Rational Inattention in Hiring Decisions *

Sushant Acharya
FRB New York

Shu Lin Wee
Carnegie Mellon University

May 14, 2017

Abstract

We provide an information-based theory of match efficiency. Rationally inattentive firms have limited capacity to process information and cannot perfectly identify suitable applicants. As losses from hiring the wrong worker are amplified in a recession, firms seek to be more selective in their hiring. Inability to obtain sufficient information about applicants causes firms to err on the side of caution and accept fewer applicants in order to avoid hiring unsuitable workers. Pro-cyclical acceptance rates drive a wedge between meeting and hiring rates, driving fluctuations in match efficiency. Quantitatively, our model accounts for changes 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

*The authors would like to thank Ronald Wolthoff (discussant), Mirko Wiederholt (discussant), Jess Benhabib, Pablo Cuba, Keshav Dogra, Laura Pilossoph and seminar and conference participants at Carnegie Mellon University, FRB St. Louis, Georgetown University, Penn State, UMD, NUS, SMU, NYU Stern IUB Workshop, SED 2016 at Toulouse, Search and Matching Conferences in Amsterdam, Philadelphia and Istanbul. 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 , Wee: shuwee@andrew.cmu.edu

1 Introduction

The Great Recession witnessed a severe spike in unemployment rates as well as a tripling in the ratio of unemployed job-seekers to 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[2] declined during the Great Recession. In this paper, we provide an information-based theory of match efficiency. In particular, we show how changing aggregate productivity and the composition of job-seekers over the business cycle affects hiring decisions of firms which in turn drive movements in match efficiency over the business cycle.

We consider a search model in which workers permanently differ in their ability. A firm's profitability is affected by aggregate productivity, a worker's ability and a match-specific component. A worker's ability is known to the worker and to her current employer, but not to a new firm. New firms can conduct interviews to learn about the suitability of an applicant. Given the information they acquire about the job-seeker, firms reject applicants whom they perceive to be unsuitable. Match efficiency is defined as the average probability that a firm accepts a worker and is distinct from the rate at which a firm contacts a worker. The average acceptance probability depends on how much information firms can acquire as well as the costliness of making a mistake in hiring the wrong worker. In our model, firms are *rationally inattentive* and have a limited capacity to process information about an applicant. When firms face a lot of uncertainty regarding the set of job-seekers they encounter, they need to process more information about the job-seeker to determine her suitability for production.

The losses associated with hiring an unsuitable worker and the uncertainty surrounding the pool of unemployed job-seekers vary over the business cycle. When aggregate productivity is high, firms are willing to hire almost any worker except those deemed to be very poor matches. During a recession, firms require a worker who can compensate for the fall in aggregate productivity. Since the losses from hiring an unsuitable worker are larger during a recession, firms seek to acquire more information about the job-seeker to determine her suitability for production. Firms, however, have finite information processing capacity. This limited capacity to decipher a job applicant's suitability for production increases the firm's incidence of making a mistake, causing firms to optimally err on the side of caution and reject applicants more often in the downturn. Overall, firms' attempts to avoid hiring unsuitable workers (Type I error) leads them to reject a larger fraction of

[1] *"Even with unemployment hovering around 9%, companies are grousing that they can't find skilled workers, and filling a job can take months of hunting."* (Cappelli, 2011) in Wall Street Journal on October 24, 2011.

[2] Match efficiency is the equivalent of a Solow Residual in a matching function. We formally define match efficiency when describing Figure 1 below.

suitable workers (Type II error).

These higher rejection rates in turn cause the distribution of unemployed job-seekers to become more varied. When firms reject job-seekers more often on average, they inadvertently reject high ability workers along with other ability types. This causes not only the average quality of the pool of unemployed job seekers to increase but also has the effect of elevating the uncertainty that the firm faces regarding these job-seekers. Higher uncertainty reinforces the firms' desire for more precise information, which in turn translates into even higher rejection rates, further weighing on measured match efficiency.

A large literature has argued that firms are more selective during downturns and their higher hiring standards causes the average quality of the unemployment pool to improve (See for example, Kosovich (2010), Lockwood (1991), Nakamura (2008) and Mueller (2015) among others.). Our paper adds to this literature and suggests that informational constraints are an important factor in generating the correct co-movement in aggregate labor market variables when the average quality of the unemployed is countercyclical. Absent any constraints on the firm's ability to process information, a higher average quality in the pool of job-seekers implies that firms meet more productive applicants on average during a downturn. This, in turn makes firms less inclined to reduce job creation, and exacerbates the employment volatility puzzle.[3] Moreover, given that firms meet higher quality applicants during the downturn, they are more likely to accept these applicants, implying an improvement as opposed to a decline in match efficiency.

Our paper resolves these issues. Crucially, the lack of precise information and the increased cost of making a Type I error (hiring the wrong worker) during a downturn raises the rejection rate of all job-seekers despite the improvement in average quality of the unemployment pool. These higher rejection rates increase the firm's uncertainty over job-seekers' types, further hampering the firm's ability to distinguish between applicants. Both the increased cost of making a mistake and higher uncertainty work against the improvement in the average quality of the unemployment pool. In response to a severe recession where productivity declines by ten percent, match efficiency dives by 10% in the rational inattention model and only completely recovers 50 months after the shock. In contrast, match efficiency in the full information model also falls on impact, but rebounds and actually rises above its steady state level by the third month as higher average quality of the unemployed counteracts the lower aggregate productivity in the economy.

The idea that firms' hiring behavior may change over the business cycle is not a new one. Davis et al. (2012) show that the implied job-filling rate from a constant-returns-to-scale matching function[4] (blue dashed line in Figure 1) diverged significantly from its empirical counterpart, the vacancy yield[5] (red solid line), during the Great Reces-

[3] Shimer (2005) showed that the standard labor search model with homogeneous workers fails in generating sufficient labor market volatility.

[4] The implied job-filling rate is computed as $(v/u)^{-0.5}$ using data from JOLTS and the CPS.

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

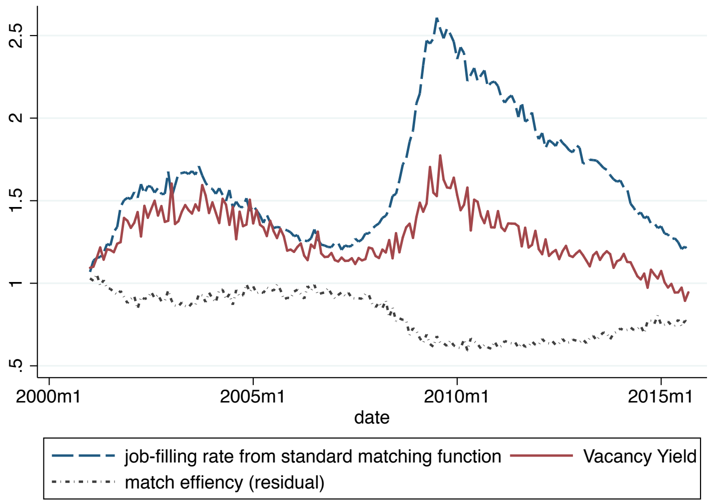


Figure 1: Vacancy Yield and Match Efficiency as computed from JOLTS and CPS

sion. Notably, measured match efficiency (gray dotted line in Figure 1) - computed as the residual variation in hires not due to the number of unemployed and vacancies - fell precisely when the two rates diverged. Davis et al. (2012) attribute this divergence to a decline in *recruiting intensity* (a catch-all term for the other instruments and screening methods firms use to increase their rate of hires). We offer a theory of recruiting intensity which is based on firms' limited ability to process information. In a recession, the desire for more information causes firms' information processing constraints to bind more, making it harder to distinguish between different types of applicants. Consequently, firms reject applicants more often. We interpret these higher rejection rates as lower recruiting intensity.

Several recent papers have also tried to examine and decompose the forces driving the decline in match efficiency. Hall and Schulhofer-Wohl (2015) study how the search effectiveness of different job-seekers over the business cycle can affect matching rates. Under the standard estimates of matching function elasticity,[6] Hornstein and Kudlyak (2016) find that search effort is countercyclical and counteracts the decline in job-finding rates stemming from procyclical compositional changes in the search efficiency of the unemployed, suggesting that firms' recruiting behavior is still important towards explaining fluctuations in aggregate match efficiency.

In closely related work, Gavazza et al. (2014) consider a model of costly recruiting effort and find that productivity shocks together with financial shocks can explain fluctuations in aggregate match efficiency. Our model does not feature an explicit cost of recruiting. Rather, fluctuations in the shadow value of information drive changes in firms' recruiting behavior/acceptance rates. Thus, our paper suggests that even in the absence of explicit recruiting costs, firms' hiring behavior can change drastically over the business

[6] See for example Petrongolo and Pissarides (2001).

cycle. In fact, our models matches movements in match efficiency in the data in response to changes only in aggregate productivity. In another related 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. The above papers, together with ours, highlight how firms' recruiting effort and time-varying hiring standards can explain fluctuations in match efficiency over the business cycle. Unlike Sedlacek (2014) and Gavazza et al. (2014), our model can also explain how firms reduce recruiting efforts even in the face of improving average quality of the unemployed during downturns. Importantly, the aforementioned papers do not feature ex-ante worker heterogeneity and are thus silent on how changing composition of the unemployment pool might affect hiring decisions.

Our paper also speaks to a large literature which argues that firms use unemployment duration as an additional tool to evaluate the suitability of a worker. Resume audit studies such as Kroft et al. (2013) and Eriksson and Rooth (2014) find evidence that firms are less likely to call back workers who have longer unemployment spells. Furthermore, Kroft et al. (2013) show that call-back rates exhibit a gentler decline with unemployment duration in areas with weak economic conditions, suggesting that unemployment duration is a poorer indicator of worker quality when labor markets are slack. Our model is able to generate both these features. In our model, firms can observe and use unemployment duration in addition to interviews to inform them about an applicant's suitability. Because firms reject applicants who they deem to be unsuitable, applicants who have been unemployed for longer are more likely to be of low quality. Since unemployment duration conveys some information about the job-seeker's type, firms in our model are less likely to employ applicants who have been unemployed for longer. However, unemployment duration is a weaker indicator of worker quality as lower vacancy postings and firms' acceptance rates for all workers cause average unemployment duration to lengthen. Because all job-seekers are unemployed for longer regardless of their type, a longer unemployment duration conveys less information about a worker's type. As such, our model features the relative job-finding rate of individuals more than 6 months unemployed to those one month unemployed as being 5% higher in a recession than in a boom.

While our paper studies how rational inattention and costly information acquisition can affect the hiring decisions of firms, there is a large literature that has focused on how rational inattention can affect other aspects of search behavior. Cheremukhin et al. (2014) consider how the costliness of processing information can affect the degree of sorting between firms and workers. Briggs et al. (2015) consider how rational inattention can rationalize increased labor mobility and participation amongst older workers late in their working life. Bartos et al. (2016) directly monitor information acquisition by firms in labor and housing markets and find that the processing of information and selection choices resembles that of decision-makers with rational inattention. Finally, Lester and Wolthoff (2016) study contracts which deliver the efficient allocations of heterogeneous workers across firms in an environment where firms face explicit interview costs.

The rest of this paper is organized as follows: Section 2 introduces the model with rational inattention and characterizes how changes in economic conditions affect a firm's hiring decisions, and consequently, match efficiency. Section 3 discusses our calibration approach while Section 4 documents our quantitative exercises. Section 5 contains a discussion regarding the key assumptions underlying our model. Section 6 concludes.

2 Model

Time is discrete. We describe the economic agents that populate this economy.

Workers The economy consists of a unit mass of risk-neutral workers who discount the future at rate β. Each worker has a permanent productivity-type z drawn from a finite set $\mathcal{Z}$. 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 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, z is the worker type 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 from a finite set $\mathcal{E}$ 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 (high z types) to be deemed as bad matches if they draw a low e shock. Likewise, low z types can 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 contacts a worker at a rate q_t. This contact rate depends on the total number of vacancies, v_t, and job-seekers, l_t, according to a constant returns to scale matching technology $m(v_t, l_t)$. 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.

