

# Rational Inattention in Hiring Decisions\*

Sushant Acharya  
FRB New York

Shu Lin Wee  
Carnegie Mellon University

February 16, 2016

[ Latest version [here](#) ]

## Abstract

The Great Recession observed a severe decline in labor market matching efficiency. Using the standard random search framework, we build a model of rational inattention in firm's hiring decisions. Firms ex-ante do not know the productivity of the worker they encounter and must expend resources to reduce their uncertainty over the worker's type. The cost of getting information, however, is rising in the dispersion of the pool of unemployed job-seekers. Selective hiring standards during a downturn further exacerbate this cost of acquiring information as the higher standards for retaining existing workers raises the disparity in the pool of job-seekers. These higher costs translate into lower acceptance probabilities and hiring rates by firms as firms now face a higher likelihood of making a mistake in terms of hiring low quality workers during the recession. Relative to the full information case, the model where firms are rationally inattentive can lead to amplifications in the unemployment rate and muted movements in the vacancy yield. These model-generated muted movements in vacancy yields and large declines in matching efficiency during a downturn are reminiscent of the behavior in hiring rates since the Great Recession.

## 1 Introduction

Accompanying the severe spike in unemployment rates during the Great recession was a tripling in the ratio of unemployed job-seekers for each job opening. Despite this sharp increase in the number of available job-seekers per vacancy, the actual number of observed hires was below

---

\*The views expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the Federal Reserve Bank of New York or the Federal Reserve System.

**Email:** Acharya: [sushant.acharya@ny.frb.org](mailto:sushant.acharya@ny.frb.org) , Wee: [shuwee@andrew.cmu.edu](mailto:shuwee@andrew.cmu.edu)

what would have been predicted by a standard matching function. This observation has led many commentators to argue that matching efficiency declined during the Great Recession due to employers becoming more selective in their recruitment process. In this paper, we show that the cost of acquiring information about workers varies over the business cycle. This in turn affects the type of recruitment strategies firms undertake over the business cycle and consequently affects the acceptance probability and hiring rates of firms.

In this paper, we argue there is a tight link between the dispersion in the pool of unemployed job-seekers and the cost of recruiting. Consider an environment where workers differ in their productivity and firms ex-ante do not know the worker's type. Firms can pay to get signals about the worker's type. Upon hiring and production, the worker's productivity is revealed and the firm must make a decision of whether to continue the employment relationship the next period. Since hiring an unproductive worker is costly as the firm can incur negative profits from hiring a low quality worker, firms would like to be selective in their hiring methods.

Selectivity in hiring however, comes at a cost. While firms do not observe workers' productivity when they encounter a worker, they can expend resources to acquire information and reduce the uncertainty about the worker's type. The firm's level of uncertainty about the worker's type however is increasing in the disparity of the pool of job-seekers. We show that selective hiring standards during a recessionary period exacerbate this cost of reducing the uncertainty over the worker's type by increasing the dispersion in the pool of unemployed job seekers.

Previous studies such as [Kosovich \(2010\)](#), [Lockwood \(1991\)](#) and [Nakamura \(2008\)](#) have argued that the quality of the unemployment pool improves during a recession as firms become more selective over the workers they hire and retain. This causes better quality workers to enter or remain in unemployment during downturns. This rise in the average quality of the unemployment pool caused by more selective hiring standards in a recession also induces a corresponding rise in the dispersion of productivity types. In a boom, only individuals deemed to have very poor productivity are left in the unemployment pool. In a recession, however, both high and low quality workers can be found within the unemployment pool as the number of workers laid off due to bad luck rises in a recession. It is this increase in the dispersion of unemployed types that raises the firm's uncertainty as to the type of worker he would encounter should he post a job. Rising uncertainty about the worker's type implies that firms in turn face higher recruiting costs if they wish to get more precise information about the worker's type. These higher recruitment costs which stem from the increased dispersion of unemployed types in the downturn translate into lower acceptance probabilities and hiring rates by firms as firms now face a higher likelihood of making a mistake in terms of hiring low quality workers rises during the recession. Unless

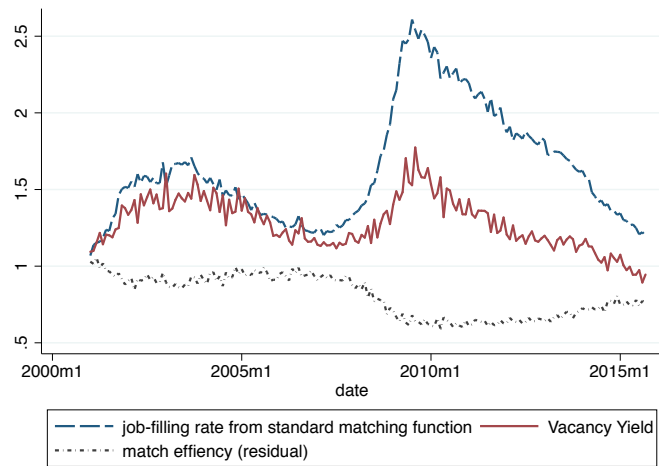


Figure 1: Vacancy Yield and Match Efficiency as computed from JOLTS and CPS

firms are willing to pay very high costs to precisely learn the worker’s type, the probability of acceptance on average declines further relative to a case where information is costless.

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.<sup>1</sup> 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 consider a particular interpretation of recruiting intensity and view it as the cost or amount expended to reduce uncertainty about the worker’s type. To this end, we consider

<sup>1</sup>The vacancy yield is defined as the ratio of hires to vacancies.

a model of rational inattention where firms optimally choose how many signals to receive and their level of precision in deciding whether to hire a worker in different states of nature. The choice of signals by the firm has implications for the firm’s acceptance probability of a worker, the latter which we view as a proxy for the match efficiency in the economy. The varying cost of information stemming from differences in the dispersion in the unemployment pool over the business cycle allows us to replicate the divergence between the rate at which firms contact workers and the rate at which workers are accepted or hired.

Several papers have tried to examine and decompose the forces driving the decline in match efficiency. [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. Closely related to our paper, [Sedlacek \(2014\)](#) considers a model with hiring standards and examines how the time-varying selectivity of the firm can explain the variation in matching efficiency. While [Sedlacek \(2014\)](#) considers a model of full information, we show how the selectivity of the firm’s hiring process adds to their recruitment cost in downturns and further exacerbates the decline in match efficiency during a recession. Given the severity of the Great Recession and the sharp increase in separation rates observed at the height of the recession, 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.

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. Next, 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  $\beta$ . 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.

