Federal Reserve Bank of New York
Staff Reports

Rational Inattention in Hiring Decisions

Sushant Acharya
Shu Lin Wee

Staff Report No. 878
February 2019

Rational Inattention in Hiring Decisions
Sushant Acharya and Shu Lin Wee
Federal Reserve Bank of New York Staff Reports, no. 878
February 2019
JEL classification: D8, E32, J63, J64

Abstract

We provide an information-based theory of matching efficiency fluctuations. Rationally inattentive firms have limited capacity to process information and cannot perfectly identify suitable applicants. During recessions, higher losses from hiring unsuitable workers cause firms to be more selective in hiring. When firms cannot obtain sufficient information about applicants, they err on the side of caution and accept fewer applicants to minimize losses from hiring unsuitable workers. Pro-cyclical acceptance rates drive a wedge between meeting and hiring rates, explaining fluctuations in matching efficiency. Quantitatively, our model replicates the joint behavior of unemployment rates and matching efficiency observed since the Great Recession.

Key words: rational inattention, hiring behavior, matching efficiency, composition of unemployed

____

Acharya: Federal Reserve Bank of New York (email: sushant.acharya@ny.frb.org). Wee: Carnegie Mellon University (email: shuwee@andrew.cmu.edu). The authors would like to thank Kenza Benhima, Ronald Wolthoff, Mirko Wiederholt, Jess Benhabib, Pablo Cuba, Keshav Dogra, Ross Doppelt, Luigi Paciello, Laura Pilossoph, and seminar and conference participants at Carnegie Mellon University, the Federal Reserve Bank of St. Louis, Georgetown University, Penn State, UMD, NUS, SMU, the 2017 NBER Summer Institute (Macro Perspectives), the NYU Stern IUB Workshop, the 2018 Barcelona Summer Forum, the 2018 Salento Macro Meetings, the 2016 Society for Economic Dynamics meeting (Toulouse), several Search and Matching Conferences (Amsterdam, Philadelphia, and Istanbul), and the 2016 Tepper-LAEF Conference.

To view the authors' disclosure statements, visit
https://www.newyorkfed.org/research/staff_reports/sr878.html.

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 *matching efficiency* declined during the Great Recession. Figure 1 shows how measured matching efficiency[2] declined as unemployment rates spiked during the height of the Great Recession. In this paper, we provide an information-based theory of fluctuations in matching 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 measured matching 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. Firms in our model, however, are *rationally inattentive* and have a limited capacity to process information about an applicant. Based on the information they acquire about the job-seeker, firms reject applicants whom they perceive to be unsuitable. Measured matching efficiency in the model then is defined as the average probability, conditional on meeting a job-seeker, that a firm accepts the worker.[3] Because limited capacity to process information implies that firms can still mistakenly hire an unsuitable applicant, average acceptance probabilities depend on the costliness of making a mistake as well as the extent to which firms are informationally constrained.

The losses associated with hiring an unsuitable worker and the firm's uncertainty regarding the type of unemployed job-seeker it meets vary over the business cycle. This in turn affects how informationally constrained firms are over booms and recessions. 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

[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] Matching efficiency is the analog of a Solow Residual in a matching function. We calculate measured matching efficiency as the wedge between meeting and hiring rates.

[3] While we define measured matching efficiency in terms of acceptance rates, it should be noted that firms can choose to accept or reject a worker after evaluating an application and even before interviewing an individual. Thus lower acceptance rates in our model manifest themselves in terms of both lower call-back rates and lower hiring rates conditional on having been interviewed.

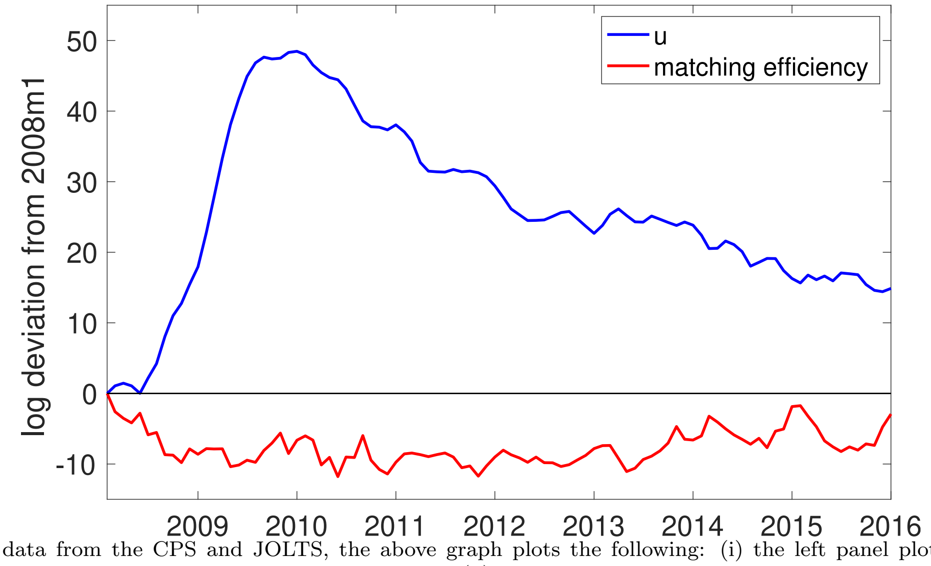
increases the firm's incidence of making a mistake, causing firms to optimally err on the side of caution and reject applicants more often in the downturn. Overall, firms' attempts to avoid hiring unsuitable workers (Type I error) leads them to reject a larger fraction of suitable workers (Type II error).

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 unemployed job seekers to increase but also has the effect of elevating the uncertainty the firm has regarding the job-seeker it meets. Higher uncertainty reinforces the firms' desire for more precise information, which in turn translates into even higher rejection rates, further weighing on measured matching 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, causing unemployment rates to rise by less. 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 measured matching efficiency.

