could be because of various reasons: (i) a worker could fail the interview if she was actually a low ability, (ii) or that she was high-ability but drew a low match-specific shock, or (iii) that the firm mistakenly rejects a worker since it may be acquiring very imprecise signals about the worker when information is costly. Whenever the cost of information is low and effect (iii) is mild, then the longer the worker has been unemployed, the more likely it is that she is a low-productivity type.

The left panel of Figure 14 highlights this feature by showing that conditional on duration of unemployment τ, the entropy of the distribution of job-seekers who have been unemployed by τ consecutive months decreases with the duration of unemployment. To appreciate this, recall that in equilibrium, higher-ability workers transition to employment from the pool of job-seekers faster then less able workers. Consequently, the firm is more certain that workers who have remained unemployed for longer are on average of low ability. From Figure 6, workers who are *long term* unemployed have on average about a 25% lower probability of transitioning back to unemployment than a worker who has been unemployed for 1 month.
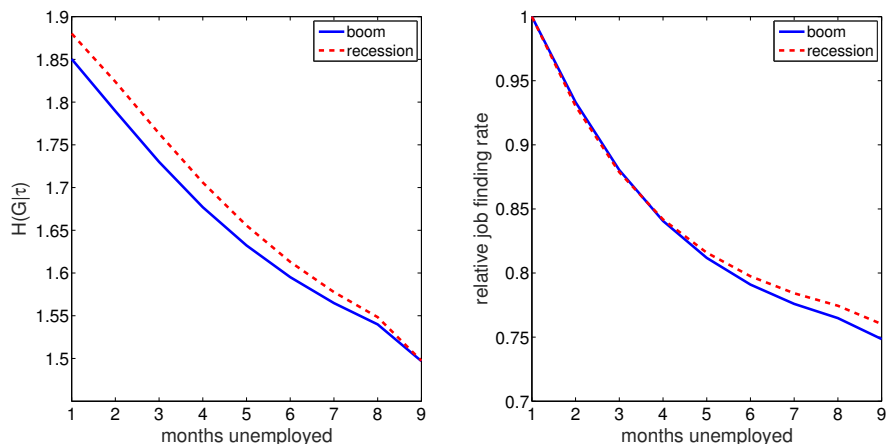

Figure 14: Entropy of the distribution of job-seekers by duration of unemployment τ.

The left panel of Figure 14 also highlights that the uncertainty about the worker's type is higher in a recession (dashed curve) than in a boom (solid curve). Firms must compensate for a fall in aggregate productivity by only retaining relatively high quality workers in order to stay profitable. Thus, the recession results in a flow of both medium-ability workers into unemployment (who were employable in normal times) as well as high-ability workers who were previously attached to low match quality jobs. This increased separation rate raises the uncertainty firms face regarding the pool of unemployed job-seekers. Overtime, given the optimal information strategy of firms, the higher-ability workers exit the pool of unemployed job-seekers at a faster rate on average than low ability types. Consequently, the pool of job-seekers at higher durations of unemployment are still dominated by lower-quality workers.

The firm's uncertainty about an applicant's ability is higher across all levels of unemployment duration in a recession relative to a boom.[18] Lower job-finding rates in recessions lengthen the unemployment duration of all job-seekers and weaken the informational content of unemployment duration as an indicator of underlying worker productivity. From equation (14), when $p(\theta)$ falls, a worker is less likely to meet a firm and hence is more likely to remain unemployed. Compounding this, when the rejection rate given by $1 - \gamma(z, e \mid \sigma, \tau)$ is high, unemployed job-seekers are also less likely to be hired and both these forces contribute to a lengthening of the unemployment duration. As can be seen in Figure 9, the aggregate job-finding rate falls by more than 5 percent in response to the fall in productivity.

The flip-side of this idea above can be seen in the right panel of the same figure. The solid curve depicts the relative job-finding rate (by duration of unemployment) in a boom and the dashed curve is the same object in a recession.[19] Both curves slope downwards, reflecting the fact that firms believe that the pool of long term unemployed is likely to be composed of low ability types than lower durations. However, the relative job finding rates for higher unemployment duration levels fall by less during recessions and implies that unemployment duration is less informative in recessions than in booms. This feature of the model is consistent with the findings of Kroft et al. (2013) who find that there is less stigma attached to longer durations of unemployment spells in areas with depressed economic activities.