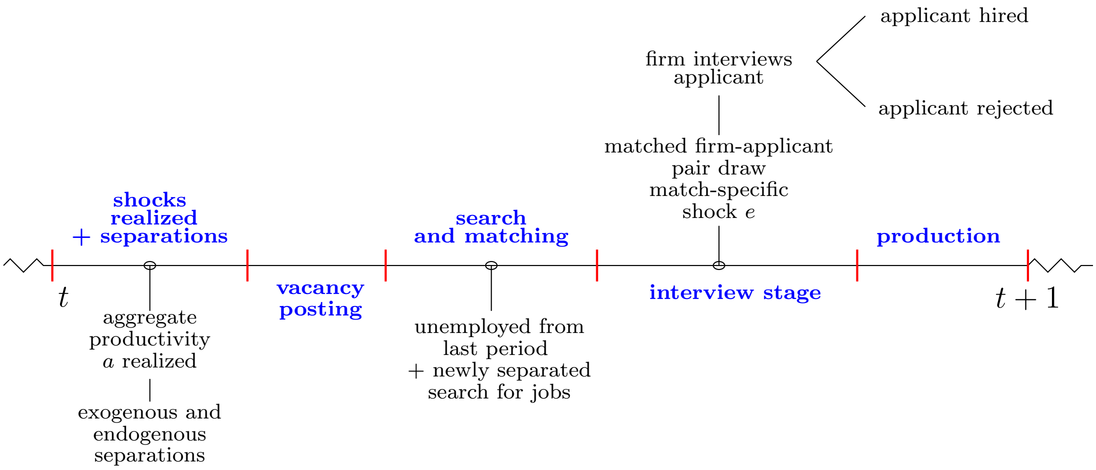


Figure 2: Timeline

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 per 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 although it can perfectly observe the applicant's unemployment duration. The firm can choose how much information to acquire to reduce its uncertainty about the worker's ze. We refer to this process as an *interview*. Since individuals who are unemployed for longer are more likely to have met a firm and failed an interview, the firm may choose to acquire different information for applicants with different unemployment spells. Given the information revealed in the interview, the firm decides whether or not to hire a worker. The firm can perfectly identify ze once production has taken place and fire a worker ex-post if she turns out to be unsuitable for the job. 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 aggregate productivity a and the distribution of (z, e) type job-seekers of each duration length τ. The hiring strategy of a firm can be described as a two-stage process. In the *first-stage*, having observed the applicant's unemployment duration, the firm must devise an information strategy, i.e. design an interview. In other words, the firm chooses how informative a signal to acquire about an applicant's true effective productivity ze. In the *second-stage*, based on the information elicited from the interview, the firm decides 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 in Section 2.3. In the meantime, it is sufficient to know that σ contains information about the level of aggregate productivity and the joint distribution of job-seekers of effective productivity (z, e) with unemployment duration τ. Further denote $\mathcal{G}(z, e \mid \sigma, \tau)$ as the conditional distribution of (z, e) types in the pool of job-seekers with unemployment duration τ given aggregate state σ and let $g(z, e \mid \sigma, \tau)$ be the associated probability mass function. In equilibrium, $\mathcal{G}(z, e \mid \sigma, \tau)$ also denotes the firm's prior belief about an applicant's (z, e) conditional on the job-seeker having duration τ in aggregate state σ.[7]

In the second-stage, the firm has already chosen an information strategy and received signals $\mathbf{s}$ about the applicant's effective productivity ze. Denote the joint-posterior belief of the firm about this applicant's true (z, e) type 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. Rejecting the applicant yields the firm a payoff of zero. As firms may be unable to ascertain the applicant's true type (z, e) even after receiving signals, the payoff from hiring an applicant can still vary based on the actual (z, e) and is thus a random variable. Since the firm can guarantee itself zero payoffs by rejecting an applicant, it only hires a worker if the expected value from hiring, under the posterior belief Γ, is non-negative. 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:*

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

where $\mathbf{x}(a, z, e)$ *denotes the payoff to the firm if it hires a type* (z, e) *applicant.*

2.1.2 First-stage Problem

In the first stage, the firm chooses an information strategy to determine the applicant's effective productivity ze. Firms, however, can only process finite amounts of information and may not be able to determine an applicant's type with certainty. We model limited information processing capacity of the firm as an entropy-based channel capacity constraint following Sims (2003). As per the rational inattention literature, uncertainty over the types of job-seekers is measured in terms of entropy, and mutual information measures the reduction of uncertainty about a worker's effective productivity. The definition below formalizes these concepts.

[7] See Section 2.3 for more information.

Definition 1. *Consider a discrete 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 an 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 $p(x \mid \mathbf{s})$. Mutual information is then given by:*

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

where $\mathcal{I}$ measures the information flow and denotes the reduction in the agent's uncertainty about X by virtue of getting signals $\mathbf{s}$*.*

We assume that the maximum amount of information that a firm can process in a period is constrained by a finite channel capacity $\chi > 0$.[8] We can then write the constraint on the information flow as:

$$\mathcal{I}(\mathcal{G}, \Gamma \mid \sigma, \tau) = H(\mathcal{G}(\cdot \mid \sigma, \tau)) - \mathbb{E}_{s} H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) \leq \chi \tag{1}$$

where $H(\mathcal{G})$ is the firm's initial uncertainty given its prior $\mathcal{G}$ and $\mathbb{E}_{\mathbf{s}} H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right)$ is the firm's residual uncertainty after obtaining signals about the worker. A firm's choice of information strategy must respect the constraint on information flows. In this respect, the firm's information strategy is akin to 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 to help it make a more informed decision over whether to accept or reject an applicant. Each additional question, however, consumes the finite processing capacity a firm possesses and limits its ability to determine the applicant's suitability. Modeling the interview process as above is particularly natural since the information flow measured in terms of reduction in entropy is proportional to the expected number of questions needed to implement an information strategy.[9] We now describe the firm's first stage problem.

Through the interview, firms can choose the informativeness of signals $\mathbf{s}$ to reduce its uncertainty about the applicant's type. More informative signals consume more channel capacity. The following definition characterizes an information strategy of the firm.

Definition 2 (Information Strategy)**.** *The information strategy of a firm who meets an applicant with unemployment duration τ in aggregate state σ is given by a joint distribu-*

[8] In an earlier version of the paper, we allowed for firms to choose their channel capacity (variable attention choice) given a constant marginal cost, λ, of paying more attention. Our results are qualitatively similar whether we allow for a variable attention choice or a fixed attention level.

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

tion of signals $\mathbf{s}$ *and types,* $\Gamma(z, e, \mathbf{s} \mid \sigma, \tau)$ *such that:*

$$\mathcal{G}(z, e \mid \sigma, \tau) = \int_{\mathbf{s}} d\Gamma(z, e, \mathbf{s} \mid \sigma, \tau) \tag{2}$$

Equation (2) is simply a consistency requirement and implies that the firm is only free to choose $\Gamma(\mathbf{s} \mid z, e, \sigma, \tau)$. Thus, an information strategy entails the firm choosing the signals it wants to observe when it meets an applicant of type (z, e).

Note that the information strategies are indexed by τ. Importantly, unemployment duration conveys additional information about the applicant's ability to the firm, allowing the firm to refine its prior about an applicant's type z. A worker could be unemployed because she failed to meet a firm or because she met a firm but was rejected. The longer the unemployment spell, the higher the likelihood that the individual has met a firm and was rejected, suggesting that applicants with longer unemployment durations possess low z. Since the pool of job-seekers with higher durations are less likely to be high z-types, the probability mass, $g(z, e \mid \sigma, \tau)$ differs across τ. As the firm's initial uncertainty over a job-seeker of duration τ and the associated expected payoffs from hiring that applicant are affected by the probability masses, $g(z, e \mid \sigma, \tau)$, the firm's information strategy may differ depending on the duration τ of the applicant it meets.

The extent to which the firm's processing constraint (1) binds, thus, depends not only on the informativeness of the signals chosen but also on the firm's prior about the distribution of job-seekers, $\mathcal{G}(z, e \mid \sigma, \tau)$. As such, how much the firm's information processing constraint binds depends on aggregate conditions, σ. The Proposition below summarizes the firm's hiring problem.

Proposition 2 (First-Stage Problem of a Firm)**.** *The firm's first-stage problem involves choosing an information strategy to maximize ex-ante payoffs from the second-stage for each duration* τ*:*

$$\mathbf{V}(\sigma, \tau) = \max_{\Gamma \in \Delta} \sum_{z} \sum_{e} \int_{s} J\left[\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right] d\Gamma(\mathbf{s} \mid z, e, \sigma, \tau) g(z, e \mid \sigma, \tau) \tag{3}$$

subject to the information processing constraint (1)*.*

$J\left[\Gamma\left(z, e \mid \mathbf{s}, \sigma, \tau\right)\right]$ denotes the ex-ante payoff from the second stage for a type (z, e) applicant given signals $\mathbf{s}$. Since the firm does not know the applicant's true (z, e), the firm's expected payoff is the weighted sum over the signals $d\Gamma\left(\mathbf{s} \mid z, e, \sigma, \tau\right)$ and the possible types of job-seekers, $g\left(z, e \mid \sigma, \tau\right)$. The problem of the firm specified in Proposition 2 is not trivial to solve as it allows firms to choose signals of any form. Fortunately, the problem can be reformulated into a more tractable form. Rather than solving for the optimal signal structure, following Matejka and McKay (2015), we solve the identical but transformed problem in terms of state-contingent choice probabilities and the associated payoffs.

Let $\mathcal{S}$ be the optimally chosen set of signals that lead the firm to take the action *hire* for an applicant of type (z, e) of duration τ in state σ. Under this optimal information strategy, the firm hires an applicant of type (z, e) with duration τ with probability $\gamma(z, e \mid \sigma, \tau)$ which is the same as the probability of drawing signal $\mathcal{S}$ conditional on the applicant being of type (z, e):

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

The average probability of hiring a worker of duration τ in state σ, denoted by $\mathcal{P}(\sigma, \tau)$, is then given by:

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

The following Lemma presents the reformulated problem:

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

$$\mathbb{V}(\sigma, \tau) = \max_{\gamma(z,e|\sigma,\tau) \in [0,1]} \sum_z \sum_e \gamma(z, e \mid \sigma, \tau) \mathbf{x}(a, z, e) g(z, e \mid \sigma, \tau) \tag{4}$$

subject to:

$$\mathcal{H}(\mathcal{P}) - \sum_z \sum_e \mathcal{H}\Big(\gamma(z, e \mid \sigma, \tau)\Big) g(z, e \mid \sigma, \tau) \quad \leq \quad \chi \tag{5}$$

where $\mathcal{H}(x) = -x \ln x - (1 - x) \ln(1 - x)$.

Proof. The proof is very similar to Appendix A of Matejka and McKay (2015). □

Intuitively, the LHS of (5) measures the information flow based on the optimal signal choices as in equation (1) but expressed in terms of choice probabilities. This equivalence follows from the fact that the information flow is a strictly convex function, implying that a firm optimally associates each action with a particular signal. Receiving multiple signals that lead to the same action is inefficient as the additional information acquired is not acted upon and uses up limited channel capacity which could have otherwise been used to make better decisions. The proposition below characterizes the optimal information strategy of a firm.

Proposition 3 (Optimal Information Strategy)**.** *Under the optimal information strategy, the firm chooses signals such that the probability of hiring an applicant of type (z, e) with unemployment duration τ in aggregate state σ is given as:*

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

where $\lambda(\sigma, \tau)$ is the multiplier on (5) and represents the shadow-value of reducing uncertainty by one nat.[10] *The average probability that a firm hires an applicant of duration τ, $\mathcal{P}(\sigma, \tau)$ is implicitly defined by:*

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

Proof. See Appendix A. $\square$

2.1.3 What affects hiring decisions?

Equation (6) reveals an important feature of the information strategy. Consider two applicants with the same e and τ, but with different worker-productivity $z_1 > z_2$. Under the optimal information strategy, the following is true:

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

Equation (8) implies that the firm chooses signals such that the induced odds-ratio of accepting the more-productive applicant is proportional to the difference in the payoffs from hiring the two types of workers. The firm chooses signals so as to reduce the incidence of making a Type II error, allowing it to accept more productive applicants more often on average. Further, equation (8) reveals that an increase in the shadow value of information, $\lambda(\sigma, \tau)$, reduces the difference between $\gamma(z_1, e \mid \sigma, \tau)$ and $\gamma(z_2, e \mid \sigma, \tau)$. As λ becomes larger, firms are starved of information and are increasingly unable to distinguish between different types. In the limit as $\lambda \to \infty$, the firm's limited processing capacity renders it incapable of distinguishing between different types of applicants. In this case, the firm's posterior belief is the same as its prior, and the firm applies the same acceptance probability $\gamma(\cdot \mid \sigma, \tau)$ to all applicants of duration τ.