Our paper resolves these issues. Crucially, tighter informational constraints stemming from a limited capacity to process more information about the job-seeker, and the increased cost of making a Type I error during a downturn raise the rejection rate of all job-seekers despite improving average quality in the pool of unemployed job-seekers. Higher rejection rates today increase future firms' uncertainty about the job-seekers they meet, further hampering their ability to distinguish between applicants. Both the increased cost of making a mistake and higher uncertainty counteract the improvement in the average quality of the unemployment pool. In response to a recession where productivity declines by five percent, measured matching efficiency falls by 2% in the rational inattention model and only completely recovers 30 months after the shock. In contrast, measured matching efficiency in the full information model actually increases on impact, as the higher average quality of job-seekers overwhelms the fall in aggregate productivity.

Taking our model to the data, we construct a series of aggregate productivity shocks which induces a sequence of unemployment rates in our model identical to that observed in the data. We then feed these shocks into our model and compute the sequence of implied matching efficiency. We show that the model with rationally inattentive firms outperforms the full information model in matching the joint behavior of unemployment rates and mea-



Notes: Using data from the CPS and JOLTS, the above graph plots the following: (i) the left panel plots the difference in HP-filtered unemployment rates from their 2008m1 value, (ii) the right panel plots the difference in HP-filtered measured matching efficiency from their 2008m1 value. Details on how matching efficiency is calculated can be found in Section 6. We plot the 3-month moving average of the series.

Figure 1: Unemployment rate and matching efficiency since the Great Recession

sured matching efficiency in the data since the Great Recession.

The idea that firms' hiring behavior varies over the business cycle is not a new one. Davis et al. (2012) attribute the divergence between the implied job-filling rate from a constant-returns-to-scale matching function and the vacancy yield[4] during the Great Recession 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 capacity 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 (conditional on a meeting) as lower recruiting intensity.

Several recent papers have also tried to examine and decompose the forces driving the decline in matching 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, Hornstein and Kudlyak (2016) find that search effort is countercyclical and counteracts the decline in job-finding rates caused by procyclical compositional changes in the search efficiency of the unemployed. This result suggests that firms' recruiting behavior is still important towards explaining fluctuations in measured matching efficiency.

In closely related work, Gavazza et al. (2017) consider a model of costly recruiting effort and find that productivity shocks together with financial shocks can explain fluctuations in aggregate matching efficiency. Our model does not feature an explicit cost of recruiting. Rather, fluctuations in the shadow value of information drive changes in firms' recruiting

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

behavior and acceptance rates. Thus, our paper suggests that even in the absence of explicit recruiting costs, firms' hiring behavior can change drastically over the business cycle.

A closely related paper by Sedlacek (2014) considers a full-information model in which firms are differentially selective over the business cycle due to the presence of firing costs. Importantly, Sedlacek (2014) does not feature any permanent worker heterogeneity. Our full information benchmark shows that adding permanent worker heterogeneity into a model similar to his set-up would weaken or even reverse the desire to accept fewer workers since the average quality of job-seekers can improve during a recession. Crucially, permanent worker heterogeneity is necessary to explain an improvement in the average quality of job-seekers during a downturn as exhibited in the data.

While our paper studies how rational inattention 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. Section 3 characterizes how changes in economic conditions affect firms' hiring decisions, and consequently, matching efficiency. Section 4 discusses our calibration approach while Section 5 describes the hiring dynamics in a recession. Section 6 shows how the model is able to replicate the joint behavior of unemployment rates and matching efficiency as observed in the data while section 7 contains additional discussion regarding the key assumptions underlying our model. Section 8 concludes.

2 Model

Time is discrete. Next, 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 $z \in \mathcal{Z}$ drawn from an exogenous and time-invariant distribution given by $\Pi_z(z)$. Workers are either employed or unemployed. Unemployed workers produce $b > 0$ as home-production. If a worker is unemployed, her duration of unemployment, denoted by τ, is publicly known.

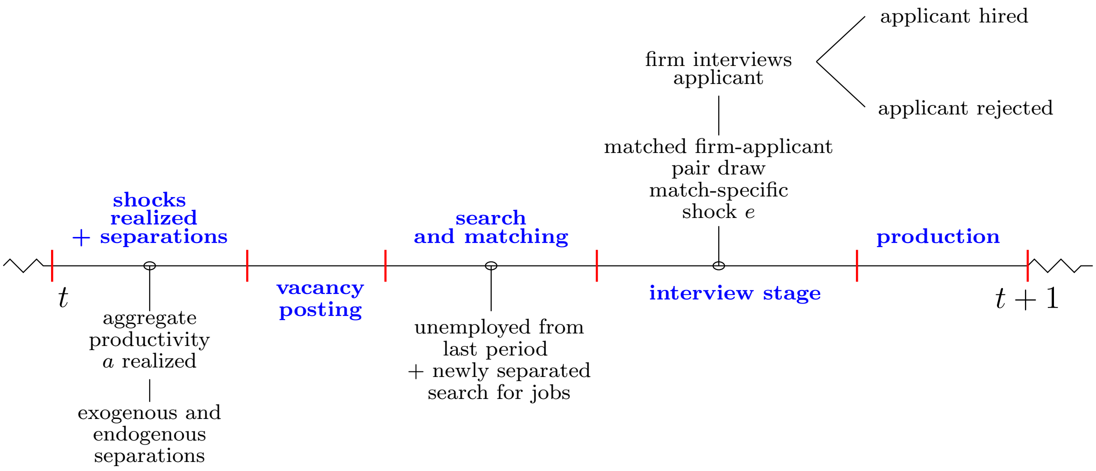


Figure 2: Timeline

Firms A job is a single firm-worker pair with per-period output given by $F(a,z,e) = aze$ where a is the level of aggregate productivity, z is the worker type and e is a match-specific shock. Aggregate productivity a follows an exogenous stationary process. When a firm and worker meet, they draw an i.i.d. match-specific shock $e \in \mathcal{E}$ from a time-invariant distribution $\Pi_e(\cdot)$. This draw of e is constant throughout the duration of the match. The match-specific shock allows for high z type workers to be deemed as bad matches if they draw low e's. Likewise, low z types can be hired if they draw high enough e's.