5 Conclusion

We present a novel channel through which firms' hiring standards affect fluctuations in match efficiency. The key insight is the presence of a tight link between match efficiency, firms' hiring strategies and the composition of unemployed job-seekers. In particular, we show that selective hiring and retention standards during a downturn cause not only the average quality of the pool of unemployed job-seekers to increase but also raises the uncertainty firms have regarding applicants. This rise in uncertainty increases the cost of processing information, or in other words the cost of screening an applicant. These endogenously arising countercyclical costs are crucial to the model's ability to replicate labor market outcomes. Increased hiring selectivity implies that firms would gather more information on workers absent any increase in screening costs. However, the cyclicality in information costs works against the firm's incentive to acquire more information and leads her to reject applicants more often in order to avoid hiring an unsuitable worker. These lower acceptance rates correspond to declines in match efficiency. Given the marked increase in separation rates observed at the height of the Great Recession, our proposed mechanism offers

---

[18]This is true except for the last bin where we have clustered workers who have 9 or more months of unemployment duration.

[19]Recall that we defined the relative job-finding rate as the UE transition rate of those with duration τ relative to the UE transition rate of those who are 1 month unemployed.

insight as to how the changes in the composition of the unemployment pool caused hiring rates to stall and vacancy yields to falter despite the large number of job-seekers available for each vacancy.

One important aspect we abstracted from in this paper was wage-setting. Rather than explicitly acquiring information about applicants, firms could potentially use contracts to reduce the information costs. However, it is far from clear in such a setting whether the firm would prefer not to expend resources directly on acquiring information. While the use of contracts to separate different ability workers defrays the cost of information acquisition, it requires the firm to give up informational rents. Furthermore, in a setting with multiple worker types, firms may not be able to design contracts to perfectly separate types. In such settings, firms may still choose to explicitly acquire information. The choice of when to issue separating contracts or pooling contracts and screen workers thereafter likely depends on the firm's prior uncertainty over the pool of workers and therefore the cost of information, both of which are changing over the business cycle. We leave this for future research.

References

Barnichon, Regis and Andrew Figura, "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics*, October 2015, *7* (4), 222–49.

Briggs, Joseph, Andrew Caplin, Daniel Martin, and Chris Tonetti, "Due Diligence: Job Search with Rationally Inattentive Workers," Working Paper 2015.

Cappelli, Peter, "Why Companies Aren't Getting the Employees They Need," http://www.wsj.com/articles/SB10001424052970204422404576596630897409182 2011.

Cheremukhin, Anton A., Antonella Tutino, and Paulina Restrepo-Echavarria, "A theory of targeted search," Working Papers 1402, Federal Reserve Bank of Dallas February 2014.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger, "Recruiting Intensity during and after the Great Recession: National and Industry Evidence," *American Economic Review*, 2012, *102* (3), 584–88.

_ , _ , and _ , "The Establishment-Level Behavior of Vacancies and Hiring," *The Quarterly Journal of Economics*, 2013, *128* (2), 581–622.

Eriksson, Stefan and Dan-Olof Rooth, "Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment," *The American Economic Review*, 2014, *104* (3), 1014–1039.

Fernald, John G., "Productivity and Potential Output before, during, and after the Great Recession," *NBER Macroeconomics Annual*, 2015, *29* (1), 1 – 51.

Fujita, Shigeru and Giuseppe Moscarini, "Recall and Unemployment," NBER Working Papers 19640, National Bureau of Economic Research, Inc November 2013.

Gavazza, Alessandro, Simon Mongey, and Gianluca Violante, "What Shifts the Beveridge Curve? Recruitment Effort and Financial Shocks," 2014 Meeting Papers, Society for Economic Dynamics 2014.

Hall, Robert E., "Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor," *Journal of Political Economy*, 04 2009, *117* (2), 281–323.

Jarosch, Gregor and Laura Pilossoph, "Hiring Practices, Duration Dependence, and Long-Term Unemployment," Technical Report 2016.

Kosovich, Stephen M, "The Value of Layoffs and Labor Market Conditions as Signals of Worker Quality," *The B.E. Journal of Economic Analysis & Policy*, March 2010, *10* (1), 1–22.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo, "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*," *The Quarterly Journal of Economics*, 2013, *128* (3), 1123–1167.