**Firms** We define jobs as a single firm-worker pair. The per-period output of a job is given by the production function  $F(a, z, e)$  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  $a$  evolves according to a finite  $n_a$ -state Markov process  $\langle \mathcal{A}, \Pi_a(a' | a) \rangle$ . The match-specific shock  $e \in \mathcal{E}$  is independent of  $a, z$  and is drawn from a time-invariant distribution  $\Pi_e(e)$ . The match specific shock is drawn at the time a firm and worker meet and persists till the end of the match.

A firm that decides to enter the market must post a vacancy at a cost  $\kappa > 0$ . We assume that the measure of firms in operation at any date  $t$  is determined by free-entry. Following this, a firm that posts a vacancy comes in contact with a worker at a rate  $q_t$ . This contact rate depends on the total number of vacancies and job-seekers according to an exogenously specified 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. We assume that  $m(\cdot, \cdot)$  is homogeneous of degree 1. In addition, our assumption of random search ensures that the probability that a firm bumps into a particular type of worker is purely random. 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. We deviate from the standard model by assuming that a firm cannot observe the type of the applicant at the time of meeting. In addition to the worker's type  $z$  being unobservable, the firm also does not directly observe the realization of the match-specific shock  $e$ . 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 the match has been formed and undertaken production once. 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 his uncertainty about the worker's effective productivity which is affected by both  $z$  and  $e$ . A more rigorous interview reveals more information about the applicant but is more costly for the firm. Depending on the information revealed in the interview, the firm decided whether to hire an applicant or not. Moreover, we assume that the state of aggregate productivity  $a$  is common knowledge. The timing of the model is summarized in Figure 2.

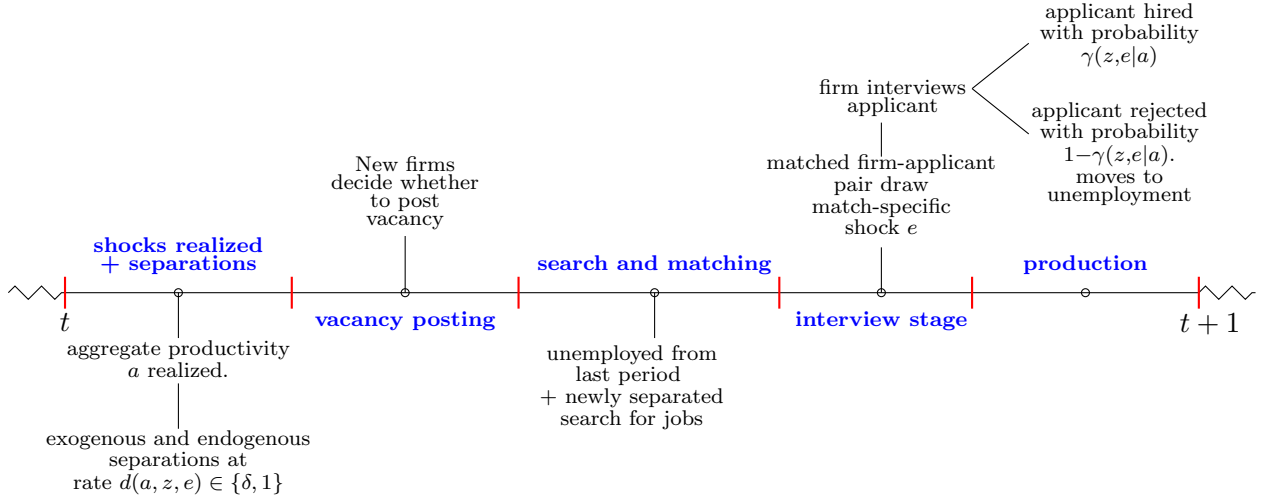


Figure 2: Timeline

## 2.1 Hiring Strategy of the Firm

Consider a firm that has posted a vacancy when the state of aggregate productivity is  $a$ . Conditional on getting matched with an applicant, the hiring strategy of a firm can be described as a two-stage process. (i) In the *first-stage*, 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  $e$ . We interpret this as how rigorous the interview conducted by the firm is. (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 choices of the firm by solving the problem backwards.

### 2.1.1 Second-stage Problem

In the second-stage, the firm has already chosen an information strategy and received signals  $\mathbf{s}$  about the applicant's type. Denote the joint-posterior belief of the firm about the applicants productivity-type  $z$  and match-specific shock  $e$  by  $\Gamma(z, e | \mathbf{s})$ . Given this belief the firm must choose whether to *hire* or *reject* the applicant. The payoff to the firm if it chooses to reject the applicant is known with certainty to be 0. However, depending on the combination of  $(z, e)$ , the payoff from hiring an applicant can 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 (unless the posterior is degenerate). The proposition below summarizes the second stage decision problem:

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

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

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

$$J(\Gamma(\cdot | \mathbf{s}), a) = \max \left\{ 0, \mathbb{E}_\Gamma[\mathbf{x}(a, z, e) | a] \right\} \quad (1)$$

*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.  $\square$

### 2.1.2 First-stage Problem

The first stage consists of choosing an information strategy for the firm or a set of signals to receive. We model costly information processing as an entropy-based cost function in the tradition of the rational inattention literature such as [Sims \(2003\)](#), [Matejka and McKay \(2015\)](#) and others. In other words, a firm can reduce the uncertainty about the applicant's type and match-specific shock 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.

Under this assumption, a choice of the information strategy can be thought of as the firm asking an applicant a series of questions. 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. [Matejka and McKay \(2015\)](#) argue that 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.<sup>2</sup>

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

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

---

<sup>2</sup>For details, see the coding theorem ([Shannon, 1948](#)).

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  $p(x | \mathbf{s})$ . Then, the mutual information between the prior and the posterior is given by:

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

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

We are now ready to describe the first stage problem.

**Definition 2** (Information Strategy). Denote the joint prior distribution about an applicants worker-productivity  $z$  and match-specific shock  $e$  (given aggregate productivity  $a$ ) by  $G(z, e | a)$ . Through the interview, firms can choose to receive signals  $\mathbf{s}$  in order to update it's belief about worker-productivity  $z$  and match-specific shock  $e$  of the applicant. More informative signals cost more than less informative ones. The information strategy of a firm is given by a joint distribution of signals and types  $\Gamma(z, e, \mathbf{s} | a)$  (given that aggregate productivity is  $y$ ) such that:

$$G(z, e | a) = \int_{\mathbf{s}} d\Gamma(z, e, \mathbf{s} | a) \tag{2}$$

Equation (2) 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  $\Gamma(\mathbf{s} | z, e, a)$ . The Proposition below summarizes the first-stage problem of the firm.