Labor market A firm that decides to post a vacancy pays a cost $\kappa > 0$. Free-entry determines the measure of vacancies v posted at any date. Search is random and a vacancy contacts a worker at a rate $q = m(v,\ell)/v$ where $m(\cdot,\cdot)$ is a constant returns to scale meeting technology with ℓ denoting the measure of job-seekers in that period. ℓ consists of the unemployed and workers newly separated from their job at the beginning of the period. Analogously, the rate at which a job-seeker meets a vacancy is given by $p = m(v,\ell)/\ell$. For simplicity, we assume that the firm has full bargaining power and makes each worker a take-it-or-leave-it wage offer of b every period.

So far the model is identical to a standard search model with the timing of the model summarized in Figure 2. As per the timeline, we deviate from the standard model by assuming that a firm cannot observe the applicant's *effective productivity* ze at the time of meeting although it can perfectly observe the applicant's unemployment duration.[5] 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*. Given the information revealed in the interview, the firm decides whether to hire the worker. The firm learns ze after production and fires a worker ex-post if the match surplus is negative.

[5]Recent research (see Kroft et al. (2013), Eriksson and Rooth (2014) and Doppelt (2016)) suggest that firms use unemployment duration to inform them about an applicant's suitability. Given that relative job-finding rates in the data exhibit duration dependence, we also allow firms in our model to condition on unemployment duration.

2.1 Hiring strategy

Consider a firm that has met a job-seeker and knows 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 τ, the firm devises an information strategy, i.e. designs an interview. In other words, the firm chooses how informative a signal to acquire about an applicant's 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. We characterize the firm's hiring strategy starting from the second stage problem.

2.1.1 Second-stage problem

Let σ denote the aggregate state variables describing the economy, which will be fleshed out in Section 2.4. 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 types (z, e) with unemployment duration τ.[6] Further denote $\mathcal{G}(z, e \mid \sigma, \tau)$ as the conditional distribution of (z, e) types in the pool of job-seekers with 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 her 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 (z, e). Given its posterior belief $\Gamma(z, e \mid \mathbf{s}, \sigma, \tau)$ about the applicant's (z, e), the firm decides whether to *hire* or *reject* the applicant. Rejecting 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. In particular, the firm hires the applicant iff $\mathbb{E}_\Gamma[\mathbf{x}(a, z, e)] > 0$ where $\mathbf{x}(a, z, e)$ denotes the payoff to the firm if it hires a (z, e) applicant. Thus, the value of such a firm can be written as:

$$J\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) = \max\Big\{0, \mathbb{E}_\Gamma\left[\mathbf{x}(a, z, e)\right]\Big\} \tag{1}$$

2.1.2 First-stage problem

In the first stage, the firm chooses an information strategy to determine the applicant's (z, e). 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

[6] While all job-seekers possess a permanent type z, at the time of meeting a firm, the applicant draws an e. We refer to such a worker as a (z, e) type.

[7] See Section 2.4 for more information.

capacity of the firm as an entropy-based channel capacity constraint following Sims (2003). As per the rational inattention literature, uncertainty about the applicant's (z, e) is measured in terms of entropy, and mutual information measures the reduction of uncertainty about a worker's (z, e). The definition below formalizes these concepts.

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 a firm's choice of information strategy must respect the constraint that information flow are bounded by the finite channel capacity $\chi > 0$:[8]

$$\mathcal{I}(\mathcal{G}, \Gamma \mid \sigma, \tau) = H(\mathcal{G}(\cdot \mid \sigma, \tau)) - \mathbb{E}_s H\left(\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right) \leq \chi \tag{2}$$

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.

The firm's information strategy is akin to asking an applicant a series of questions to reduce its uncertainty about the worker's (z, e). Each additional question provides the firm with incremental information to help it make a more informed decision over whether to accept or reject an applicant. However, each additional question consumes the finite processing capacity a firm possesses. The lower χ is, the lower the firm's 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] Through the interview, firms 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*

[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).

applicant with unemployment duration τ in aggregate state σ is given by a joint distribution of signals $\mathbf{s}$ *and types,* $\Gamma(z, e, \mathbf{s} \mid \sigma, \tau)$ *such that:*

$$\mathcal{G}(z, e \mid \sigma, \tau) = \int_{\mathbf{s}} d\Gamma(z, e, \mathbf{s} \mid \sigma, \tau) \tag{3}$$

Equation (3) is 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 strategy is indexed by τ. Importantly, unemployment duration τconveys 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 latter has occurred, 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, $g(z, e \mid \sigma, \tau)$ 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, $g(z, e \mid \sigma, \tau)$, the firm's information strategy differs depending on the duration τ of the applicant it meets.

The extent to which the firm's processing constraint (2) 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. Formally, the firm's first-stage problem conditional on meeting an applicant with duration τ can be written as:

$$\mathbf{V}(\sigma, \tau) = \max_{\Gamma \in \Delta} \sum_{z} \sum_{e} \int_{s} J\left[\Gamma(\cdot \mid \mathbf{s}, \sigma, \tau)\right] d\Gamma(\mathbf{s} \mid z, e, \sigma, \tau) g(z, e \mid \sigma, \tau) \tag{4}$$

subject to the information processing constraint (2). $J\left[\Gamma\left(z, e \mid \mathbf{s}, \sigma, \tau\right)\right]$ denotes the ex-ante payoff from the second stage for a type (z, e) applicant given signals $\mathbf{s}$ as defined in (1). As the firm does not know the applicant's true (z, e), its 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 firm's problem as specified in (4) is not trivial to solve as it allows firms to choose signals of any form. Rather than solving for the optimal signal structure, we follow Matejka and McKay (2015) and solve the identical but transformed problem in terms of state-contingent choice probabilities and their associated payoffs.