Lemma 2 (Information Strategy with Costless Information)**.** *Given any distribution of $G(z, e)$, if $\chi \geq H(G)$, then firms can replicate the hiring decisions that arise under full information. Consequently, the induced probability of hiring a particular type of worker (z, e) with duration τ under the optimal information strategy is given by:*

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

Proof. See Appendix B. $\square$

[10] When using a logarithm with exponential base, entropy is measured in nats. An equivalent but alternative way to measure entropy is to use a logarithm with base 2. In this case, the measure of entropy would be in terms of bits.

If the information processing constraint does not bind, the firm can ascertain the applicant's (z, e) type 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 $\mathbf{x}(a, z, e) < 0$.

Uncovering the forces at play: A Static Limit To highlight the forces that affect a firm's hiring decision, we consider the static limit of the model in which $\beta = 0$. We shut-down the match-quality e dimension of heterogeneity and assume that there are only two types of workers $z_H > z_L$ in proportion $\alpha \geq 0.5$ and $1 - \alpha$ respectively. In this case, the firm's hiring problem becomes:

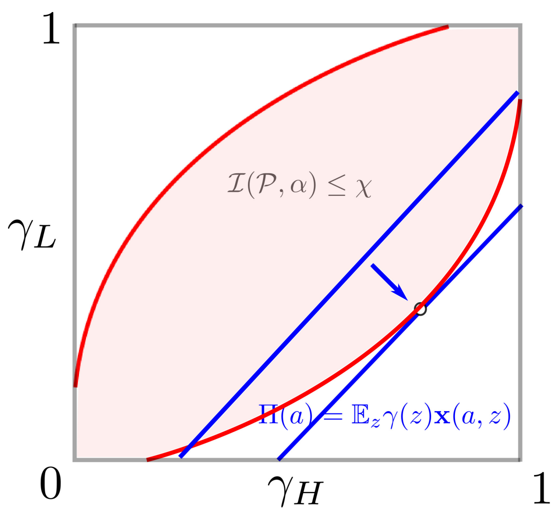
$$\Pi(a) = \max_{(\gamma_H, \gamma_L) \in [0,1]^2} \alpha \gamma_H \mathbf{x}(a, z_H) + (1 - \alpha) \gamma_L \mathbf{x}(a, z_L)$$

s.t.

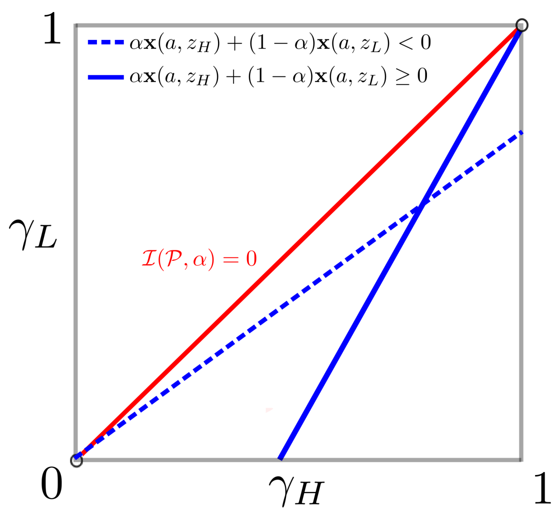
$$\mathcal{H}\Big(\mathcal{P}(a)\Big) - \alpha \mathcal{H}(\gamma_H) - (1 - \alpha) \mathcal{H}(\gamma_L) \leq \chi \tag{10}$$

where $\mathbf{x}(a, z) = az - b$ and $\mathcal{P}(a) = \alpha \gamma_H + (1 - \alpha) \gamma_L$ denotes the average probability that the firm hires an applicant when aggregate productivity is a.

Suppose aggregate productivity is such that $a \in \left[\frac{b}{z_H}, \frac{b}{z_L}\right]$. In this interval, $\mathbf{x}(a, z_H) > 0 > \mathbf{x}(a, z_L)$ and a firm only wants to hire the z_H applicant. Given a limited channel capacity χ, the firm, however, may not be able to identify a z_H applicant perfectly. Figure 3 depicts the hiring decisions for different values of χ.



(a) $0 < \chi < \overline{\chi}$



(b) $\chi = 0$

Figure 3: Optimal Hiring Decisions

The firm's unconstrained choice is given by $(\gamma_H = 1, \gamma_L = 0)$. The level of χ determines how close a firm's decision can be to the unconstrained choices. For $\chi <$

$-\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha) \equiv \overline{\chi}$, the unconstrained choice is not feasible. The shaded area in Figure 3a shows the feasible choices that a firm can make for a channel capacity χ for all $\chi \in (0, \overline{\chi})$. Notice that not discriminating between the types $\gamma_H = \gamma_L \in [0,1]$ is always feasible but not necessarily optimal. In fact, Figure 3b shows that if the firm cannot process any information, choosing $\gamma_H = \gamma_L$ (the diagonal) is its only feasible choices. Overall, Figure 3 reveals that an information strategy that tries to distinguish between applicant types by choosing (γ_H, γ_L) away from the diagonal corresponds to more informative signals and thus, requires more channel capacity.

The parallel blue lines in figure 3a depict the iso-profit curves with profits increasing in the south-east direction and the highest profit achieved at the point $(\gamma_H = 1, \gamma_L = 0)$. Correspondingly, the optimal choice of (γ_H, γ_L) must lie on the south east frontier of the feasible set. In any interior optimum, the following is true:

$$\frac{\alpha \mathbf{x}(a, z_H)}{(1-\alpha)\mathbf{x}(a, z_L)} = \frac{\alpha \left[\mathcal{H}'\Big(\mathcal{P}\Big) - \mathcal{H}'(\gamma_H)\right]}{(1-\alpha)\left[\mathcal{H}'\Big(\mathcal{P}\Big) - \mathcal{H}'(\gamma_L)\right]} \tag{11}$$

Equation (11) implies that in any interior optimum, the iso-profit curves are tangent to the boundary of the constraint set.[11] The firm wants to choose the highest γ_H and the lowest γ_L possible but equation (11) reveals that in trying to increase γ_H, the firm is forced to choose a higher γ_L so as to respect the information processing constraint. Thus, the constraint limits the firm's ability to acquire signals which help it distinguish between the z_H and z_L applicant. In the extreme where $\chi = 0$, the only feasible choices lie along the diagonal, i.e. $\gamma_H = \gamma_L = \gamma$. The optimal choices are then characterized by a *bang-bang* solution with firms hiring any applicant, $\gamma = 1$ if $\alpha\mathbf{x}(a, z_H) + (1-\alpha)\mathbf{x}(a, z_L) \geq 0$, and rejecting all applicants, $\gamma = 0$, otherwise. The former is depicted by the intersection of the solid blue and red lines in Figure 3b, and the latter by the intersection of the dashed blue line and red line.

How does a fall in aggregate productivity affect optimal choices? Next, we show how a decline in aggregate productivity - *a recession* - affects hiring decisions. While changes in a do not affect the information processing constraint, and thus the set of feasible choices, changes in a do affect the slope of the iso-profit curves. Figure 4a[12] shows that a lower a corresponds to flatter iso-profit curves[13], shifting the point of tangency south-west, and lowering optimal γ_H and γ_L.

[11] The LHS is the slope of the iso-profit curve while the RHS is the slope of the constraint set.

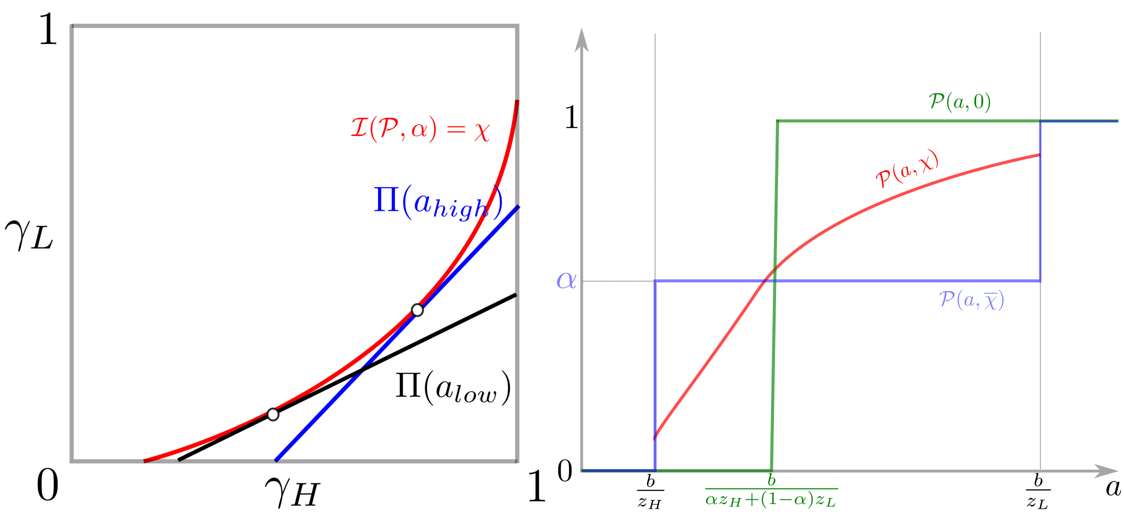
[12] Since the optimal choices of γ_H and γ_L lie on the south-east frontier of the constraint set, we omit drawing the north-west frontier in Figures 4a and 5a to avoid clutter.

[13] The slope of iso-profit curves is given by $\frac{\mathbf{x}(a,z_H)}{\mathbf{x}(a,z_L)} = \frac{az_H - b}{az_L - b}$ which is decreasing in a.

Lemma 3 (Comparative Statics)**.** *Under the optimal information strategy,* $\gamma_H(a)$ *and* $\gamma_L(a)$ *are increasing in* a*. Thus,* $\frac{\partial \mathcal{P}(a)}{\partial a} > 0$.

Proof. See Appendix C. $\square$

While profits are diminished in a recession even if firms correctly identify and hire a z_H applicant, losses are magnified if firms mistakenly hire a z_L applicant.[14] Accordingly, firms err on the side of caution in recessions and reject applicants more often. As a result, γ_H, γ_L and consequently $\mathcal{P}$ decrease with the fall in a. As outlined in Section 2.2.2, declines in $\mathcal{P}$ correspond to falls in match efficiency in our model.



(a) Change in aggregate productivity a

(b) $\mathcal{P}$ as a function of a

Figure 4: Effect of χ on $\mathcal{P}(a)$

Figure 4b shows that the average acceptance probability $\mathcal{P}(a)$ is more sensitive to changes in aggregate productivity when firms are more informationally constrained. The blue line denotes the optimal $\mathcal{P}^{FI}(a)$ under full information. In this case, firms always observe the applicant's type and never make mistakes in hiring.