_ , _ , _ , and Lawrence F. Katz, "Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation," *Journal of Labor Economics*, 2016, *34* (S1), S7 – S54.

Lester, Benjamin and Ronald Wolthoff, "Interviews and the Assignment of Workers to Firms," Technical Report 2016.

Lockwood, Ben, "Information Externalities in the Labour Market and the Duration of Unemployment," *The Review of Economic Studies*, 1991, *58* (4), 733–753.

Matejka, Filip and Alisdair McKay, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, January 2015, *105* (1), 272–98.

Menzio, Guido and Shouyong Shi, "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, July 2010, *145* (4), 1453–1494.

___ **and** ___ , "Efficient Search on the Job and the Business Cycle," *Journal of Political Economy*, 2011, *119* (3), 468 – 510.

Mirkin, Kenneth, "The Informational Content of Unemployment and Labor Market Dynamics," Working Paper, UCLA 2016.

Mueller, Andreas I., "Separations, Sorting and Cyclical Unemployment," Technical Report 2015.

Nakamura, Emi, "Layoffs and lemons over the business cycle," *Economics Letters*, April 2008, *99* (1), 55–58.

Oberholzer-Gee, Felix, "Nonemployment stigma as rational herding: A field experiment," *Journal of Economic Behavior and Organization*, 2008, *65* (1), 30 – 40.

Pissarides, Christopher A., "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica*, 2009, *77* (5), 1339–1369.

Sedlacek, Petr, "Match efficiency and firms' hiring standards," *Journal of Monetary Economics*, 2014, *62* (C), 123–133.

Shannon, C.E., "A mathematical Theory of Communication," *Bell Technical Journal*, 1948, *3* (27), 379–423.

Shimer, Robert, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, March 2005, *95* (1), 25–49.

Silva, Jose Ignacio and Manuel Toledo, "Labor Turnover Costs And The Cyclical Behavior Of Vacancies And Unemployment," *Macroeconomic Dynamics*, May 2009, *13* (S1), 76–96.

Sims, Christopher A., "Implications of rational inattention," *Journal of Monetary Economics*, April 2003, *50* (3), 665–690.

Tauchen, George, "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 1986, *20* (2), 177–181.

APPENDIX

A Proofs

A.1 Proof of Proposition 3

Without loss of generality, we suppress the dependence of the firm's problem on τ, the duration of unemployment for simplicity. This is the case where firms are unable to use such information to defray the cost of information. The reformulated first-stage problem in Lemma 1 can be expressed as the following Lagrangian:

$$
\begin{aligned}
\mathcal{L} = {} & \sum_z \sum_e \gamma(z, e \mid \sigma) \mathbf{x}(a, z, e) g(z, e \mid \sigma) \\
& -\lambda \Bigg[-\mathcal{P}(\sigma) \log \mathcal{P}(\sigma) - [1 - \mathcal{P}(\sigma)] \log [1 - \mathcal{P}(\sigma)] \\
& + \sum_z \sum_e \Big\{ \gamma(z, e \mid \sigma) \log \gamma(z, e \mid \sigma) + [1 - \gamma(z, e \mid \sigma)] \log [1 - \gamma(z, e \mid \sigma)] \Big\} g(z, e \mid \sigma) \Bigg] \\
& + \sum_z \sum_e \zeta(z, e \mid \sigma) \gamma(z, e \mid \sigma) g(z, e \mid \sigma) - \sum_z \sum_e \mu(z, e \mid \sigma) \left(\gamma_i(z, e \mid \sigma) - 1 \right) g(z, e \mid \sigma)
\end{aligned}
$$

where $\zeta(z, e)$ and $\mu(z, e)$ are the multipliers on the non-negativity constraint and the upper bound of 1 respectively. Taking first order conditions with respect to $\gamma(z, e \mid \sigma)$:

$$
\mathbf{x}(a, z, e) - \lambda \left[-\ln \frac{\mathcal{P}(\sigma)}{1 - \mathcal{P}(\sigma)} + \ln \frac{\gamma(z, e \mid \sigma)}{1 - \gamma(z, e \mid \sigma)} \right] + \zeta(z, e \mid \sigma) - \mu(z, e \mid \sigma) = 0 \tag{18}
$$

with complementary slackness conditions

$$
\mu(z, e \mid \sigma) \left[1 - \gamma(z, e \mid \sigma) \right] = 0 \tag{19}
$$
$$
\zeta(z, e \mid \sigma) \gamma(z, e \mid \sigma) = 0 \tag{20}
$$