Let $\mathcal{S}$ be the optimally chosen set of signals that lead the firm to *hire* an applicant of type (z, e) with duration τ in state σ. Under this information strategy, the firm hires a type (z, e) applicant with duration τ with probability $\gamma(z, e \mid \sigma, \tau)$ which is equal to 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, $\mathcal{P}(\sigma, \tau)$, is given by:

$$\mathcal{P}(\sigma, \tau) = \sum_z \sum_e \gamma(z, e \mid \sigma, \tau) g(z, e \mid \sigma, \tau) \tag{5}$$

Finally, conditional on meeting a worker, $\mathcal{P}(\sigma)$ is the average probability of hiring a worker in state σ which corresponds to measured matching efficiency in our model (see Section 2.3 for more details). The following Lemma presents the reformulated problem:

Lemma 1 (Reformulated First-Stage Problem). *The problem in* (4) *is equivalent to the transformed problem below:*

$$\mathbb{V}(\sigma, \tau) = \max_{\gamma(z,e|\sigma,\tau) \in [0,1]} \sum_z \sum_e \gamma(z, e \mid \sigma, \tau) \mathbf{x}(a, z, e) g(z, e \mid \sigma, \tau) \tag{6}$$

subject to:

$$\mathcal{H}\Big(\mathcal{P}(\sigma, \tau)\Big) - \sum_z \sum_e \mathcal{H}\Big(\gamma(z, e \mid \sigma, \tau)\Big) g(z, e \mid \sigma, \tau) \ \leq \ \chi \tag{7}$$

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

Proof. See Appendix D. □

Intuitively, the LHS of (7) measures the information flow based on the optimal signal choices as in (2) but expressed in terms of choice probabilities.[10] 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 1 (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:*

$$\gamma\left(z, e \mid \sigma, \tau\right) = \frac{\mathcal{P}\left(\sigma, \tau\right) e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma,\tau)}}}{1 + \mathcal{P}\left(\sigma, \tau\right) \left[e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma,\tau)}} - 1\right]} \tag{8}$$

[10] Appendix D uses the symmetry property of mutual information to show that the constraint (2) is equivalent to (7).

where $\mathcal{P}(\sigma, \tau)$ is defined in (5) and $\lambda(\sigma, \tau)$ is the multiplier on (7) and represents the shadow-value of reducing uncertainty by one nat.[11]

Proof. See Appendix A. $\square$

Equation (8) reveals an important feature of the information strategy. Consider two applicants with the same e and τ, but with different worker-productivity $z_1 > z_2$. Using (8):

$$\log \frac{\gamma\left(z_1, e \mid \sigma, \tau\right)}{1-\gamma\left(z_1, e \mid \sigma, \tau\right)} - \log \frac{\gamma\left(z_2, e \mid \sigma, \tau\right)}{1-\gamma\left(z_2, e \mid \sigma, \tau\right)} = \frac{\mathbf{x}(a, z_1, e) - \mathbf{x}(a, z_2, e)}{\lambda(\sigma, \tau)} \tag{9}$$

Equation (9) implies that firms choose 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. In other words, signals are chosen so as to reduce the incidence of Type II errors, allowing firms to accept more productive applicants more often on average. Further, (9) 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 $\gamma(z_2, e \mid \sigma, \tau)$. A higher λ implies that firms are starved of information and are less able to distinguish between different types. In the limit as $\lambda \to \infty$, 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. The other extreme is the case in which firms have no informational constraints, i.e. $\chi \to \infty$. In other words, firms have full information about the type of the applicant implying that the shadow value of information $\lambda \to 0$ in which case (8) becomes:[12]

$$\gamma(z, e \mid \sigma, \tau) = \begin{cases} 1 & \text{if } \mathbf{x}(a, z, e) \geq 0 \\ 0 & \text{else} \end{cases} \tag{10}$$

Since firms can perfectly ascertain an applicant's (z, e), they only hire an applicant if $\mathbf{x}(a, z, e) \geq 0$. Even with full-information, $\mathcal{P}(\sigma, \tau) < 1$ if some applicants have $\mathbf{x}(a, z, e) < 0$.

2.2 Value of a firm

Thus far, we have not specified how the payoff from hiring a (z, e) applicant is determined. 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:

$$\mathbf{x}(a, z, e) = F(a, z, e) - b + \beta \mathbb{E}_{a'|a} \Big(1 - d(a', z, e)\Big) \mathbf{x}(a', z, e) \tag{11}$$

where $d(a, z, e) \in \{\delta, 1\}$. Since the firm learns the worker's ze perfectly after production, the firm can choose to fire the worker, $d(a, z, e) = 1$, if it finds the worker to be unsuitable

[11] Using a logarithm with base e, entropy is measured in nats. An equivalent but alternative way to measure entropy is to use a logarithm with base 2. The measure of entropy would then be in terms of bits.

[12] See Appendix B for details.

to retain, i.e. $\mathbf{x}(a, z, e) < 0$. Even if the worker is deemed suitable, she can still be separated from the firm at an exogenous rate δ.

2.3 Free entry

Free entry determines the equilibrium market-tightness and the rate at which firms and workers meet. Let $g_\tau(\tau \mid \sigma)$ be the probability mass of job-seekers of duration τ in aggregate state σ:

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

Then from the free-entry condition, we have:

$$\kappa \geq q(\theta) \sum_\tau \mathbb{V}(\sigma, \tau) g_\tau(\tau \mid \sigma) \tag{12}$$

where $\theta = v/\ell$ is the labor market tightness. $\theta = 0$ if (12) holds with a strict inequality. Unlike the standard DMP model, the job-filling rate in our model can be decomposed into two components. Free entry pins down the first component - the *contact* rate - $q(\theta)$, 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, $\mathcal{P}(\sigma, \tau)$.[13] Formally, the aggregate job-filling rate in our model is 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 matching efficiency in the model. $\mathcal{P}(\sigma)$ forms a wedge between the job-filling rate and the contact rate.

2.4 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 $t-1$. σ_t is common knowledge and hence firms 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:

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

[13] This is subsumed inside $\mathbb{V}(\sigma, \tau)$ in equation (12).