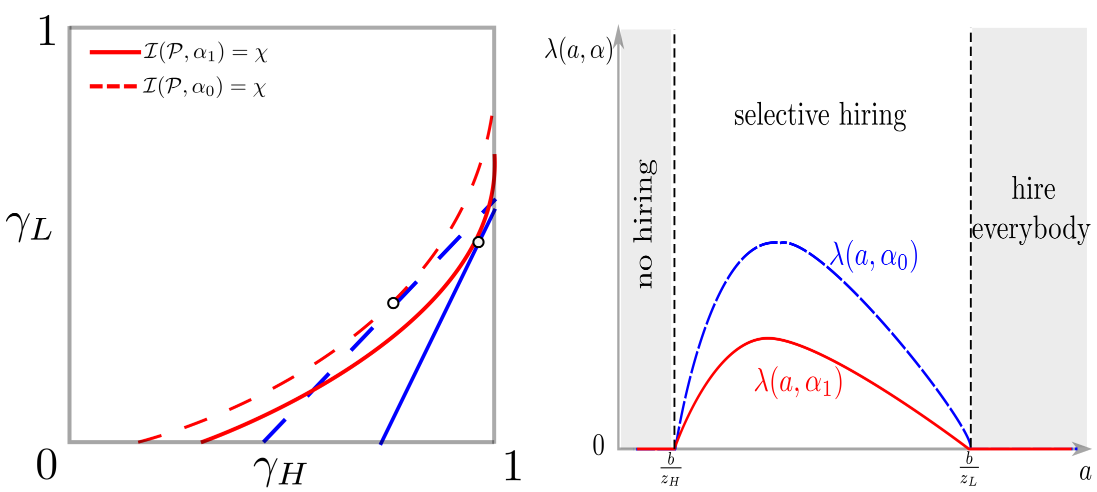
For very low levels of a: i.e. $a < \frac{b}{z_H}$, firms incur losses from hiring any applicant, as such $\gamma_H^{FI} = \gamma_L^{FI} = \mathcal{P}^{FI} = 0$ in this range. For very high a: $a > \frac{b}{z_L}$, the firm is happy to hire any applicant and $\gamma_H^{FI} = \gamma_L^{FI} = \mathcal{P}^{FI} = 1$ in this range. Figure 4b shows that in these regions, firms' decisions under FI and RI are identical. This is because even if firms are unable to process any information, choosing an information strategy in which $\gamma_H = \gamma_L \in \{0, 1\}$ is always feasible. In these regions, the shadow value of information $\lambda = 0$ and information constraints do not matter.

Next consider $a \in [{}^{b}/_{z_H}, {}^{b}/_{z_L}]$. In this range, under FI, firms only hire the H type, $\gamma_H = 1, \gamma_L = 0$. As a result, the optimal $\mathcal{P}^{FI}(a) = \alpha$ in this region as shown by the

[14]Here we still assume that a recession observes a lower a but a is still in the region $[{}^{b}/_{z_H}, {}^{b}/_{z_L}]$. If the new a was lower than ${}^{b}/_{z_H}$, then the firm would make negative profits if it hired any type of applicant.

blue line in Figure 4b. Under rational inattention, firms may not be able to distinguish z_H types, implying that the shadow value of information in this range may be positive (Figure 5b). In fact, as long as $\chi < \overline{\chi}$, Proposition 3 shows that ${}^{\partial \mathcal{P}(a)}/_{\partial a} > 0$ in this range and the implied $\mathcal{P}(a)$ under the optimal information strategy is depicted by the upward sloping red-line in Figure 4b. Thus, relative to the full information case, small changes in a only result in large changes in $\mathcal{P}(a)$, our measure of match efficiency, when firms are informationally constrained.[15]

The relationship between $\lambda(a)$ and a is non-monotonic. For extremely low or high levels of aggregate productivity, firms do not value information and their decisions are unaffected by limited information processing capacity, implying that $\lambda = 0$. For $a \in [\frac{b}{z_H}, \frac{b}{z_L}]$, hiring a z_L type worker incurs losses and firms value information as they would like to avoid hiring a z_L worker. When aggregate productivity a is close to $\frac{b}{z_H}$, the firm wants to strongly avoid z_L types. Firms value an additional unit of processing capacity at this point as it would allow them to better identify the z_L applicant, hence $\lambda > 0$. As a increases away from $\frac{b}{z_H}$, the attractiveness of hiring z_H applicants increases, causing the shadow value of information to go up since firms still want to avoid hiring z_L types. As a approaches $\frac{b}{z_L}$, losses from hiring a z_L applicant decline. The firm's concern over wrongly hiring a z_L applicant is outweighed by the benefit of hiring a z_H type, lowering the firm's need to distinguish between types and reducing λ.



(a) Effect of changing α on γ_H, γ_L

(b) Shadow value of information

Figure 5: Comparative Statics

[15] The green line corresponds to the case where $\chi = 0$. This line has an infinite slope at $a = \frac{b}{\alpha z_H + (1-\alpha) z_L}$ implying that small declines in a around this point could lead to large changes in $\mathcal{P}(a)$ causing it to fall from 1 to 0. No such change would occur in the full information case.

How does a change in the distribution of job-seekers affect optimal choices? Unlike a, an increase in α which corresponds to the fraction of z_H types in the population, affects both the slope of the indifference curve and the constraint set. Figure 5a shows that an increase from $\alpha_0 \geq 0.5$ to $\alpha_1 > \alpha_0$ makes the iso-profit curves steeper since there are more z_H types in the population (the solid line curves represent the case with α_1 while the dashed ones correspond to α_0 case). At the same time, an increase in α away from 0.5 also reduces the uncertainty associated with the distribution of applicant types,[16] implying an expansion in the feasible set and allowing firms to pick more informative signals. This is depicted in Figure 5a by an outward shift in the frontier of feasible information strategies from the red-dashed line to the red solid line.[17] At the new optimum, γ_H is higher. Even though γ_L is also higher, in expectation this strategy still leads to higher expected profits since there are more z_H types in the population.

A higher α gives rise to a relatively lower shadow value of information for any given value of a. Consequently, the solid red curve in Figure 5b, corresponding to α_1 lies weakly below the dashed blue curve, corresponding to $\alpha_0 < \alpha_1$. Intuitively, with a higher $\alpha_1 > \alpha_0$ firms are faced with lower initial uncertainty and require less information to distinguish between types.

While in this static limit, the distribution of job-seekers is given exogenously, in the dynamic model that follows, the distribution endogenously evolves over the business cycle. As we show in section 4.2, more indiscriminate rejection of applicants during a recession implies higher productivity applicants are less likely to be filtered out of the pool of job seekers. Higher rejection rates elevate uncertainty over the pool of job-seekers for a sustained period of time, and weigh on firms' hiring activities in the recovery.

2.2 Closing the Model

With the hiring strategy characterized, all that remains is to close the model. This entails specifying how equilibrium meeting rates are determined.

2.2.1 Value of a Firm

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:

$$\mathbf{x}(a, z, e) = F(a, z, e) - b + \beta \mathbb{E}_{a'|a}\Big(1 - d(a', z, e)\Big)\mathbf{x}(a', z, e) \tag{12}$$

[16] Recall that the entropy associated with the prior is given by $-\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha)$ is the largest for $\alpha = 0.5$. Thus, the firm's uncertainty regarding his job applicant is at its maximum whenever $\alpha = 0.5$. For $\alpha \in [0.5, 1]$, an increase in α lowers the initial uncertainty and allows the firm to be less constrained in processing more precise information to lower his posterior uncertainty compared to the case with a lower α.

[17] This outward shift makes information choices which are closer to the unconstrained optimal $gamma_H = 1, \gamma_L = 0$ available to the firm.

We assume that the firm learns the worker's effective productivity perfectly after production. Thus, the firm can choose to fire the worker, $d(a', z, e) = 1$, if it finds the worker to be unsuitable to retain, i.e. $\mathbf{x}(a', z, e) < 0$. Even if the worker is deemed suitable, she may still be separated from the firm at an exogenous rate δ.

2.2.2 Free Entry Condition

A free-entry condition determines the total number of firms that post vacancies in a particular period, pinning down equilibrium market-tightness and 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:[18]

$$\kappa \geq q(\theta) \sum_\tau \mathbb{V}(\sigma, \tau) g_\tau(\tau \mid \sigma) \tag{13}$$

and $\theta = 0$ if (13) holds with a strict inequality. Unlike the standard DMP model, the job-filling rate in our model can be decomposed into two components. The free entry condition pins down the first component - the *contact* rate - $q(\theta) = m(v, l)/v$, which is the rate at which a firm meets a job-seeker. The second component that affects a firm's hiring rate of a worker of duration τ is given by the firm's acceptance rate, $\mathcal{P}(\sigma, \tau)$.[19] Formally, we can now express the aggregate job-filling rate in our model as the product of these two components:

$$\text{Job-filling rate} = q(\theta) \times \mathcal{P}(\sigma)$$

where $\mathcal{P}(\sigma) = \sum_\tau g_\tau(\tau \mid \sigma)\mathcal{P}(\sigma, \tau)$ is the average (across all durations) acceptance probability and corresponds to measured match efficiency in the model. $\mathcal{P}(\sigma)$ forms a wedge between the job-filling rate and the contact rate.[20] Correspondingly, the job-finding rate is given by $p(\theta) \times \mathcal{P}(a)$ where $p(\theta) = m(v, l)/l$.

[18]Since we assume that search is random, the probability that a firm meets a particular type of applicant just depends on the proportion of that type of candidate in the pool of job-seekers.

[19]This is subsumed inside $\mathbb{V}(\sigma, \tau)$ in equation (13).

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

2.3 Composition of job seekers over the business cycle

We are now in a position to define the state variables σ for our economy. At any date t, the economy can be fully described by $\sigma_t = \{a_t, n_{t-1}(z,e), u_{t-1}(z,\tau)\}$ where a_t denotes the prevailing aggregate productivity, $n_{t-1}(z,e)$ 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 z type workers with duration τ at the end of $t-1$. Each firm knows σ_t at the beginning of date t and hence can always compute the distribution of (z,e) across job-seekers of different duration τ.

In equilibrium, the evolution of the mass of job-seekers of duration τ with worker productivity z in period t can be written as:

$$l_t(z,\tau) = \begin{cases} \sum_e d(a,z,e)n_{t-1}(z,e) & \text{if } \tau = 0 \\ u_{t-1}(z,\tau) & \text{if } \tau \geq 1 \end{cases} \tag{14}$$

The first part of equation (14) shows that job-seekers of type z with zero unemployment duration are comprised of workers who were employed at the end of $t-1$ but who were separated from their firms at the beginning of the current period, t. The second line in equation (14) refers to all the z-type unemployed with duration τ at the end of the $t-1$. By construction, all unemployed individuals at the end of a period have duration $\tau \geq 1$. The evolution of the mass of z-type unemployed workers of duration τ is given as:

$$u_t(z,\tau) \;=\; l_t(z,\tau-1)\Big\{1 - p(\theta_t) + p(\theta_t)\sum_e \pi_e(e)\Big(1 - \gamma(z,e \mid \sigma,\tau-1)\Big)\Big\} \tag{15}$$

The first term on the RHS of equation (15) refers to all z-type job-seekers of duration $\tau-1$ at the beginning of t. With probability $1-p(\theta)$, such a job-seeker fails to meet a firm and remains unemployed. With probability $p(\theta)$, the job-seeker meets a firm, draws match productivity e with probability $\pi_e(e)$, but is rejected with probability $1-\gamma(z,e \mid \sigma,\tau-1)$ and remains unemployed. Failure to find a job within period t causes unemployment duration to increase by 1 period from $\tau-1$ to τ. Thus, all $l_t(z,\tau-1)$ job-seekers who fail to find a job within t form the mass of unemployed, $u_t(z,\tau)$, at the end of t.

Similarly, the law of motion for the employed of each (z,e) type match is given as:

$$n_t(z,e) \;=\; [1 - d(a_t,z,e)]\, n_{t-1}(z,e) + p(\theta_t)\pi_e \sum_{\tau=0}^{\infty} \gamma(z,e \mid \sigma,\tau) l_t(z,\tau) \tag{16}$$

Equation (16) shows that the mass of employed workers with effective productivity ze in period t is composed of two terms. The first term denotes the fraction of employed workers at the end of $t-1$, with effective productivity ze, who are not separated from the firm at the beginning of t. The second term refers to all job-seekers at date t who meet a firm with probability $p(\theta)$, draw match specific e with probability $\pi_e(e)$ and who are hired by a firm after the interview. For a type z applicant with τ unemployment duration, the latter occurs 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:

$$\sum_{\tau} u_t(z, \tau) + \sum_{e} n_t(z, e) = \pi_z(z) \ , \forall z \in \mathcal{Z}$$

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.

$$\begin{aligned} l_t(\tau) &= \sum_{z} l_t(z, \tau) \\ l_t &= \sum_{\tau} l_t(\tau) \end{aligned}$$

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

$$g_z(z \mid \sigma, \tau) = \frac{g_{z,\tau}(z, \tau \mid \sigma)}{g_\tau(\tau \mid \sigma)} \equiv \frac{l_t(z,\tau)/l_t}{l_t(\tau)/l_t} = \frac{l_t(z, \tau)}{l_t(\tau)} \ , \forall \tau \geq 0 \tag{17}$$

where $g_z(z \mid \sigma, \tau)$ is defined simply as 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.