**Proposition 2** (First-Stage Problem of a Firm). The firm chooses an information strategy to maximize ex-ante payoffs from the second-stage less the cost of information:

$$\max_{\Gamma \in \Delta} \int_{z, e, \mathbf{s}} J[\Gamma(\cdot | \mathbf{s}, a) | a] d\Gamma(\mathbf{s} | z, e, a) dG(z, e | a) - c(\Gamma)$$

subject to equation (2) and where the cost of information  $C(\Gamma)$  is given by:

$$c(\Gamma) = \lambda [H(G) - \mathbb{E}_{\mathbf{s}} H(\Gamma(\cdot | \mathbf{s}, a))] \tag{3}$$

where  $\lambda$  is the marginal cost of reducing an additional unit of uncertainty measured in terms of mutual information.

The problem above allows firms to construct signals of any form they wish as long as they satisfy the consistency requirement in equation (3). This is not an easy problem to characterize



a solution to. 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 state-contingent choice probabilities and the associated payoffs. The following Lemma presents the reformulated problem:

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

$$\mathbb{V}(a) = \max_{\{\gamma(z,e|a) \in [0,1]\}} \int_{z,e} \gamma(z, e | a) \mathbf{x}(a, z, e) dG(z, e | a) - c(\mathcal{P}, G) \quad (4)$$

subject to:

$$0 \leq \gamma(z, e | a) \leq 1, \forall z, e$$

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

$$\begin{aligned} c(\mathcal{P}, G) = & \lambda \left[ -\mathcal{P}(a) \log \mathcal{P}(a) - [1 - \mathcal{P}(a)] \log [1 - \mathcal{P}(a)] \right. \\ & \left. + \int_{z,e} \left\{ \gamma(z, e | a) \log \gamma(z, e | a) + [1 - \gamma(z, e | a)] \log [1 - \gamma(z, e | a)] \right\} dG(z, e | a) \right] \end{aligned} \quad (5)$$

where  $\gamma(z, e | a)$  is the induced probability of hiring a worker of type  $(z, e)$  in aggregate state  $a$  and  $\mathcal{P}(a) = \int_{z,e} \gamma(z, e | a) dG(z, e | a)$  is the induced unconditional probability of hiring an applicant in aggregate state  $a$ .

*Proof.* See [Matejka and McKay \(2015\)](#). □

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  $e$  in a state with aggregate productivity  $a$  with probability  $\gamma(z, e | a)$  which can be written as:

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

Consequently, the unconditional probability that a firm hires an applicant after meeting her (when aggregate productivity is  $a$ ) is implicitly defined by:

$$1 = \int_{z, e} \frac{e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 + \mathcal{P}(a) \left[ e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} - 1 \right]} dG(z, e | a) \quad (7)$$

*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 worker-productivity  $z_1$  and  $z_2$  with  $z_1 > z_2$ . Then the optimal information strategy implies the following:

$$\log \frac{\gamma(z_1, e | a)}{1 - \gamma(z_1, e | a)} - \log \frac{\gamma(z_2, e | a)}{1 - \gamma(z_2, e | a)} = \frac{\mathbf{x}(a, z_1, e) - \mathbf{x}(a, z_2, e)}{\lambda} \quad (8)$$

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  $\lambda$ , 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  $\lambda \rightarrow \infty$ , 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 ( $z, e$ ) under the optimal information strategy is given by:*

$$\gamma(z, e | a) = \begin{cases} 1 & \text{if } \mathbf{x}(a, z, e) \geq 0 \\ 0 & \text{else} \end{cases} \quad (9)$$

*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,  $\mathcal{P}(a) < 1$  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

The previous subsections characterized the hiring decisions of a firm conditional on meeting an applicant of type  $(z, e)$  given  $\mathbf{x}(a, z, e)$ . Here we close the model by specifying this payoff. Given our assumption that the workers type is revealed after one period of production,  $\mathbf{x}(a, z, e)$  can be written as:

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

where  $\delta$  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.  $\mathcal{J}(a, z, e)$  is the value of a firm that knows the type of his worker when aggregate productivity is  $a$  and is given by:

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

and  $d^*$  can be written as:

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

This constitutes a full-description of the payoffs of a firm that posts a vacancy and meets a worker.

### 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. Under free entry, each firm posting a vacancy makes zero profits in expectation. This condition allows us to back out market-tightness:

$$\begin{aligned} \kappa &\geq q(\theta) \mathbb{V}(a) \\ [\kappa - q(\theta) \mathbb{V}(a)] \theta &= 0 \end{aligned} \quad (12)$$

where  $\mathbb{V}(a)$  denotes the value of a firm from hiring a worker net of interview costs and is defined in equation (4).  $\theta$  denotes the market-tightness and  $q(\theta) = \frac{m(v,l)}{l}$  is the rate at which a firm meets a job-seeker. We refer to this as the *contact rate*. The average *job-filling rate* in our model is given by  $q(\theta)\mathcal{P}(a)$ . This wedge arises because a firm can choose to reject an applicant after interviewing the applicant.<sup>3</sup>

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

Recall that the information strategies specified above depended on the prior  $G(z, e | a)$  that firms held about an applicant who she meets. However, up till now, we have not specified how a firm forms these priors. We address this issue next. Since aggregate productivity is common knowledge, the aggregate laws of motions of each type of worker are known by all firms. In particular, the evolution of the mass of unemployed workers with worker-productivity can be written as:

$$\begin{aligned} u'(z) &= (1 - p[\theta(a)]) \left\{ u(z) + \int_e d(a, z, e)n(z, e) \right\} \\ &\quad + p[\theta(a)] \left\{ u(z) + \int_e d(a, z, e)n(z, e) \right\} \int_{\hat{e}} \left[ 1 - \gamma(z, \hat{e} | a) \right] \pi_e(\hat{e}) \end{aligned} \quad (13)$$

$$n'(z, e) = [1 - d(a, z, e)]n(z, e) + p[\theta(a)]\pi_e(e)\gamma(z, e | a) \left\{ u(z) + \int_{\hat{e}} d(a, z, e)n(z, \hat{e}) \right\} \quad (14)$$

The first term on the RHS of equation (13) reflects the mass of job-seekers with worker-productivity  $z$  in the current period who were unable to meet a firm. This pool consists of workers who were unemployed in the last period and also those workers who get separated at the beginning of the period. The second term on the RHS of equation (13) reflects the mass of job-seekers with worker productivity  $z$  who did manage to match with a firm but were rejected after the interview stage. Of the job-seekers with worker productivity  $z$ , a fraction  $p(\theta)$  meet a firm. Out of this group which got matched with a firm, a fraction  $\pi_e(\hat{e})$  drew a match-specific shock of  $\hat{e}$  but got rejected after the interview with probability  $1 - \gamma(z, \hat{e} | a)$ .<sup>4</sup>

Similarly, equation (14) describes the evolution of the mass of workers with worker productivity  $z$  who are employed in jobs with match-specific productivity  $e$ . The first term on the RHS reflects the mass of existing jobs of type  $(z, e)$  which do not get destroyed. The last term on

<sup>3</sup>As explained earlier, this wedge also potentially exists in our model with no information costs because of the presence of worker heterogeneity. The cost of information affects the size and cyclical nature of this wedge.