The first part of (13) shows that job-seekers of type z with zero unemployment duration are comprised of newly separated workers who were employed at the end of $t-1$. The second line in (13) refers to all the z-type unemployed with duration τ at the end of $t-1$. 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 by:

$$u_t(z,\tau) \;=\; \ell_t(z,\tau-1)\Big\{1-p(\theta_t)+p(\theta_t)\sum_e \pi_e(e)\Big(1-\gamma(z,e \mid \sigma_t,\tau-1)\Big)\Big\} \tag{14}$$

The first term on the RHS of (14) refers to all z-type job-seekers of duration $\tau-1$ at the beginning of t. With probability $1-p(\theta_t)$, such a job-seeker fails to meet a firm and remains unemployed. With probability $p(\theta_t)$, the job-seeker meets a firm, draws match productivity e with probability $\pi_e(e)$, but is rejected with probability $1-\gamma(z,e \mid \sigma_t,\tau-1)$ 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 $\ell_t(z,\tau-1)$ 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 is given as:

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

Equation (15) shows that the mass of employed workers of type (z,e) in period t is composed of two terms. The first term denotes the fraction of employed workers of type (z,e) at the end of $t-1$ who remain with the firm at the beginning of t. The second term refers to all job-seekers at date t who meet a firm with probability $p(\theta_t)$, 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 $\gamma(z,e \mid \sigma_t,\tau)$.

Finally, we have the accounting identity that the sum of employed and unemployed workers of type z must equal to the population of type z workers 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 $\ell_t(\tau)$ as the mass of job-seekers of duration τ and ℓ_t as the total mass of job-seekers, i.e.

$$\ell_t(\tau)=\sum_z \ell_t(z,\tau) \quad \text{and} \quad \ell_t=\sum_\tau \ell_t(\tau)$$

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

$$g_z(z \mid \sigma_t,\tau)=\frac{g_{z,\tau}(z,\tau \mid \sigma_t)}{g_\tau(\tau \mid \sigma_t)}=\frac{\ell_t(z,\tau)}{\ell_t(\tau)}\ ,\ \forall\tau\geq 0 \tag{16}$$

where $g_z(z \mid \sigma_t, \tau)$ is the share of job-seekers of duration τ who are of type z in state σ_t. Since the match-specific shock e is drawn independently of z and of any past realizations whenever a worker matches with a firm, the joint probability mass of a type (z, e) job-seeker is simply given by $g_z(z \mid \sigma_t, \tau)\pi_e(e)$, i.e.

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

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. Thus, a firm's prior at any date t about a worker's type (z, e) given their unemployment duration τ and σ is captured by the joint distribution $\mathcal{G}(z, e \mid \sigma_t, \tau)$.

3 Uncovering the forces that affect hiring decisions

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 $\alpha \leq 0.5$ and $1 - \alpha$ respectively. Then, the firm's hiring problem can be written :

$$\Pi(a, \alpha) = \max_{(\gamma_H, \gamma_L) \in [0,1]^2} \alpha\gamma_H \mathbf{x}(a, z_H) + (1 - \alpha)\gamma_L \mathbf{x}(a, z_L) \tag{17}$$

s.t.

$$\mathcal{I}(\alpha) \equiv \mathcal{H}(\mathcal{P}) - \alpha\mathcal{H}(\gamma_H) - (1 - \alpha)\mathcal{H}(\gamma_L) \leq \chi \tag{18}$$

where $\mathbf{x}(a, z) = az - b$ and $\mathcal{P} = \alpha\gamma_H + (1 - \alpha)\gamma_L$ denotes the average probability that the firm hires an applicant. Suppose aggregate productivity is such that $a \in [b/z_H, b/z_L]$. 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, i.e. if the firm could observe the applicant's type, it would choose $\gamma_H = 1, \gamma_L = 0$. However, limited channel capacity χ may prevent the firm from identifying a z_H applicant perfectly. Figure 3 depicts the hiring decisions for different values of χ.

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 choice. 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.

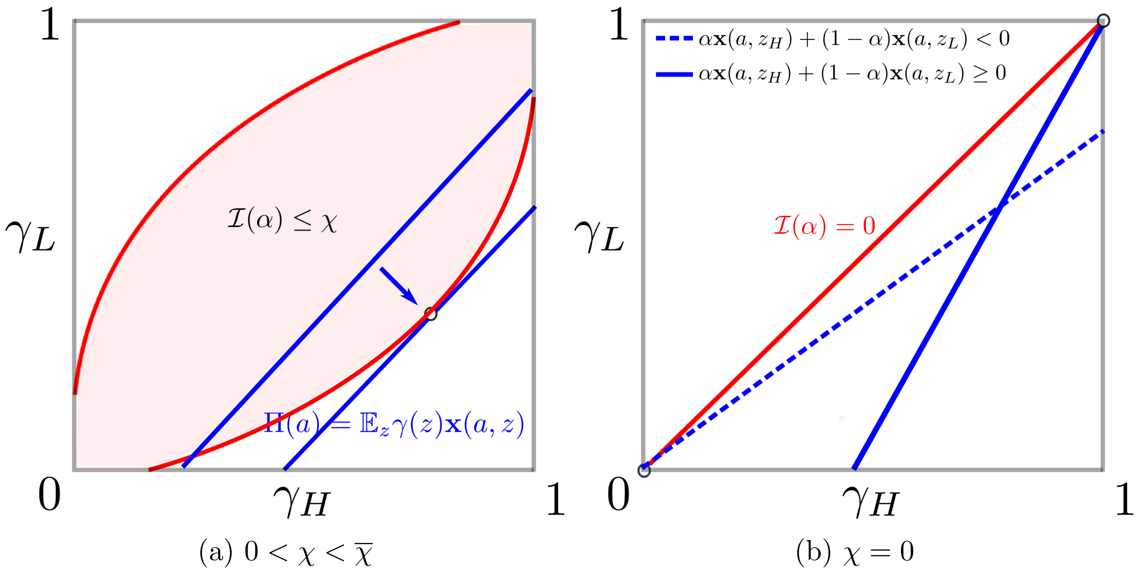


Figure 3: Optimal Hiring Decisions

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 iso-profit curves are tangent to the boundary of the constraint set:

$$-\frac{\alpha \mathbf{x}(a, z_H)}{(1-\alpha)\mathbf{x}(a, z_L)} = -\frac{\alpha \left[\mathcal{H}'\Big(\alpha\gamma_H + (1-\alpha)\gamma_L\Big) - \mathcal{H}'(\gamma_H)\right]}{(1-\alpha)\left[\mathcal{H}'\Big(\alpha\gamma_H + (1-\alpha)\gamma_L\Big) - \mathcal{H}'(\gamma_L)\right]} \tag{19}$$