Thus, as long as $0 < \gamma(z, e \mid \sigma) < 1$, it must be the case that $\zeta(z, e \mid \sigma) = \mu(z, e \mid \sigma) = 0$ and $\gamma(z, e \mid \sigma)$ can be written as:

$$
\gamma(z, e \mid \sigma) = \frac{\mathcal{P}(\sigma) e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 - \mathcal{P}(\sigma) \left[1 - e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} \right]} \tag{21}
$$

Summing across (z, e) and dividing both sides by $\mathcal{P}(\sigma)$, one can show that:

$$1 = \sum_z \sum_e \frac{e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 - \mathcal{P}(\sigma)\left[1 - e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}\right]} g(z, e \mid \sigma) \tag{22}$$

A.2 Proof of Lemma 2

Recall that under the optimal information strategy, the induced probability of accepting an applicant of type (z, e) is given by:

$$\gamma(z, e \mid \sigma) = \frac{\mathcal{P}(\sigma) e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 + \mathcal{P}(\sigma)\left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda}} - 1\right]}$$

Now consider the costless information case which corresponds to the limit in which $\lambda \to 0$. First consider an applicant (z, e) such that $\mathbf{x}(a, z, e) < 0$. If the firm hired this worker, the firm would surely make losses. Under the optimal information strategy, the firm rejects this worker with probability 1.

$$\lim_{\lambda \to 0} \gamma(z, e \mid \sigma) = \lim_{\lambda \to 0} \frac{\mathcal{P}(\sigma) e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 + \mathcal{P}(\sigma)\left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda}} - 1\right]} = 0 \tag{23}$$

Next, consider an applicant (a, z) such that $\mathbf{x}(a, z, e) \geq 0$. If the firm hired this worker, the firm would surely have positive per-period profits. Then under the optimal information strategy, this applicant is hired with probability 1.

$$\lim_{\lambda \to 0} \gamma(z, e \mid \sigma) = \lim_{\lambda \to 0} \frac{\mathcal{P}(\sigma) e^{\frac{\mathbf{x}(a,z,e)}{\lambda}}}{1 + \mathcal{P}(\sigma)\left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda}} - 1\right]} = \lim_{\lambda \to 0} \frac{\mathcal{P}(\sigma) \mathbf{x}(a, z, e) e^{\frac{-\mathbf{x}(a,z,e)}{\lambda^2}}}{\mathcal{P}(\sigma) \mathbf{x}(a, z, e) e^{-\frac{\mathbf{x}(a,z,e)}{\lambda^2}}} = 1 \tag{24}$$

where the second equality follows from L'Hospital's Rule.

B Numerical Implementation

We assume that firms observe a top-coded distribution of unemployment durations. Firms can observe the exact duration of unemployment τ as long as $0 \leq \tau < \bar{\tau}$. For all worker unemployed for a duration of at least $\bar{\tau}$, the firm cannot see the exact duration of unemployment but knows that the duration is at least $\bar{\tau}$. Then the transition equations for this top-coded model can be

written as:

$$
l_t(z, \tau) = \begin{cases} \int_e d(a_t, z, e) n_{t-1}(z, e) & \text{if } \tau = 0 \\[2ex] u_{t-1}(z, \tau) & \text{if } 1 \leq \tau < \bar{\tau} \\[2ex] u_{t-1}(z, \bar{\tau}) & \text{if } \tau \geq \bar{\tau} \end{cases} \tag{25}
$$