<sup>4</sup> $\pi_e$  refers to the p.m.f of the match specific shock.

the RHS reflects those unemployed workers with productivity  $z$  who match with a firm with probability  $p(\theta)$ , draw match-specific productivity  $e$  and then get hired after the interview with probability  $\gamma(z, e | a)$ .

Following equations (13) and (14), the mass of each unemployed type is given by:

$$g_z(z | a) = \frac{u(z) + \int_e d(a, z, e)n(z, e)}{u + \int_{z,e} d(a, z, e)n(z, e)} \quad (15)$$

$g_z(z | a)$  denotes the p.m.f of the distribution of workers with worker-productivity  $z$  in the pool of job-seekers. 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  $g_z(z | a)\pi_e(e)$ . Denote by  $G(z, e | a)$  the joint-distribution of  $(z, e)$ . Since this is an environment with random search, a firm's prior about any workers type  $(z, e)$  (without any additional information) is simply given by the joint distribution  $G(z, e | a)$ . This concludes the description of the model. In the next section, we proceed to analyze how our model of rational inattention compares against a model where firms are perfectly informed.

### 3 Calibration

We calibrate the parameters of the model using data on the aggregate flows of workers in the US labor market. We then use the calibrated model to observe how unemployment, vacancy yields and transition rates from unemployment back to employment are affected under a one standard deviation drop in aggregate productivity. We compare how our model with costly information compares against a model of full information, i.e. the standard DMP search and matching model. We take a period to be a month in our model. Both workers and firms discount the future with discount factor  $\beta$ . We set  $\beta = 0.9967$  which is consistent with an annualized risk free rate of about 4 percent. We assume that the rate at which a worker meets a firm  $p(\theta)$  takes the form of  $p(\theta) = \theta(1 + \theta^\iota)^{-1/\iota}$  and set  $\iota$  to be 0.5 as standard in the literature.<sup>5</sup>

We assume that the production function takes the following form  $F(a, z, e) = a \times z \times e$ . We assume that  $\log(a)$  follows an AR(1) process with persistence  $\rho_a$ , unconditional mean  $\mu_a$  and standard deviation  $\sigma_a$ . We approximate the stochastic process of  $a$  with a seven-state Markov process using the algorithm specified in [Tauchen \(1986\)](#). We set  $\mu_a = 0$  and the persistence  $\rho_a = 0.983$ . We set the standard deviation  $\sigma_a = 0.042$  which gives us a quarterly standard

---

<sup>5</sup>See for example [Menzio and Shi \(2011\)](#).

deviation in output of about 0.01, similar to the standard deviation of the HP filtered quarterly log GDP series for the period spanning 2000Q4 to 2015Q4. To estimate the distribution of the worker’s permanent productivity component, we use CPS data from the fourth and eighth survey months for prime age white males from age 25-55 and run a fixed effects wage regression on age, age- squared, education, the national unemployment rate and year dummies. We find that the standard deviation in unobserved fixed effects of workers is about 0.49 while the dispersion in the residual to be about 0.18. We treat the fixed effects from our wage regression to be indicative of the unobserved heterogeneity of workers’ productivity. Thus, we assume that  $\log(z)$  is a five point approximation of a normal distributed with mean  $\mu_z = 0$  and  $\sigma_z = 0.49$ . For match quality,  $e$ , we use the Beta distribution with support  $[0, 1]$  and fix the first shape parameter to be equal to 2. We then choose the other shape parameter so as to get a standard deviation of 0.18, which was the dispersion in the residual of our wage regression.

The remaining parameters,  $b, \delta, \kappa$ , and  $\lambda$  are chosen such that statistics from the simulated data match their empirical counterparts. [Hall and Milgrom \(2008\)](#) find that the replacement rate is roughly seventy percent. In our model, firms have all the bargaining power and pay workers their value from home production. As such, we calibrate  $b$  such that it is 70 percent of output. We target a monthly unemployment rate of 6.5%, which is the average unemployment rate in the data over the same coverage period as JOLTS. [Shimer \(2005\)](#) finds a monthly job-finding rate of 0.45. In our model, both the rate at which workers meet firms and the firm’s acceptance probability affects the flow of workers from unemployment to employment. As such, we jointly calibrate  $\lambda$ , the per unit cost of reducing uncertainty, and  $\kappa$ , the vacancy posting cost to meet the target of a 6.5 percent unemployment rate and a job-finding rate of 0.45. Finally, [Shimer \(2005\)](#) reports that the average job lasts 2.5 years. We target a monthly separation rate of 3%. [Tables 1 and 2](#) summarize the parameter values we use.

## 4 Results

We simulate a recession as a one standard deviation drop in productivity. We compare our results against the full information model where firms, upon meeting the worker, discover perfectly the values of worker-productivity  $z$  and match-specific productivity  $e$  and can choose whether to accept or reject the worker. Note that the full information model is similar to the standard DMP search model where the firm has all the bargaining power. Firms reject workers whenever the surplus cannot support the match. [Figure 3](#) plots the response of aggregate variables to a one standard deviation drop in productivity in both the full and costly information model. For

Variable	Value	Description/Source
$\beta$	0.9967	annualized interest rate = 4%
$\mu_a$	0	Mean of $\log(a)$
$\sigma_a$	0.042	Target standard deviation of cyclical component of $\log$ GDP
$\rho_a$	0.983	persistence of $\log(y)$
$\mu_z$	0	Mean of $\log(z)$
$\iota$	0.5	<a href="#">Menzio and Shi (2010)</a>
$\sigma_z$	0.49	Standard deviation of fixed effect term in wage regression
$\Pi_e$	$\sim \text{Beta}(2,3.91)$	Target standard deviation of residual from wage regression

Table 1: Parameter Values

Variable	Value	Target	Data	Model Generated
$b$	0.40	70% of output	0.7	0.689
$\delta$	0.024	Mean tenure = 2.5 years	0.03	0.0308
$\kappa$	0.15	Monthly job finding rate	0.45	0.4577
$\lambda$	0.4	Average unemployment rate	0.065	0.0654

Table 2: Calibrated Parameters

both the costly information and full information models, we compute the steady state values for each of these models and record the log deviation of the aggregate variables from their steady state value.