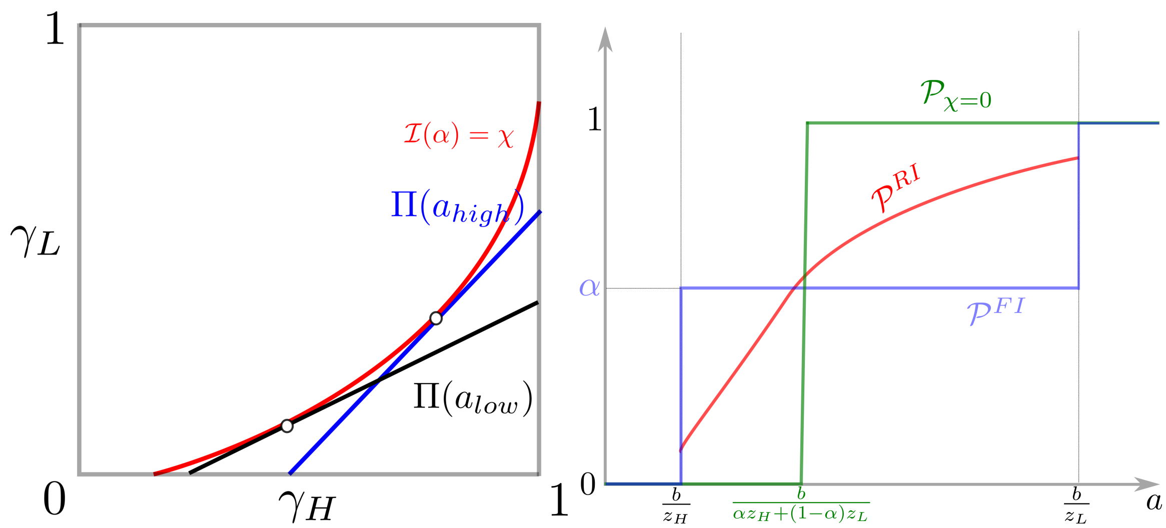
Notice that the LHS of (19) is the slope of the iso-profit curve while the RHS is the slope of the constraint set (18). The optimal choices γ_H and γ_L are the solution to (18) and (19). Intuitively, the firm wants to choose the highest γ_H and the lowest γ_L possible but (19) reveals that in trying 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 with $\chi = 0$, the only feasible choices lie along the diagonal, i.e. $\gamma_H = \gamma_L = \gamma$. The optimal choices are then described by a *bang-bang* solution - the firm hires any applicant, $\gamma = 1$ if $\alpha \mathbf{x}(a, z_H) + (1-\alpha)\mathbf{x}(a, z_L) \geq 0$, and rejects 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. (18) shows that while changes in α alter the set of feasible (γ_H, γ_L), changes in a do not. Changes in a, however, do affect the expected payoff from hiring in (17). From the LHS of (19), one can see that the slope of the iso-profit curves flatten as a falls.[14]. Correspondingly, both γ_H, γ_L must decline to maintain the equality.[15] Figure 4a[16] shows that a lower a leads to flatter iso-profit curves, shifting the point of tangency to the constraint towards the south-west, lowering optimal γ_H and γ_L.

Lemma 2 (Comparative Statics)**.** *Under the optimal information strategy, γ_H and γ_L are increasing in a. Thus, average acceptance probability $\mathcal{P}$ is increasing in a.*

Proof. See Appendix C. □

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.[17] 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 2.3, declines in $\mathcal{P}$ correspond to falls in measured matching efficiency in our model.



(a) Change in aggregate productivity a

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

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

Figure 4b shows that the average acceptance probability $\mathcal{P}$ is more sensitive to changes in aggregate productivity when firms are informationally constrained. The blue line denotes the optimal $\mathcal{P}^{FI}$ under full information (FI). In this case, firms always observe the applicant's type and never make mistakes in hiring. For $a \in [b/z_H, b/z_L]$, FI firms only hire the z_H type, ($\gamma_H = 1, \gamma_L = 0$). As a result, optimal acceptance rates under full information in this

[14]The slope of the iso-profit curves is increasing in a since $-\frac{\mathbf{x}(a,z_H)}{\mathbf{x}(a,z_L)}$ is increasing in a.

[15]See Appendix C for details.

[16]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.

[17]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.

region are constant at $\mathcal{P}^{FI} = \alpha$ as shown by the blue line in Figure 4b. Under rational inattention (RI), 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 2 shows that $\partial \mathcal{P}/\partial a > 0$ in this range as depicted by the upward sloping red line labeled $\mathcal{P}^{RI}$ in Figure 4b.[18] Thus, relative to the full information case, small changes in a only result in large changes in $\mathcal{P}$, our measure of matching efficiency, when firms are informationally constrained.[19]

How does a change in the distribution of job-seekers affect optimal choices? Unlike a, an increase in the fraction of z_H types, α, affects both expected payoffs (17) and the set of feasible choices of (γ_H, γ_L) as in (18). Figure 5a shows that an increase from $\alpha_0 \leq 0.5$ to $\alpha_1 \in (\alpha_0, 0.5]$ makes the iso-profit curves steeper (LHS of (19)) since there are more high types in the population. The solid blue lines represent the set of steeper iso-profit curves while the dashed blue line represents the flatter iso-profit curves associated with lower α.

Holding the constraint set fixed, a higher fraction of z_H types (or higher average quality of job-seekers) induces the firm to raise acceptance rates - denoted by a move from the choice $\mathcal{E}_0$ to $\mathcal{E}_1$. The higher average quality of job-seekers induces firms to raise γ_H. However, in doing so, firms' limited processing capacity forces them to also increase γ_L as they raise their hiring probability of z_H types.