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

Our assumption that search is random implies that the probability that a firm meets a particular type of applicant is the same as the proportion of that type of applicant in the pool of job seekers. Consequently, a firm's prior about any workers type (z, e) after observing their unemployment duration τ and σ is simply given by the joint distribution $\mathcal{G}(z, e \mid \sigma, \tau)$. 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. 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%. 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 probabilities are bounded between 0 and 1. We set ι to 0.5 as standard in the literature.[21] We assume that (log) aggregate productivity follows an AR(1) process:[22] $\ln a_t = \rho_a \ln a_{t-1} + \sigma_a \varepsilon_t$ where $\varepsilon_t \sim N(0,1)$. We set the persistence $\rho_a = 0.9$. We set the standard deviation $\sigma_a = 0.0165$ as in Shimer (2005).

The remaining parameters are chosen to minimize the distance between a set of moments from model 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%. This is the average exit probability in the data over the period of 1950-2016 implying that the average tenure of a worker lasts roughly 2.8 years.[23] 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), 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. $z \sim Beta(A_z, B_z) + 0.5$ while the match quality shock is drawn from the Beta distribution $e \sim Beta(A_e, B_e)$.[24]

Since the vacancy posting cost, κ, the capacity processing constraint, χ 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 5.8%, which is the average unemployment rate in the data over the period 1950-2016. We use data on unemployment duration and unemployment-to-employment transitions (UE) from the Current Population Survey (CPS) to calculate the relative job-finding rates in the data. As in Kroft et al. (2016), we estimate a weighted non-linear least squares regression on the relative job-finding rate against unemployment duration of the following form: $\frac{\mathrm{UE}(\tau)}{\mathrm{UE}(1)} = \pi_1 + (1-\pi_1)exp(-\pi_2\tau)$ where τ is the unemployment duration, and $\frac{\mathrm{UE}(\tau)}{\mathrm{UE}(1)}$ is the average job-finding rate of an unemployed individual of duration τ relative to an unemployed individual with 1 month of unemployment duration. We cluster all those who are more than 9 months unemployed into a single bin. We target the fitted relative job-finding rates from this regression.

Heterogeneity in both individual fixed productivity z and match-specific productivity e are crucial for replicating the decline in relative job-finding rates as in the data. If the only form of heterogeneity across matches arose from match-specific productivity, relative job-finding rates across duration would be flat as draws of e are i.i.d and independent

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

[22] 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.

[23] We calculate the exit probabilities as in Shimer (2012).

[24] Specifically we set the number of worker productivity types to be $n_z = 7$ and the number of match-specific shocks to $n_e = 15$. See Appendix D.1 for details on the construction of $\mathcal{Z}$ and $\mathcal{E}$.

of past matches. As such, unemployment duration would provide no useful information about the applicant's type since her productivity is a random draw at the time of each new meeting. If instead the only form of heterogeneity stemmed from workers' fixed productivity types z, then relative job-finding rates would be strictly declining in duration and would not exhibit any flattening out. Longer spells of unemployment suggest a higher number of rejections and signal that the applicant is of a low z type. Heterogeneity in e alongside z is required to generate the convex shape in relative job-finding rates as in the data. We use this feature in relative job-finding rates to discipline our choice of parameters regarding the heterogeneity in z and e.

In summary, we have 8 parameters to estimate $\{\chi, \kappa, \delta, b, A_z, B_z, A_e, B_e\}$ and we target 8 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 greater than one month. Tables 1 summarizes both the fixed and inferred parameters.

Table 1: Model Parameters

Fixed Parameters

Parameter	Description	Value	Source
β	discount factor	0.9967	annualized real return = 4%
σ_a	std. dev. of a	0.0165	Shimer (2005)
ρ_a	autocorr. of a	0.9	Shimer (2005)
ι	matching func. elasticity	0.5	Menzio and Shi (2011)

Inferred Parameters

Parameter	Description	Value
b	home production	0.2581
δ	exog. separation rate	0.0078
κ	vacancy posting cost	0.0001
χ	capacity processing constraint (in nats)	0.1118
A_z	shape parameter - worker ability	1.3451
B_z	shape parameter - worker ability	8.5290
A_e	shape parameter - match productivity	1.9975
B_e	shape parameter - match productivity	7.3234

The model is able to match the moments in data. Under the parametrization, home production is 68.3% of average output (the target was 70% of average output). The model generates an average EU rate of 1.8% and an average unemployment rate of 5.9%. 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. The drop-off in the model's implied relative job-finding rates for individuals unemployed for nine months or more reflects the fact that we have clustered these individuals with

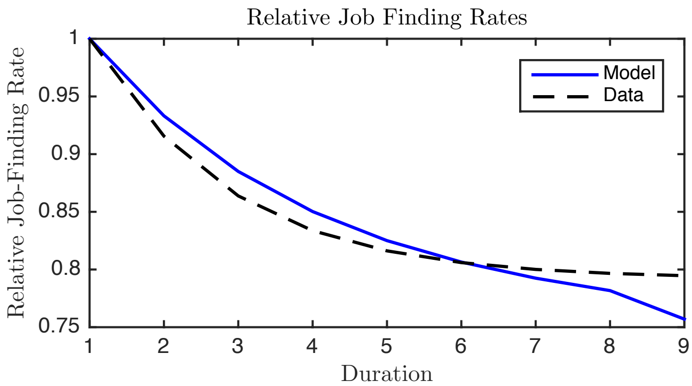


Figure 6: Relative Job finding rates by duration of unemployment

unemployment duration greater than equals to nine months into one bin.

The shadow cost of processing information In steady state, the average shadow value of information, $\lambda = \sum_{\tau} g_{\tau}(\tau)\lambda(\tau) = 0.42$. The shadow value can be interpreted in terms of how much output a firm is willing to give up for one more unit of information. Survey evidence on turnover and recruitment costs suggests that the cost of hiring is not trivial. Using data from the California Establishment Survey, Dube et al. (2010) report that the average cost per recruit is about 8% of annual wages while Hamermesh (1993), using a 1979 national survey, suggests that depending on the occupation, hiring costs range from \$680 to \$2200 dollars. In our model, the average λ is around 13% of annual wage income. Thus, our value for λ suggests that the implicit cost of processing information about applicants is sizable and is in the order of magnitude as measured hiring costs.

4 Results

4.1 What happens in recessions?

As our first exercise, we simulate a moderate 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 (RI) to responses of the full information model (FI). All results are presented in terms of log deviations. For the FI economy, we assume that firms have no constraints on the amounts of information they can process ($\chi \to \infty$) and are able to determine an applicant's type with certainty.

The left panel of Figure 7 plots the path of aggregate productivity in our experiment while the right panel shows how the firing rate responds to the fall in productivity. We measure the firing rate in period t as the fraction of employed individuals from $t-1$ who are newly separated at the start of t. On impact, the moderate decline in a does not

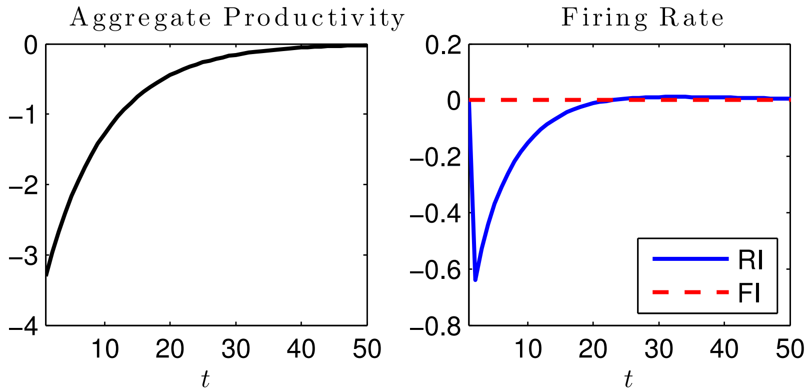


Figure 7: Moderate Decline in Aggregate Productivity

induce any change in the firing rate in both the FI and RI models, implying that the set of workers who were already employed before the shock are still profitable to retain in the current period.[25] Consequently, any impact to unemployment rates comes from the hiring margin and not from higher separation rates. In fact, in the RI model, the firing rate declines in subsequent periods. As we discuss next, this decline in firing rates is explained by more selective hiring by RI firms.

A moderate decline in a increases the loss a firm incurs from hiring an unsuitable worker. As such, firms desire more information to avoid hiring the wrong worker. However, limited processing capacity prevents firms from acquiring more informative signals about the applicant, raising the average shadow value of information λ. The top left panel of Figure 8 shows that λ rises by about 2.5% on impact. It should be noted that since firing rates did not change on impact, the composition of the unemployment pool remains the same on impact and the initial rise in λ is solely due to changes in a. Similar to the forces we highlighted in our static example in Section 2.1.3, a fall in a without any change in the composition of job-seekers when firms are information-constrained translates into lower acceptance rates as firms err on the side of caution and reject more often to avoid hiring the wrong worker. Since RI firms reject more applicants, the incidence of hiring unsuitable workers and firing them ex-post also decreases. In contrast, the FI model witnesses no decline in firing rates since FI firms never hire the wrong worker and thus never need to fire them.

The initial fall in a drives movements in unemployment, job-finding rates and match efficiency in the RI model but generates negligible response in the FI model. The second cell in the top row of Figure 8 shows that the average acceptance rate (which is also measured match efficiency), $\mathcal{P}(\sigma)$, falls by close to 1% in the RI model. In contrast, Figure 8 shows that changes in measured match efficiency are zero on impact and negligible

[25]This result is due to the discreteness of types in our model. For larger shocks, the firing margin is triggered and the distribution of types in the unemployment pool changes on impact. We illustrate this outcome in Section 4.2.

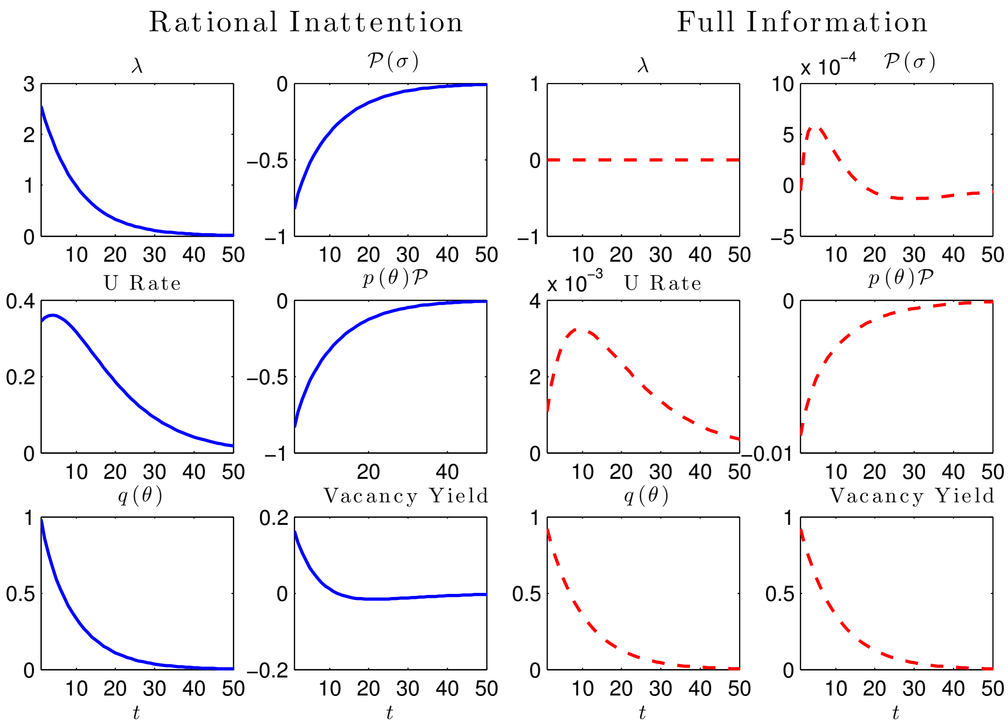


Figure 8: Response of Key Labor Market Variables in RI vs. FI models

thereafter in the FI model.[26] Because a moderate decline in a did not change the firm's willingness to hire the same set of workers as they did in periods prior to the shock, the composition of unemployed job-seekers did not observe any change on impact. Given negligible change in the pool of unemployed, the FI firm's acceptance decisions is relatively unaffected in a moderate downturn as it can always observe and screen out the worker it wants.