Noticeably, the model where information is costly provides more amplification in terms of the response of unemployment to an aggregate productivity shock relative to the full information model; the unemployment rate jumps by 40 percent in the costly information model while it rises by about 30 percent in the full information model. This amplification arises as the cost of information increases during a recession and hence further squeezes firms' profits during a downturn. This in turn lowers the firm's incentive to create jobs. However, the critical force driving the amplification of unemployment rates in the costly information model stems from the greater decline in the firm's ex-ante probability of acceptance  $\mathcal{P}(a)$  relative to the full information model. We explain the crucial role of information costs and its impact on  $\mathcal{P}(a)$  below.

The cost of recruiting, i.e. the cost of getting information about workers, rises during the recession as the drop in aggregate productivity causes existing firms to only retain workers who have relatively higher effective productivity  $ze$ . This lower retention rate on impact is reflected by the sudden spike in the flows from employment to unemployment, i.e. the EU transition

rate in Figure 4. The increase in firms' standards both in terms of the workers they are willing to hire and retain causes the dispersion in the pool of job seekers to rise. It is important to note that the pool of job seekers is not necessarily made up of only low productivity ( $z$ ) workers during a recession but also consists of high productivity workers who drew middling to low match quality shocks ( $e$ ) in their previous jobs. As such, the firm still has an opportunity to hire a worker with high permanent productivity  $z$  in the downturn but he faces a more disparate pool of job-seekers. Figure 5 demonstrates how both average quality and dispersion in the pool of job seekers rise with respect to an adverse productivity shock in the costly information model. The increases in average quality and dispersion of the job-seeker pool are consistent with the findings from previous studies such as Lockwood (1991). As the dispersion in the unemployment pool increases in the downturn, the firm during the downturn has to expend more resources to figure out if the worker he contacts will be meet his hiring standards. This increased cost of getting information does not arise in a boom where firms are willing to hire almost anyone.

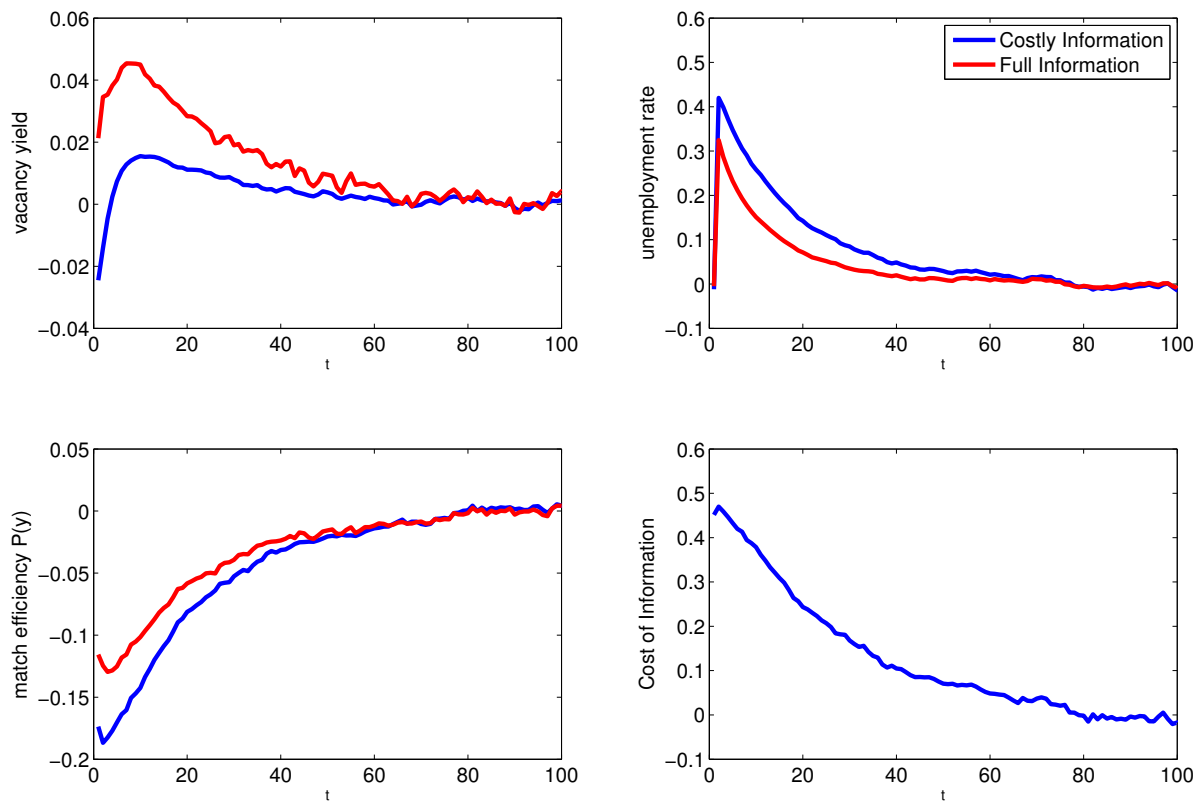


Figure 3: Response in aggregate variables to a 1 standard deviation drop in productivity