Importantly, there is also a countervailing force which tends to reduce acceptance rates. An increase in α closer to 0.5 also increases the initial uncertainty, $H(\mathcal{G})$, the firm has over the type of job-seeker he meets,[20] implying a contraction in the feasible set as depicted in Figure 5a by an inward shift in the frontier of feasible information strategies from the red-dashed line to the red solid line. This tighter informational constraint on the firm causes the shadow value of information to rise as depicted by the higher blue line in Figure 5b. Recall from (9) that a higher λ restricts the firm's ability to distinguish between worker types, forcing the optimal choice of (γ_H, γ_L) towards the 45 degree line.

[18] 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 $\mathcal{P}$ causing it to fall from 1 to 0.

[19] While the example considered a case with two types, it should be noted that for a larger but finite number of types, $\mathcal{P}^{FI}$ would resemble a step function and still be constant for small changes in a. In contrast, under rational inattention, $\mathcal{P}^{RI}$ would be more sensitive to changes in a. Thus, the result does not depend on the number of types.

[20] 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 $\alpha \in [0, 0.5]$, an increase in α raises the initial uncertainty and causes the firm to be more constrained in processing information to lower his posterior uncertainty compared to the case with a lower α.

An alternative way to see this is by rewriting (19) as:

$$\frac{\mathbf{x}(a,z_H)}{\mathbf{x}(a,z_L)} = \frac{\mathcal{H}'\Big(\alpha\gamma_H + (1-\alpha)\gamma_L\Big) - \mathcal{H}'(\gamma_H)}{\mathcal{H}'\Big(\alpha\gamma_H + (1-\alpha)\gamma_L\Big) - \mathcal{H}'(\gamma_L)}$$

The LHS is independent of changes in α. However, since $\mathcal{P} = \alpha\gamma_H + (1-\alpha)\gamma_L$ is affected by α, (γ_H, γ_L) must adjust to make the RHS the same as the unchanged LHS. As mentioned above, there are two opposing forces affecting how (γ_H, γ_L) adjust. The increase in α raises the average quality of job-seekers and is a force towards firms raising acceptance probabilities. However, a higher α (closer to 0.5) also raises the initial uncertainty $H(\mathcal{G})$ and this lowers acceptance probabilities. Figure 5a depicts the case in which the increase in average quality is overwhelmed by higher uncertainty which has the effect of lowering γ_H and raising γ_L.[21] At the new optimum (denoted by $\mathcal{E}_2$ in the figure), γ_H is lower.

Interestingly and unlike the case with changes in a, even small changes in α can affect the acceptance rate of firms in the FI model. In fact, $\mathcal{P}^{FI}$ changes one-for-one with the changes in α, the proportion of z_H job-seekers. The absence of a counteracting force from increasing uncertainty in the FI model implies that the increase in α only serves to raise the average quality of job-seekers, causing FI firms to increase their acceptance rates. Notably, this implies that $\mathcal{P}^{FI}$ can actually increase during a recession if the average quality of job-seekers improves. We highlight this phenomenon in greater detail in Sections 5 and 5.2.

While the distribution of job-seekers (and hence $H(\mathcal{G})$) is given exogenously in this static limit, the distribution endogenously evolves over the business cycle in the dynamic model. As we show in Section 5.1, 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. This in turn raises the $H(\mathcal{G})$ and weighs on firms' hiring activities in the recovery.

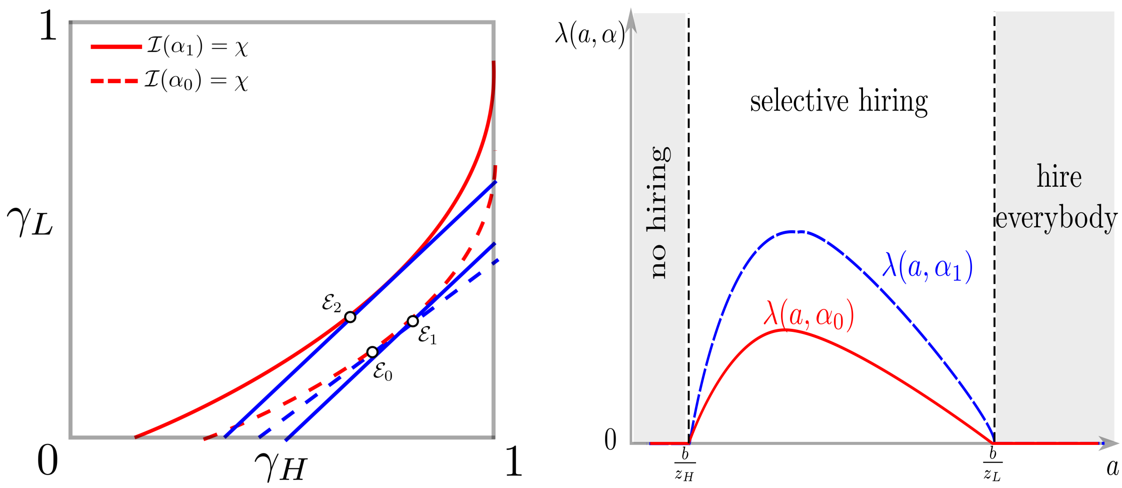
The Shadow Value of Information The shadow value of information λ (the multiplier on the information processing constraint) summarizes the extent to which a firm is informationally constrained. Since the shadow value of information in the FI model is always 0,[22] we focus on how λ is affected by a and α in the RI model.

The relationship between λ and aggregate productivity 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 $\lambda = 0$.[23] For $a \in [b/z_H, b/z_L]$, firms value information as identifying and hiring a z_H worker

[21] The figure depicts the case we find to be quantitatively relevant in our numerical exercises. More generally, depending on whether the shift in of the constraint set is larger or the change in slope of the iso-profit curves is bigger, the first effect might dominate.

[22] This is because firms in the FI model can observe the applicant's type perfectly.