The shock's impact on match efficiency in turn affects unemployment and job-finding rates in the economy. The first panel on the second row of Figure 8 shows that in the RI model, the unemployment rate rises by close to 0.35% while the job-finding rate, measured as $p(\theta)\mathcal{P}(\sigma)$, falls by close to 1% (second panel of the second row of Figure 8). In contrast, the unemployment rate rises by a mere 0.003% in the FI model. The differential response of match efficiency with respect to a moderate downturn is the primary factor behind the different unemployment experiences between the two models. While these numbers may seem small, this is because our model suffers from the usual Shimer puzzle, i.e. that the labor search model is unable to generate enough volatility to match the fluctuations in the data. Importantly, the FI model still fares much worse in predicting the response of unemployment to moderate downturns.

[26] Match efficiency in the FI model actually observes a slight uptick in subsequent periods following the initial adverse shock. This stems from the fact that lower vacancy creation during the recession reduces the outflow from unemployment and leaves the average quality of the job-seeker pool marginally higher than in a boom. Higher average quality of the unemployed causes FI firms to accept more often on average. As such, match efficiency in the FI model rises by 0.0005%.

Even though the moderate decline in a was not large enough to change firms' retention decisions, the decline nonetheless reduces the value of creating a vacancy. Consequently, fewer firms enter the labor market, causing $q(\theta)$ to rise as in the bottom first and third cell in Figure 8. Importantly, the bottom rightmost panel of Figure 8 reveals that vacancy yield in the FI model exactly mirrors the rise in $q(\theta)$. In contrast, the RI model predicts a muted increase in vacancy yields in response to a shock. In the RI model, the decline in average acceptance rates, $\mathcal{P}$ mitigates the rise in $q(\theta)$, lowering the number of hires even though firms meet workers at higher rates. These differential responses in vacancy yield in the RI and FI model are reminiscent of Figure 1 where the standard full information labor search model with negligible changes in match efficiency predicts a much higher vacancy yield than its empirical counterpart.

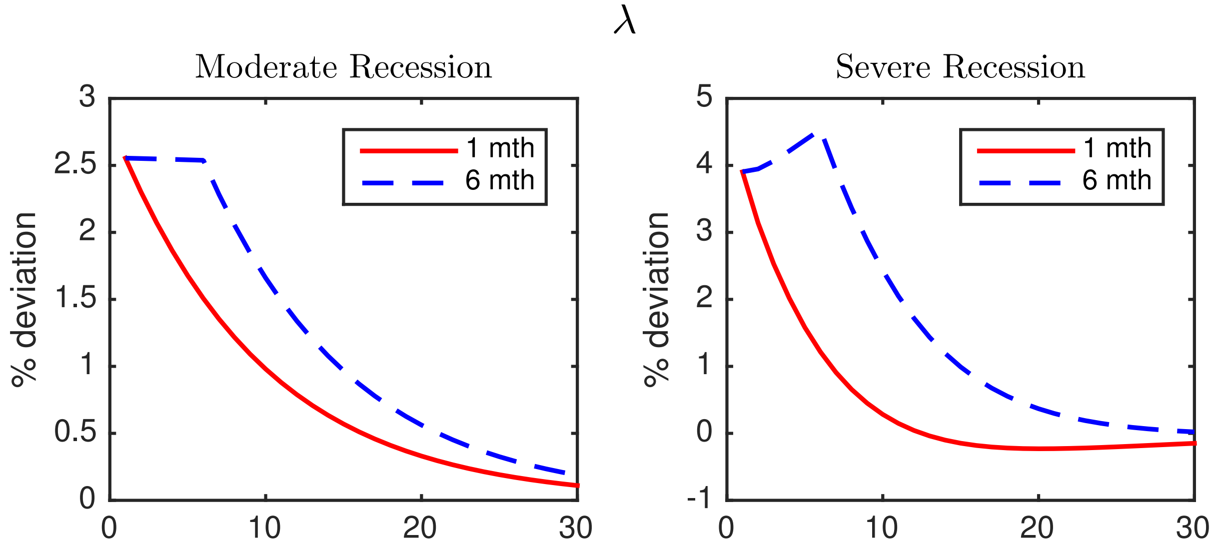


Figure 9: Decomposing Changes in λ

In our model, there are two forces that affect λ, the shadow value of information, which in turn affects the firm's ability to distinguish workers and its hiring probabilities. Both a and the distribution of job-seekers affect the extent to which a firm is informationally constrained and hence affects the response in λ. To show which force drives the increase in λ during a moderate downturn, we conduct a separate exercise where aggregate productivity a falls and remains about 3% (two standard deviations) below steady state for 6 consecutive periods. In this experiment, while a is fixed over these 6 periods, the composition of the unemployed continues to change and hence any further change on λ stems only from changes in the distribution of job-seekers.

The blue dashed line in the left panel of Figure 9 depicts the response of λ when a is kept fixed at about 3% below steady state for 6 consecutive periods and then recovers while the red solid line is the same as in the left topmost panel of Figure 8 . Because unemployment rates in the RI model rose by only about 0.52% when a fell and remained about 3% below steady state for 6 periods, the composition of unemployed job-seekers did not change significantly. Initial uncertainty (measured by entropy) grows a cumulative 0.03% by the sixth period. As such, our calibrated model suggests that for a moderate recession, most of the impact on λ stems from the changing level of aggregate productivity, a, as opposed to changes in the distribution. This can be seen by the fact that λ does not increase in the 6 periods where a is held fixed.

4.2 What Happens in Severe Recessions?

To simulate a severe recession, we subject the economy with a ten percent fall in aggregate productivity. Figure 10 depicts the response of the FI and RI models to this shock. Unlike the case of a moderate decline in a, the top leftmost of Figure 10 shows that in response to a severe negative shock, firing rates spike on impact in both the RI and FI models, implying that both the composition of the unemployed as well as aggregate productivity changes on impact in a severe recession. The percentage increase in firing rates in the FI model is larger than in the RI model since the FI model only witnesses exogenous separations in steady state as firms can perfectly observe the applicant's type and never hire a worker who is unsuitable for production. In contrast, in the RI model, both exogenous and endogenous separations occur since firms in steady state can still make mistakes and hire the wrong workers.

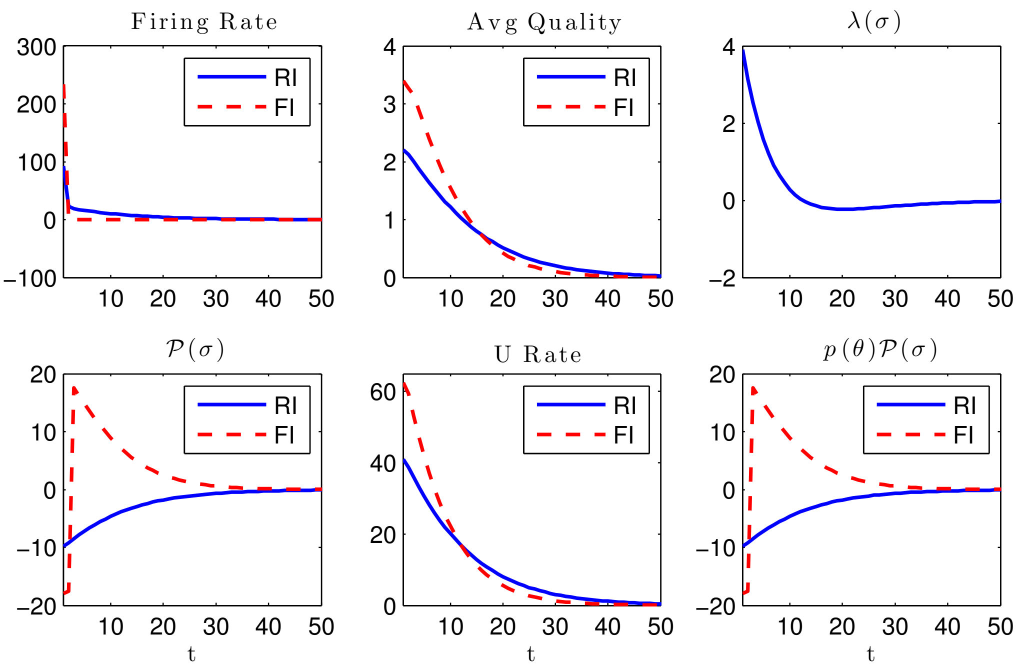


Figure 10: Response to a Severe Recession

The higher firing rates triggered in a severe recession have the effect of raising the average quality of unemployed job-seekers in both the FI and RI model. This is in line with the findings of Mueller (2015) who shows that the composition of the unemployed shifts towards high quality types during a recession. The higher percentage increase in the FI model relative to the RI model can be explained by the difference in the steady state average quality in both models. Firms in the FI model correctly identify and select the high (z, e) types for production, leaving behind a pool of unemployed job-seekers of lower average quality. In contrast, in the RI model, firms sometimes mistakenly reject high (z, e) applicants, causing the pool of unemployed to be of higher average quality in steady state than in the FI model.

A large decline in a magnifies the loss from hiring the wrong worker to RI firms. The possibility of more severe losses prompt RI firms to require even more information about the applicant to guide their hiring decisions. The top rightmost panel of Figure 10 shows that on impact, the average shadow value of information, λ, rises by 4%, implying that firms desire more precise information about an applicant but are unable to acquire this information. As per equations (6) and (7), the combination of a lower payoff from hiring, i.e., lower $\mathbf{x}(a, z, e)$ for all (z, e) types, together with a rise in λ implies that firms optimally reject workers more often despite the fact that the average quality of the unemployed job-seekers has improved. Consequently, match efficiency in the RI model falls by close to 10% as is depicted in the bottom leftmost panel of Figure 10.

Match efficiency, however, behaves very differently in the FI model. On impact, match efficiency in the FI model falls by close to 20% relative to steady state. However, in subsequent periods, match efficiency recovers and in fact rises 20% above its steady state level. This rapid reversal in the decline in match efficiency stems from the counteracting effects of an improved average quality of the unemployment pool. Since firms raise their retention standards on impact of the shock, newly separated workers at the start of $t = 1$ tend to be comprised of higher $z-$types relative to job-seekers with duration $\tau \geq 1$. As the economy recovers, the higher average quality of the unemployed of duration zero who were not re-employed at the end of $t = 1$ mitigates the lower aggregate productivity. The combination of recovering aggregate productivity after the initial shock, and higher average quality of job-seekers lead FI firms to raise their average acceptance probabilities. Thus, average job finding rates and measured match efficiency rebound and rise above their steady state level in the FI model.

The key to understanding the differential response between the two models is the fact that FI firms can costlessly observe the applicant's true (z, e) while RI firms cannot. In other words, λ is always equal to zero for FI firms, while λ increases precisely during downturns when RI firms desire more information. As aggregate productivity recovers, FI firms take advantage of the improving average quality in the pool of unemployed job-seekers and increase their acceptance rates, causing match efficiency to rapidly rebound and even overshoot its steady state level for an extended period. In contrast, the rise in λ together with a fall in a implies that RI firms are reluctant to hire in spite of improvements in the average quality of the unemployed, since both the cost of hiring the wrong worker and the extent to which the firm is informationally constrained have increased. In addition, higher rejection rates for all types of job-seekers raise the amount of uncertainty in the pool of unemployed going forward, further hampering the firm's ability to distinguish applicants and keeping acceptance rates depressed. Both these forces counteract the effect of an increase in average quality on measured match efficiency in the RI model, keeping it persistently below its steady state level for 50 months following the shock.