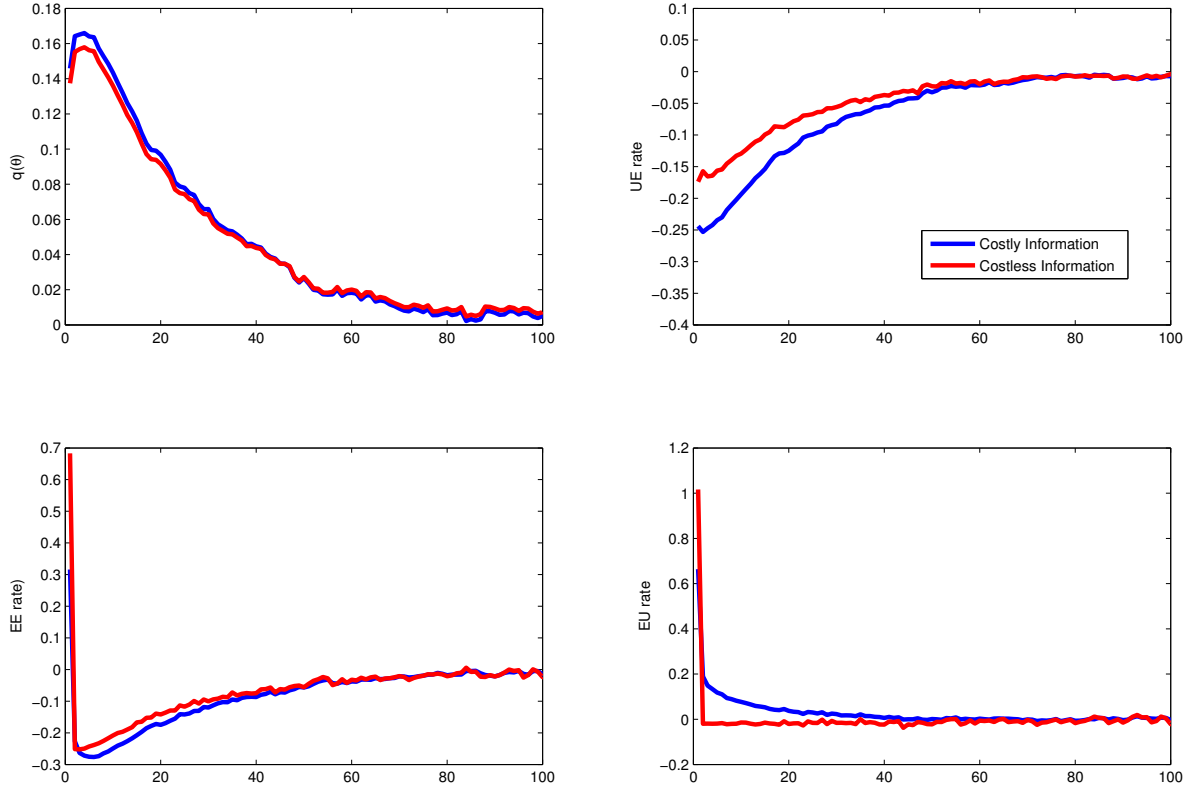


Figure 4: Response in labor flows to a 1 standard deviation drop in productivity

Importantly, while the costly information model amplifies the response of unemployment to an adverse aggregate productivity shock, it dampens the rise in the vacancy yield - defined as the ratio of hires to vacancies - relative to the full information model. Vacancy yield in the costly information model, in fact, falls on impact and subsequently increases to a high of less than 2 percent. In contrast, the vacancy yield in the full information model jumps up on impact and increases by about 4 percent. Recall, that the full information model is akin to the standard DMP search model which over-predicts the response of the vacancy yield relative to the data. In contrast, our model predicts that the rise in the vacancy yield is muted in a recession, similar to the response of the vacancy yield in the data during the Great Recession.

The divergence in the job-filling rate as suggested by the standard search model and the observed vacancy yield during the Great Recession led many commentators to suggest that matching efficiency had declined substantially over that period. In both the full information and costly information model, match efficiency is an endogenous object and stems from the firm's

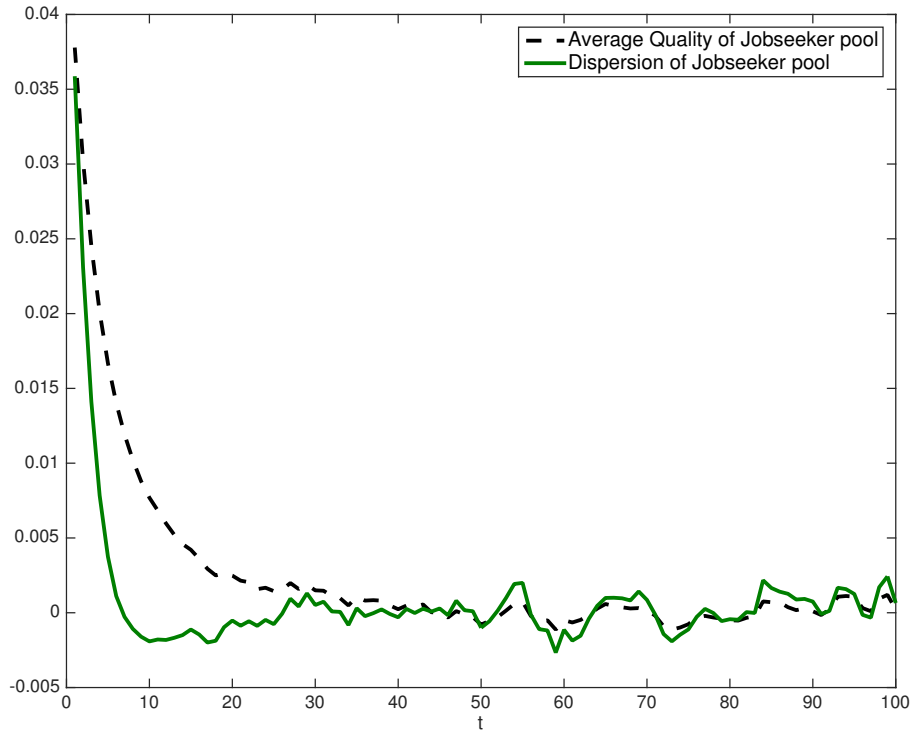


Figure 5: Average quality of job-seeker pool and its dispersion rises in a recession

recruiting decisions. The standard DMP search model makes the assumption that all workers are homogeneous and hence presumes no fluctuations in match efficiency. Here, match efficiency can change over the business cycle as workers are heterogeneous and firms have an option to reject workers perceived to be low quality. The key differences between the results of the full information and costly information models stem from the degree of precision in their information gleaned about the worker they meet. Thus, in both these models, the level of matching efficiency is represented by the firm's probability of acceptance,  $\mathcal{P}(a)$ .

Note that in the lower left hand corner of Figure 3, the firm's probability of acceptance in the costly information model dives to a further extent than in the full information model. Recall that at the start of the recession, the cost of information rises by above 40 percent. Getting precise signals to discover the worker's type is highly costly. As the firm does not receive perfect information about the worker's effective productivity at the time of meeting, he faces a risk of making a mistake, earning negative profits with the worker in the current period and firing him in the next period. This non-zero probability of making a mistake causes the firm's unconditional

probability of accepting a worker,  $\mathcal{P}(a)$ , decline further than in the full information case. This further fall in  $\mathcal{P}(a)$  is what causes the vacancy yield to be muted in the costly information model as the probability of filling a job now depends not only on the rate at which firms meet workers but also on their acceptable probability.

Finally, we turn to the implications of our model for labor flow rates during the downturn. Notably, the biggest differences come from the UE transition rates. Again, this can be explained by the fact that rising cost of information causes the probability of acceptance to decline further in the costly information model and hence induces a larger reduction in the UE transition rate. Separation rates jump on impact in both model. After the initial shocks, firms in the full information model know the effective productivity,  $z_e$ , of the workers they meet perfectly. As such, firms in the full information model only hire workers that they will retain for sure. The separation rate in this case goes back to its steady state value after one period. In contrast, in the costly information model, firms can still make mistakes over which type of workers they hire since getting precise signals is now more costly in the recession. As such, the separation rate remains slightly elevated above its steady state level even after the initial recessionary shock.