$$
u_t(z, \tau) = \begin{cases} l_t(z, \tau - 1)\Big\{1 - p\left(\theta[\sigma_t]\right) + p\left(\theta[\sigma_t]\right) \sum_e \pi_e(e)(1 - \gamma\left[z, e \mid \sigma_t, \tau - 1\right])\Big\} & \text{if } 1 \leq \tau < \bar{\tau} \\[2ex] l_t(z, \bar{\tau} - 1)\Big\{1 - p\left(\theta[\sigma_t]\right) + p\left(\theta[\sigma_t]\right) \sum_e \pi_e(e)(1 - \gamma\left[z, e \mid \sigma_t, \bar{\tau} - 1\right])\Big\} + \\ l_t(z, \bar{\tau})\Big\{1 - p\left(\theta[\sigma_t]\right) + p\left(\theta[\sigma_t]\right) \sum_e \pi_e(e)(1 - \gamma\left[z, e \mid \sigma_t, \bar{\tau}\right])\Big\} & \text{if } \tau \geq \bar{\tau} \end{cases} \tag{26}
$$

We use this top-coded model in our numerical exercises. For the purpose of our numerical exercises we set $\bar{\tau} = 9$ months. Thus, we label all individuals who have been unemployed for more than 9 months into one group.

C   Parameterization of Full Information model

We re-calibrate the full information model such that the simulated moments from the full information model match our target moments. These are the parameters used to generate the bottom panel of Figure 13. The following parameter values used are as listed below:

Inferred Parameters

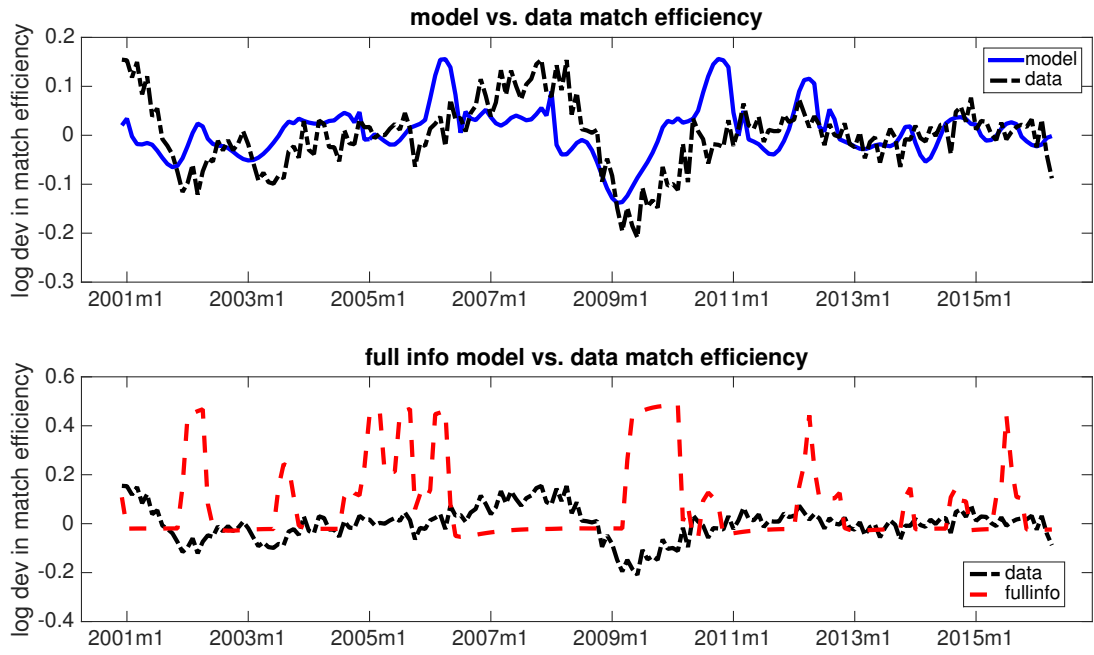| Parameter	Description	Value
b	home production	0.5213
δ	exog. separation rate	0.0270
κ	vacancy posting cost	0.0004
λ	marginal cost of information	0
A_z	shape parameter - worker ability	0.5216
B_z	shape parameter - worker ability	1.2893
A_e	shape parameter - match productivity	4.4241
B_e	shape parameter - match productivity	8.4046

Figure 15: Match efficiency: model vs. data

Response of full information model

We also compare the full information keeping the same parameters as in 1 but setting $\lambda = 0$. This gives us qualitatively the same outcomes: match efficiency in the full information model struggles to replicate the declines in match efficiency during recessions and over-predicts the rise in match efficiency during booms.