Unlike a moderate recession case, the fall in hiring rates has tangible effects on the composition of job-seekers going forward, which in turn impacts the extent to which a firm is informationally constrained. To highlight how uncertainty can affect changes in λ during a severe downturn, we conduct a similar experiment as before, where we fix a to be lower for 6 consecutive periods. The blue dashed line in the right panel of Figure 9 shows λ in response to a falling and remaining 10% below steady state for 6 periods before recovering while the red solid line is the same graph as the top rightmost panel in Figure 10. Because a is held fixed for the first 6 periods, the additional increase in λ after the first period can only be attributed to changes in the distribution of job-seekers. Unlike the moderate recession, initial uncertainty grows a cumulative 4% by the sixth period. The rise in uncertainty causes λ to increase an additional 0.5 percentage points by the sixth month, which is akin to 12.5% increase from its initial rise of 4% in period 1.

Overall, the impulse response functions show that the RI model generates propagation properties that better fit the data than the FI model. In response to smaller shocks, the RI model observes a rise in unemployment rates and a decline in match efficiency while the FI model observes negligible change. In response to large shocks, the improvement in the average quality of the unemployed mitigates the decline in aggregate productivity in the FI model and predicts a rise in match efficiency after the third period. In contrast, the RI model observes persistently low match efficiency. We now examine how well simulated match efficiency from the rational inattention model fits actual match efficiency in the data.

4.3 Model vs. data - match efficiency

Having described how the model responds in both moderate and severe recessions, we now assess how well match efficiency in the RI model compares to actual match efficiency in the data. Using the same matching function as in our model, we compute (log) match efficiency ξ_t as:

$$\ln \xi_t = \ln m_t - \ln \left(\frac{u_t v_t}{(u_t^{\iota} + v_t^{\iota})^{1/\iota}} \right)$$

where 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. As per our calibration, we set $\iota = 0.5$. We detrend the resulting series for match efficiency using a cubic time trend.

To assess our model's ability to replicate match efficiency as in the data, we use information on unemployment rates for the coverage period of JOLTS, i.e. 2000m12 - 2016m12. We use the unemployment rate as the observable and assume that $u_t^{data} = u_t^{model} + e_t$ where we treat $e_t \sim N(0, \sigma_e^2)$ as measurement error. Conditioning on the unemployment rate, we use the particle filter to filter out the sequence of productivity shocks which would generate the sequence of observed unemployment rates in the model as in the data.[27] Next, we feed the filtered series of productivity shocks into the calibrated model and compute the implied measured-match efficiency in the model. We perform this exercise with both the RI and FI models.[28]

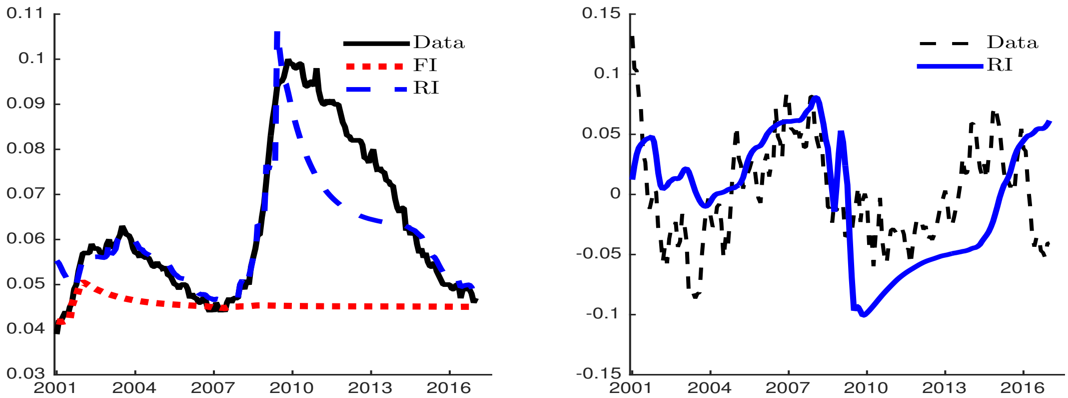


Figure 11: Data vs. Model: Unemployment Rates and Match Efficiency

The left panel of Figure 11 demonstrates the models' ability to match unemployment rates as in the data. Noticeably, the FI model struggles to match unemployment rates as in the data and the particle filter attributes most of the deviations from the mean unemployment rate to measurement error e_t.[29] This is perhaps not surprising since the FI model displayed very little response in unemployment with respect to moderate shocks. Furthermore, in response to severe shocks, the sharp reversal in match efficiency in the FI model ensured that spikes in unemployment rates decayed at a rapid rate following the initial shock. In contrast, the RI model provides a much better fit for the unemployment rates in the data, reaffirming our earlier findings which showed that the impulse response of unemployment in the RI model features more amplification and persistence.

With the filtered productivity series, the RI model is also able to replicate match

[27]The implementation of the particle filter follows Fernandez-Villaverde et al. (2016).

[28]In order to keep the comparison fair, we calibrate the FI model such that the moments from the FI model matches the same targets we used to calibrate the RI model. For the calibrated parameters used in the full information model, please see the appendix E.

[29]Figure 13 in Appendix E plots the measurement error component for the FI and RI model.

efficiency in the data fairly well and matches the turning points in the data. Note that this provides some external validation for the model since we only used information on unemployment to infer the implied shocks, and did not use any direct information on match efficiency. The right panel of Figure 11 shows the model's ability to replicate match efficiency in the data. Through the lens of the FI model, the particle filter attributes a large fraction of the changes in observed unemployment to measurement error. Thus, the filtered changes in productivity and hence, match efficiency are negligible.[30]

In addition to the exercise above, using the Kolmogorov-Smirnov test, we test the hypothesis that the measured-match efficiency in each model and the measured-match efficiency in data are drawn from the same distribution. We cannot reject this hypothesis for the RI model but can reject it for the FI model.[31] Hence, we find support for the claim that the model with rational inattention is a better fit for explaining match efficiency movements in the data.

4.4 Duration of unemployment as signal of quality

Research on resume audit studies has found evidence of firms using unemployment duration to filter applicants. Our model suggests that duration as an indicator of worker quality is a less informative signal during recessions. As recessions are periods where the value of the firm declines, firms post fewer vacancies and the rate at which a worker contacts a job declines. Alongside the lower meeting rates, firms in the RI model also reject applicants more often to avoid hiring the wrong worker. These lower acceptance rates compound the lower meeting rates of job-seekers and raise the likelihood of a longer unemployment spell for all job-seekers. As such, duration provides less information about an applicant's true ability during a recession.

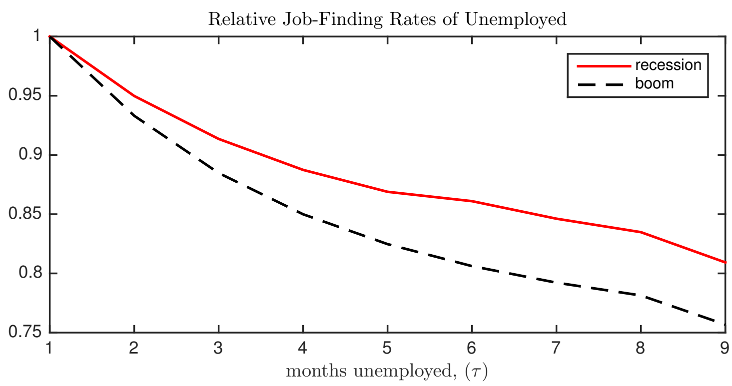


Figure 12: Relative Job Finding Rate of Job-seekers by Duration τ.

Figure 12 shows the relative job-finding rates of unemployed job-seekers relative to a job-seeker with 1 month unemployment duration across booms and severe recessions.

[30]Figure 14 in Appendix E compares measured-match efficient in the FI model and the data.

[31]The p-value for the rational inattention model is 0.4255 implying that the null hypothesis cannot be rejected. In contrast, the p-value for the full information model is given by 3.9e-14.

In booms, the job-finding rate of job-seekers 6 months unemployed is about 20% lower than that of a job-seeker with an unemployment spell of 1 month. In a severe recession, however, the job-finding rate of an individual with 6 months unemployment duration is about 15% lower than that of an individual with one month unemployment duration. These results are consistent with the findings of Kroft et al. (2013) who find that there is a gentler decline in callback rates in areas with depressed economic activity.

5 Discussion

5.1 Alternative cost structures of processing information

A natural question that arises is why modeling information costs in terms of entropy is better suited to address the question at hand. One alternative popular specification commonly used is that of a fixed cost of acquiring information. Under this alternative specification, once firms pay a fixed cost, they learn the type of a worker perfectly.[32] In such a setting, the acceptance rate of firms would then be similar to that generated by the full information model. As firms value information more in a recession, they are willing to pay the fixed cost during downturns, learning perfectly the applicant's type before making the hiring decision. As such, firms always perfectly screen out the correct candidates for production and match efficiency behaves as in the full information model. In fact, fixed costs of acquiring information, when activated, present an additional cost of creating a vacancy and merely act toward further depressing vacancy posting and raising the contact rate of firms during recessions. Thus, much like the full-information model, such a cost structure would be unable to explain the large changes in match efficiency.[33]

Alternatively, one could have modeled the problem of the firm as one of noisy information where firms receive a noisy signal about the object they wish to learn about. In terms of our model, one could think of firms receiving a signal (upon meeting an applicant) of the form: $\mathbf{s} = ze + \eta$, where $\eta \sim N\left(0, \sigma_\eta^2\right)$ is Gaussian noise. It is important to realize that in our characterization of the firm's optimal information strategy in terms of acceptance probabilities, we did not restrict the firm from getting signals of this form. As such, if firms found such signals to be optimal, they would choose it and that would not affect our characterization of the optimal solution in terms of choice probabilities. However, in general a firm would never want to choose a signal of this form in our setting (even though it could) since this signal is very expensive in terms of entropy and is

[32] This specification allows for interviews to either be fully informative or not informative at all.

[33] Even staying within the family of rational inattention models, there is an alternative specification under which firms must pay a physical cost which depends on the reduction in uncertainty measured in terms of entropy (See for example Paciello and Wiederholt (2014) and Stevens (2015)). As aforementioned, an earlier version of this paper featured this specification and the results are qualitatively the same. These results are available upon request.

not optimal. Signals of this form are only optimal if agents have a quadratic objective which is not the case in our model.[34] Restricting the information structure to be normal imposes additional costs on firms' processing capacity and prevents firms from designing more cost-effective information strategies.

5.2 Wage determination

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 incentivize applicants to reveal their true types, thus, circumventing the constraints posed by the firm's limited information processing capacity. While the use of contracts to separate different ability workers does potentially help overcome the limited processing capacity of the firm, it requires the firm to give up informational rents in order to incentivize applicant to reveal their types truthfully. Thus, depending on how constrained firms are in terms of channel capacity, they may or may not choose to use contracts. 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 shadow cost of information, both of which are changing over the business cycle. We leave this for future research.

6 Conclusion

We present a novel channel through which firms' hiring standards affect fluctuations in measured-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. Inability to get more precise information leads firms to err on the side of caution and reject more job-seekers. These lower acceptance rates are reflected in lower measured-match efficiency. Further, lack of adequate screening by firms in filtering out suitable applicants today leads to elevated uncertainty in the pool of unemployed in the future, thus, providing an additional propagation mechanism for initial shocks to affect and amplify unemployment rates in the future. Overall, our mechanism offers insight on how constraints on information processing can cause hiring rates to stall and vacancy yields to falter despite the large number of job-seekers available for each vacancy.

[34] Sims (2006) shows that Gaussian posterior uncertainty is optimal only with quadratic loss functions.