The possibility of making mistakes in hiring initially causes the employment-to-employment (EE) transition rates to differ initially between the two models. We define the EE transition rate as the ratio of newly separated workers who found employment within the same period  $t$  to the total employed at the end of last period,  $t - 1$ . While the EE rate shows little difference between the two models over time, on impact, however, the EE transition rate is higher in the full information case as firms know exactly the effective productivity of the worker at the time of meeting and readily hire any newly separated worker with high effective productivity,  $z_e$ . In contrast, firms in the costly information model do not know if the newly separated worker they meet is of high quality and hence the initial EE transition rate is more muted. Because our definition of EE transitions revolve around the newly separated workers, over time, the EE transition rates between the two models overlap as there is little difference in separation rules between the costly and costless information model after the initial shock.

Importantly, the top left panel of Figure 4 reveals that the main driver between the differences in employment (or unemployment) and the vacancy yield stems from the differential decline in  $\mathcal{P}(a)$ , our stand-in for match efficiency. In both models, the decline in the expected value of creating a job causes a similar rise in  $q(\theta)$ , the rate at which a firm meets a worker. However, it is the combination of changes in the contact rate,  $q(\theta)$ , **and** changes in the firm's acceptance rate that drives the differences in vacancy yield, unemployment rates and the transition rate from unemployment to employment.

A question arises whether the same effects are observed in a boom. Figure 6 traces out the responses of aggregate variables to a one standard deviation increase in productivity. As expected, unemployment falls by more in the costly information model and the vacancy yield observes less of a decline. However, notice that the cost of information falls drastically in a boom, by almost 200 percent. In a boom, both the dispersion and average quality of job seekers fall. The fall in the dispersion of job-seekers drives down the cost of information. Since information is cheap and since firms accept almost any worker in a boom, the probability of acceptance  $\mathcal{P}(a)$  rises to the same extent on impact in a boom.

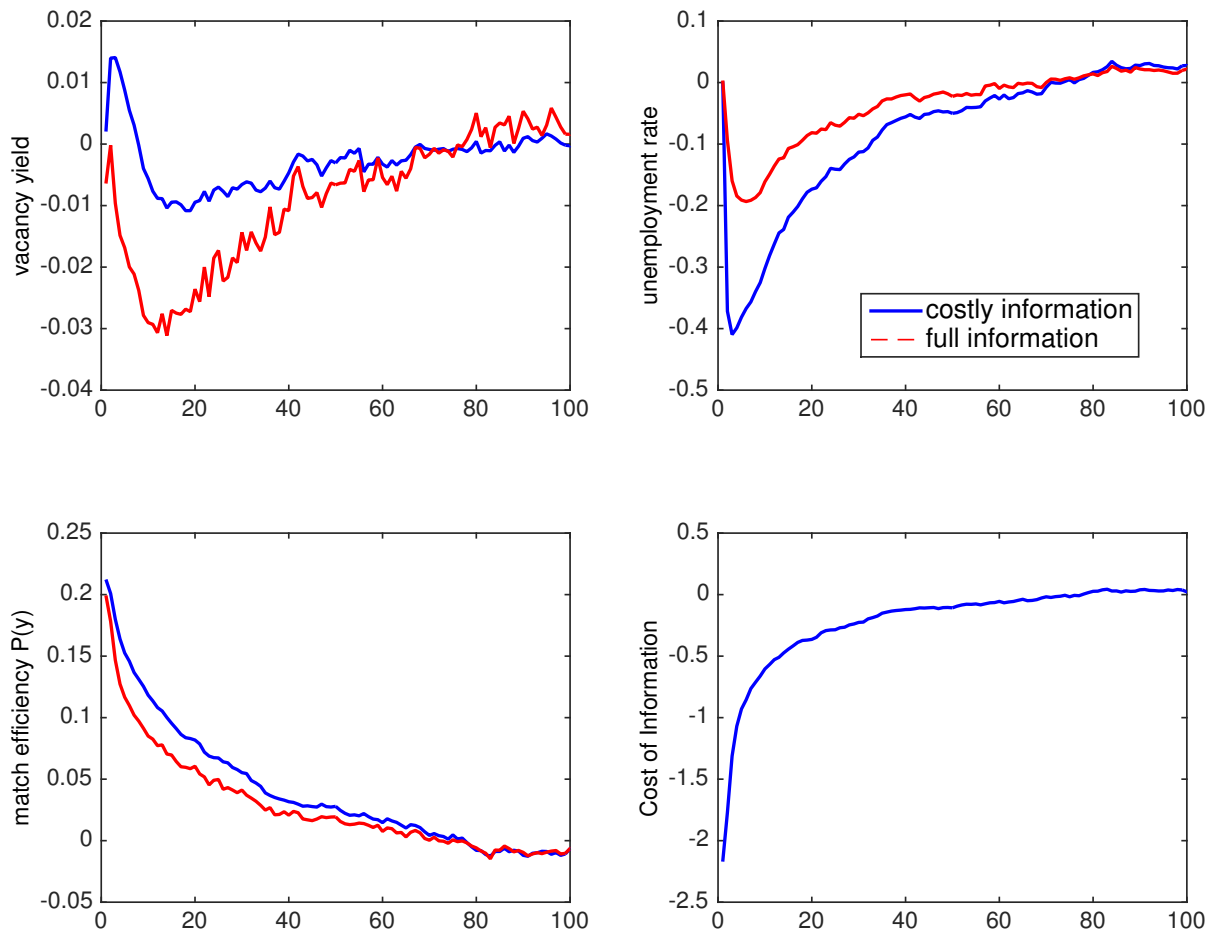


Figure 6: Response of aggregate variables in a boom

Although the probability of acceptance  $\mathcal{P}(a)$  is largely the same across both the full and

costly information models, the increase in probability of acceptance has differential impact on employment flows across the two models. Figure 7 reveals that drastic decline in the cost of information and consequent rise in the probability of acceptance leads to larger increases in the EE and elevated UE rates in the costly information model relative to its steady state value. As always, these transition rates lag their counterparts in the full information model as firms still do not have complete information and are subject to making mistakes. In this case, firms may reject individuals who they should have instead hired. As such, the UE rate and EE rate remain elevated as firms may fail to hire the correct individual. In contrast, the firms under the full information model always correct hire individuals who exceed the acceptable threshold of effective productivity.

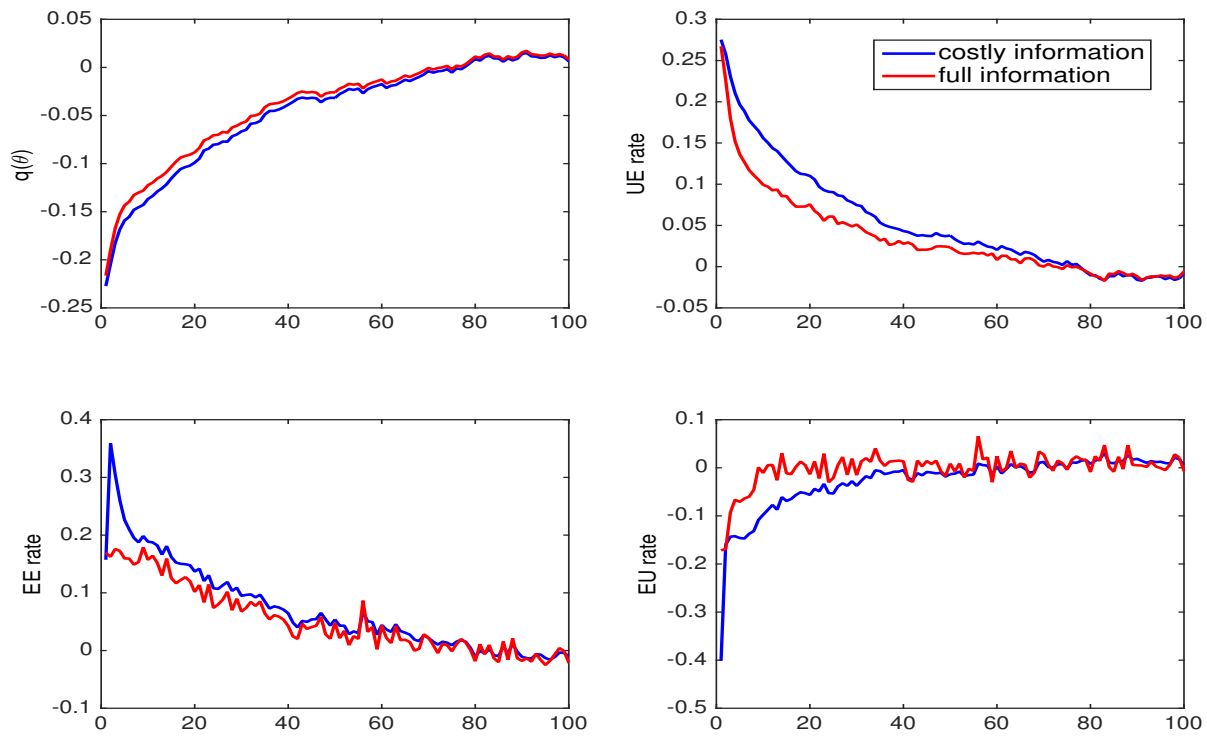


Figure 7: Transition rates in a boom

## 5 Conclusion

This paper shows that in a model with heterogeneous workers, selective hiring standards can cause the composition and dispersion of the unemployment pool to vary over the business cycle and lead to countercyclical movements in the costs of recruiting. These countercyclical movements in the cost of recruiting have strong implications for the firm’s decision on how much information to obtain about a job candidate and consequently, on his hiring decision. We show how the firm’s probability of acceptance can proxy for the labor market match efficiency and demonstrate how the model with rationally inattentive employers can deliver more amplification in the unemployment rate and a more muted hiring rate in recessions than the full information case.

## References

- Barnichon, Regis and Andrew Figura**, “Labor Market Heterogeneity and the Aggregate Matching Function,” *American Economic Journal: Macroeconomics*, October 2015, 7 (4), 222–49.
- 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.
- Hall, Robert E. and Paul R. Milgrom**, “The Limited Influence of Unemployment on the Wage Bargain,” *American Economic Review*, September 2008, 98 (4), 1653–74.
- 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.
- 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.
- Mongey, Simon, Gianluca Violante Violante, and Alessandro Gavazza**, “What Shifts the Beveridge Curve? Recruitment Effort and Financial Shocks,” 2014 Meeting Papers, Society for Economic Dynamics 2014.
- Nakamura, Emi**, “Layoffs and lemons over the business cycle,” *Economics Letters*, April 2008, *99* (1), 55–58.
- 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.
- 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

The reformulated first-stage problem in Lemma 1 can be expressed as the following Lagrangian:

$$\begin{aligned} \mathcal{L} = & \int_{z,e} \gamma(z, e | a) \mathbf{x}(a, z, e) dG(z, e | a) \\ & - \lambda \left[ -\mathcal{P}(a) \log \mathcal{P}(a) - [1 - \mathcal{P}(a)] \log [1 - \mathcal{P}(a)] \right. \\ & \left. + \int_{z,e} \left\{ \gamma(z, e | a) \log \gamma(z, e | a) + [1 - \gamma(z, e | a)] \log [1 - \gamma(z, e | a)] \right\} dG(z, e | a) \right] \\ & + \int_{z,e} \zeta(z, e) \gamma(z, e | a) dG(z, e | a) - \int_{z,e} \mu(z, e) (\gamma_i(z, e | a) - 1) dG(z, e | a) \end{aligned}$$

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  $\gamma(z, e | a)$ :

$$\mathbf{x}(a, z, e) - \lambda \left[ -\ln \frac{\mathcal{P}(a)}{1 - \mathcal{P}(a)} + \ln \frac{\gamma(z, e | a)}{1 - \gamma(z, e | a)} \right] + \zeta(z, e) - \mu(z, e) = 0 \quad (16)$$

with complementary slackness conditions

$$\mu(z, e) [1 - \gamma(z, e | a)] = 0 \quad (17)$$

$$\zeta(z, e) \gamma(z, e | a) = 0 \quad (18)$$

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

$$\gamma(z, e | a) = \frac{\mathcal{P}(a) e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 - \mathcal{P}(a) \left[ 1 - e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} \right]} \quad (19)$$

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

$$1 = \int_{z,e} \frac{e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 - \mathcal{P}(a) \left[ 1 - e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} \right]} dG(z, e | a) \quad (20)$$



## 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:

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

Now consider the costless information case which corresponds to the limit in which  $\lambda \rightarrow 0$ . First consider an applicant  $(a, z)$  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.

$$\lim_{\lambda \rightarrow 0} \gamma(z, e | a) = \lim_{\lambda \rightarrow 0} \frac{\mathcal{P}(a) e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 + \mathcal{P}(a) \left[ e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} - 1 \right]} = 0 \quad (21)$$

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.

$$\lim_{\lambda \rightarrow 0} \gamma(z, e | a) = \lim_{\lambda \rightarrow 0} \frac{\mathcal{P}(a) e^{\frac{\mathbf{x}(a, z, e)}{\lambda}}}{1 + \mathcal{P}(a) \left[ e^{\frac{\mathbf{x}(a, z, e)}{\lambda}} - 1 \right]} = \lim_{\lambda \rightarrow 0} \frac{\mathcal{P}(a) \mathbf{x}(a, z, e) e^{-\frac{\mathbf{x}(a, z, e)}{\lambda^2}}}{\mathcal{P}(a) \mathbf{x}(a, z, e) e^{-\frac{\mathbf{x}(a, z, e)}{\lambda^2}}} = 1 \quad (22)$$

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