References

Bartos, Vojtech, Michal Bauer, Julie Chytilova, and Filip Matejka, "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review*, June 2016, *106* (6), 1437–1475.

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.

Dube, Arindrajit, Eric Freeman, and Michael Reich, "Employee replacement costs," *Institute for Research on Labor and Employment*, 2010.

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.

Fernandez-Villaverde, Jesus, Juan F. Rubio Ramirez, and Frank Schorfheide, "Solution and Estimation Methods for DSGE Models," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, 1 ed., Vol. 2a of *Handbook of Macroeconomics*, Elsevier, 2016, chapter 10.

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.

_ **and Sam Schulhofer-Wohl**, "Measuring Job-Finding Rates and Matching Efficiency with Heterogeneous Jobseekers," NBER Working Papers 20939, National Bureau of Economic Research, Inc February 2015.

Hamermesh, Daniel, *Labor Demand*, Princeton University Press, 1993.

Hornstein, Andreas and Marianna Kudlyak, "Estimating Matching Efficiency with Variable Search Effort," Working Paper Series 2016-24, Federal Reserve Bank of San Francisco October 2016.

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

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.

Paciello, Luigi and Mirko Wiederholt, "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *Review of Economic Studies*, 2014, *81* (1), 356–388.

Petrongolo, Barbara and Christopher A. Pissarides, "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, June 2001, *39* (2), 390–431.

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.

_ , "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*, April 2012, *15* (2), 127–148.

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

_ , "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review*, May 2006, *96* (2), 158–163.

Stevens, Luminita, "Coarse Pricing Policies," Staff Report 520, Federal Reserve Bank of Minneapolis November 2015.

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

Appendix

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

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

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 $\gamma(z,e\mid\sigma)$:

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

with complementary slackness conditions $\mu(z,e\mid\sigma)\left[1-\gamma(z,e\mid\sigma)\right]=0$ and $\zeta(z,e\mid\sigma)\gamma(z,e\mid\sigma)=0$. Thus, for $0<\gamma(z,e\mid\sigma)<1$, 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:

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

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

$$1=\sum_z\sum_e\frac{e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma)}}}{1-\mathcal{P}(\sigma)\left[1-e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma)}}\right]}g(z,e\mid\sigma) \tag{19}$$

B Proof of Lemma 2

For $\chi>H(G)$, the shadow value of an additional nat, $\lambda=0$. Evaluating (6) for (z,e) combinations such that $\mathbf{x}(a,z,e)<0$ in the limit as $\lambda\to 0$ yields:

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

Next, consider an applicant (a,z) such that $\mathbf{x}(a,z,e)\geq 0$. Under the optimal information strategy, this applicant is hired with probability 1:

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

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

C Proof of Lemma 3

Let $\gamma_H(a)$ and $\gamma_L(a)$ denote the optimal choices when aggregate productivity is given by a. Differentiating (10) with respect to aggregate productivity a yields:[35]

$$[\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_H)]\, \alpha\gamma_H'(a) + [\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_H)]\,(1-\alpha)\,\gamma_H'(a) = 0 \tag{20}$$

where $\gamma_i'(a) = {}^{\partial\gamma_i(a)}/_{\partial a}$ for $i \in \{H, L\}$. We also know that the optimal choices of $\gamma_H(a)$ and $\gamma_L(a)$ satisfy the first order conditions:

$$\mathbf{x}(a, z_i) = \lambda(a)\,[\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_i(a))]\ \ ,\, i \in \{H, L\} \tag{21}$$

where $\lambda(a)$ is the multiplier on the constraint (10). Using this, rewrite equation (20) as:

$$\frac{\gamma_H'(a)}{\gamma_L'(a)} = -\frac{(1-\alpha)\,\mathbf{x}(z_L, a)}{\alpha\mathbf{x}(z_H, a)} \tag{22}$$

First, consider $a \in [{}^{b}/_{z_H}, {}^{b}/_{z_L}]$. In this range, $\mathbf{x}(a, z_H) \geq 0$ and $\mathbf{x}(a, z_L) \leq 0$. Thus, (22) implies that $\gamma_H'(a)$ and $\gamma_L'(a)$ are the same sign. It remains to show that the sign is positive. To see this, recall that that the optimal choices are characterized by equation (11). For $a \in [{}^{b}/_{z_H}, {}^{b}/_{z_L}]$, the LHS of (11) is a negative number. Also, when a goes up marginally, the LHS becomes a larger negative number. Suppose $\gamma_H(a)$ and $\gamma_L(a)$ were decreasing in a. Then the RHS must be a smaller negative number since the feasible set of choices is a convex set following from the properties of entropy which implies a contradiction. Thus, $\gamma_H(a)$ and $\gamma_L(a)$ are increasing for a in this interval.

Now consider the range $a < {}^{b}/_{z_H}$. In this range, $\mathbf{x}(a, z_L) < \mathbf{x}(a, z_H) < 0$ and thus, the firm does not hire any applicants. In this range of aggregate productivity, $\gamma_H(a) = \gamma_L(a) = 0$ and thus is constant in a. Similarly, if the level of aggregate productivity is very high, $a > {}^{b}/_{z_L}$, $\mathbf{x}(a, z_H) > \mathbf{x}(a, z_L) > 0$ and the firm is willing to hire both types: $\gamma_H(a) = \gamma_L(a) = 1$. Since the unconditional probability of accepting an applicant $\mathcal{P}(a) = \alpha\gamma_H(a) + (1-\alpha)\gamma_L(a)$, it is also weakly increasing in a.

[35] This condition holds as long as the constraint holds with an equality at the optimum which is always the case since we assumed that $\chi < \overline{\chi}$.

Online Appendix (not for publication)

D Numerical Implementation

D.1 Constructing the sets $\mathcal{Z}$ and $\mathcal{E}$

This Appendix details how we construct the discrete sets $\mathcal{Z}$ and $\mathcal{E}$ from the continuous $Beta(A, B)$ distributions used to calibrate the model. We only describe the process for constructing $\mathcal{Z}$ as the same procedure is used to construct $\mathcal{E}$. In what follows, we denote the cumulative distribution of the $Beta(A, B)$ distribution by $\mathcal{F}(\cdot; A, B)$.

1. Construct a sequence $\{p_i\}_{i=1}^{n_z+1}$ where each $p_1 = 0$, $p_{n_z+1} = 1$ and for $i \in \{2, \cdots, n_z\}$ p_i is given by:
$$p_i = i/n_z$$
where n_z denotes the cardinality of the set $\mathcal{Z}$.

2. Using the sequence $\{p_i\}$ construct a sequence of intervals denoted $\{m_i\}_{i=1}^{n_z+1}$ such that
$$m_i = \mathcal{F}^{-1}(p_i; A, B)$$
where $\mathcal{F}^{-1}(\cdot; A, B)$ denotes the inverse cdf.

3. Next construct the sequence $\{z_i\}_{i=1}^{n_z}$ as:
$$z_i = \frac{m_i + m_{i+1}}{2} \text{ , for } i \in \{1, \cdots, n_z\}$$

The set $\mathcal{Z}$ is just defined as $\{z_1, \cdots, z_{n_z}\}$ and the probability mass associated with each z_i is given by $1/n_z$.

D.2 Solving the model

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

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

and the evolution of the mass of unemployed workers of type z and unemployment duartion τ can be written as follows. For unemployment durations $1 \leq \tau < \bar{\tau}$, we can write the transition equation as:

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

while for $\tau \geq \overline{\tau}$ we can write it as:

$$\begin{aligned} u_t(z,\tau \geq \overline{\tau}) &= l_t(z,\bar{\tau}-1)\Big\{1 - p\left(\theta[\sigma_t]\right) + p\left(\theta[\sigma_t]\right)\sum_e \pi_e(e)(1-\gamma\left[z,e \mid \sigma_t, \bar{\tau}-1\right])\Big\} + \\ & \quad l_t(z,\bar{\tau})\Big\{1 - p\left(\theta[\sigma_t]\right) + p\left(\theta[\sigma_t]\right)\sum_e \pi_e(e)(1-\gamma\left[z,e \mid \sigma_t, \bar{\tau}\right])\Big\} \end{aligned}$$

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

E Particle Filter

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. We keep fixed the parameters governing the heterogeneity of workers and match specific productivity as in the rational inattention model. This implies that the unconditional distribution of individuals have the same effective productivity, ze, as in the rational inattention model. In additional, the full information model sets χ to infinity, i.e. there is no fixed capacity processing constraint. Table 2 details used in the full information model.

Table 2: Model Parameters for Full Information

Inferred Parameters		
Parameter	Description	Value
b	home production	0.2894
δ	exog. separation rate	0.0029
κ	vacancy posting cost	0.0281
A_z	shape parameter - worker ability	1.3451
B_z	shape parameter - worker ability	8.5290
A_e	shape parameter - match productivity	1.9975
B_e	shape parameter - match productivity	7.3234

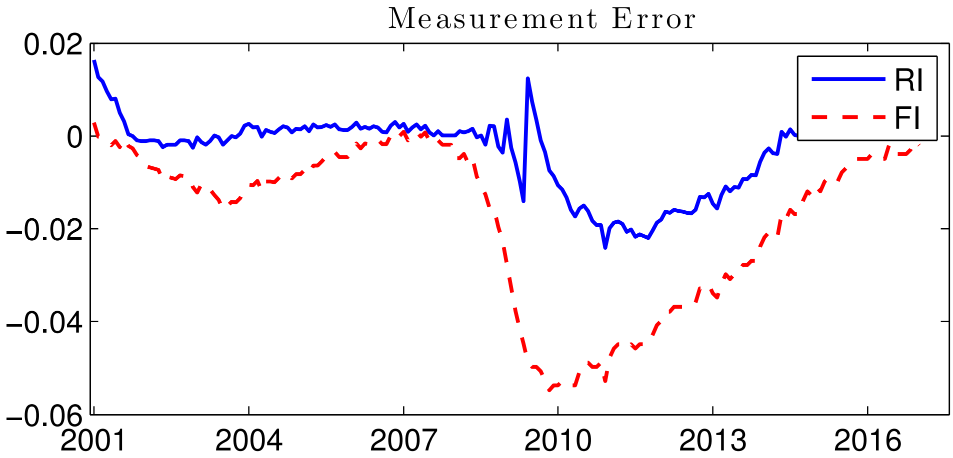


Figure 13: Data vs. Full Information Match Efficiency

Filtered Match Efficiency in the FI model

Since the particle filter attributes most of the variation in the observed unemployment rate to measurement error, the resulting fluctuations in productivity are small. Consequently, in the FI model, match efficiency barely changes during the Great Recession and the subsequent period of elevated unemployment.

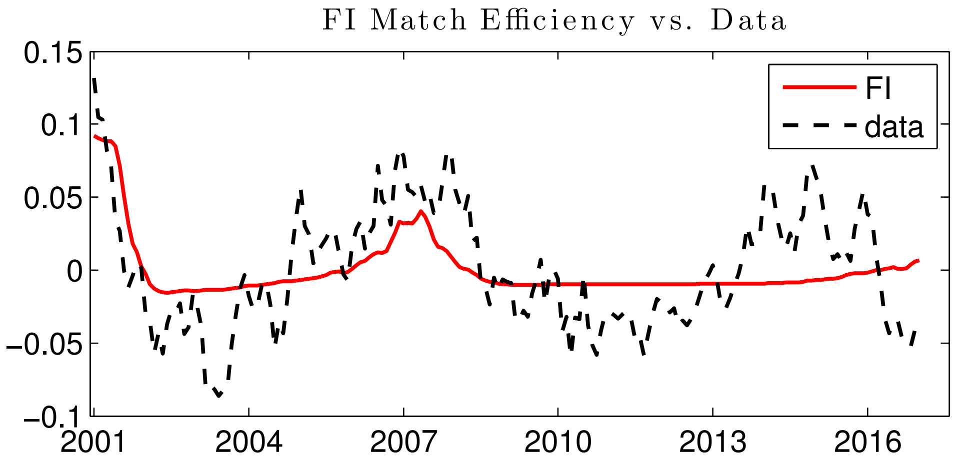


Figure 14: Data vs. Full Information Match Efficiency