[23] For very low levels of a and very high levels of a firms' decisions under FI and RI are identical. For very low a: $a < \frac{b}{z_H}$, firms incur losses from hiring any applicant and both RI and FI firms choose $\gamma_H = \gamma_L = \mathcal{P} = 0$ in this range. Similarly, for very high a: $a > \frac{b}{z_L}$, both RI and FI firms are happy to hire any applicant and $\gamma_H = \gamma_L = \mathcal{P} = 1$ in this range. Since implementing information strategies with $\gamma_H = \gamma_L$



(a) Effect of changing α on γ_H, γ_L (b) Shadow value of information

Figure 5: Comparative Statics

brings them positive payoffs. When a falls from b/z_L the firm increasingly wants to avoid hiring z_L types but would still like to hire z_H types as they still provide positive payoffs. As such, firms value an additional unit of processing capacity at this point as it helps them to better distinguish the two types, and hence λ is higher. As a continues falling and approaches b/z_H, the payoffs from hiring either type is low. The firm's concern over hiring a z_H applicant is outweighed by the cost of mistakenly hiring a z_L type, lowering the firm's need to distinguish between types and reducing λ. Thus, λ starts to decline as $a \to b/z_H$.[24]

In contrast, given aggregate productivity a, a lower α relatively decreases λ as the distribution becomes skewed towards a particular z type worker. Consequently, the solid red curve in Figure 5b corresponding to α_0 lies weakly below the dashed blue curve, corresponding to $\alpha_1 \in (\alpha_0, 0.5]$. When α approaches 0.5, firms have more initial uncertainty and require more information to distinguish between types.

4 Calibration

We discipline the model using data on US aggregate labor market flows. A period in our model is a month. Consistent with an annualized risk free rate of 4%, we set $\beta = 0.9967$. The rate at which a worker meets a firm takes the form of $p(\theta) = \theta(1 + \theta^\iota)^{-1/\iota}$ where $\iota = 0.5$ as standard in the literature.[25] We assume that (log) aggregate productivity follows an AR(1) process: $\ln a_t = \rho_a \ln a_{t-1} + \sigma_a \varepsilon_t$ where $\varepsilon_t \sim N(0, 1)$. We set $\rho_a = 0.9$, and $\sigma_a = 0.0165$ as in Shimer (2005).

is always feasible under RI, RI firms can make the same decisions as FI firms. In these regions, the shadow value of information $\lambda = 0$ and information constraints do not matter.

[24]Despite this non-monotonicity, in our quantitative exercises, λ rises when a falls implying that a never falls to such a low level such that the labor market shuts down completely with firms rejecting all workers.

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

The remaining parameters are chosen to minimize the distance between model generated moments and their empirical counterparts. We use the following moments to discipline our model. We target an employment-to-unemployment transition rate (EU) of 3.2% to discipline our choice of δ, the exogenous separation rate in our model. 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 2.5 years.[26] 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. Following Hall (2009), we set the value of home production, b, such that it is equal to 70% of output.

Following a large literature that has assumed that worker heterogeneity is drawn from a Beta distribution (see Jarosch and Pilossoph (2016), Lise and Robin (2017) for example), we assume that the unobserved worker fixed effect, z, is drawn from a discretized Beta distribution, i.e. $z \sim Beta(A_z, B_z) + \underline{z}$. We set $\underline{z} = 0.5$ which is a normalization that ensures that the lowest z type worker is still employable if she draws a high enough match quality e. The match quality, e, is drawn from the Beta distribution $e \sim Beta(A_e, B_e)$.[27]

The vacancy posting cost, κ, the channel capacity, 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. Thus, in addition to the aggregate unemployment rate, we use information on the relative job-finding rates across workers of different unemployment duration to govern these parameters. We target an aggregate unemployment rate of 6%, 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 CPS to calculate the relative job-finding rates in the data. These relative job-finding rates help us to pin down the parameters governing heterogeneity in z and e. 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 form: $\frac{\text{UE}(\tau)}{\text{UE}(1)} = \pi_1 + (1-\pi_1)exp(-\pi_2\tau)$ where τ is the unemployment duration, and $\frac{\text{UE}(\tau)}{\text{UE}(1)}$ is the job-finding rate of an unemployed individual of duration τ relative to an unemployed individual with 1 month of unemployment duration.[28] We target the fitted relative job-finding rates from this regression.

To see why the decline in relative job-finding rates contains crucial information which helps us discipline the distribution of z and e, we consider what happens when there is only heterogeneity in z and when there is only heterogeneity in e. For example, if there were no permanent worker types, i.e. all individuals have the same z, relative job-finding

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

[27] Specifically we set the number of worker productivity types to be $n_z = 15$ and the number of match-specific shocks to $n_e = 15$. See Appendix E.1 for details on the construction of $\mathcal{Z}$ and $\mathcal{E}$. In previous versions, we have tried different combinations of n_z and n_e. Qualitatively, this does not change any of our results. We pick $n_z = n_e = 15$ so as reduce the computational burden.

[28] In our regression, we use the extended model as in Kroft et al. (2016) and include controls for gender, age, race, education and gender interactions for age, race and education variables. We cluster all those who are more than 10 months unemployed into a single bin as relative job-finding rates are relatively flat for those unemployed for more than 6 months.

rates across duration would be flat as draws of e are i.i.d and independent of past matches - unemployment duration would not provide any useful information about the applicant's suitability. 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. This is because longer spells of unemployment would then suggest a higher number of rejections and signal that the applicant is of a low z type. When we discuss the model's fit, we will demonstrate that in order to match the empirical relative job-finding rates (depicted by the black dashed line in Figure 6a), both heterogeneity in e and z are necessary to generate the sharp initial decline and subsequent flattening out across durations. Therefore, relative job-finding rates help us discipline the parameters governing the heterogeneity in z and e.

We have 8 parameters to estimate $\{\chi, \kappa, \delta, b, A_z, B_z, A_e, B_e\}$ and we target 8 moments: average monthly separation rate, aggregate unemployment rate, unemployment benefits worth 70% of output and relative job-finding rates for unemployment spells greater than one month. Table 1 summarizes both the fixed and inferred parameters.

Table 1: Model Parameters

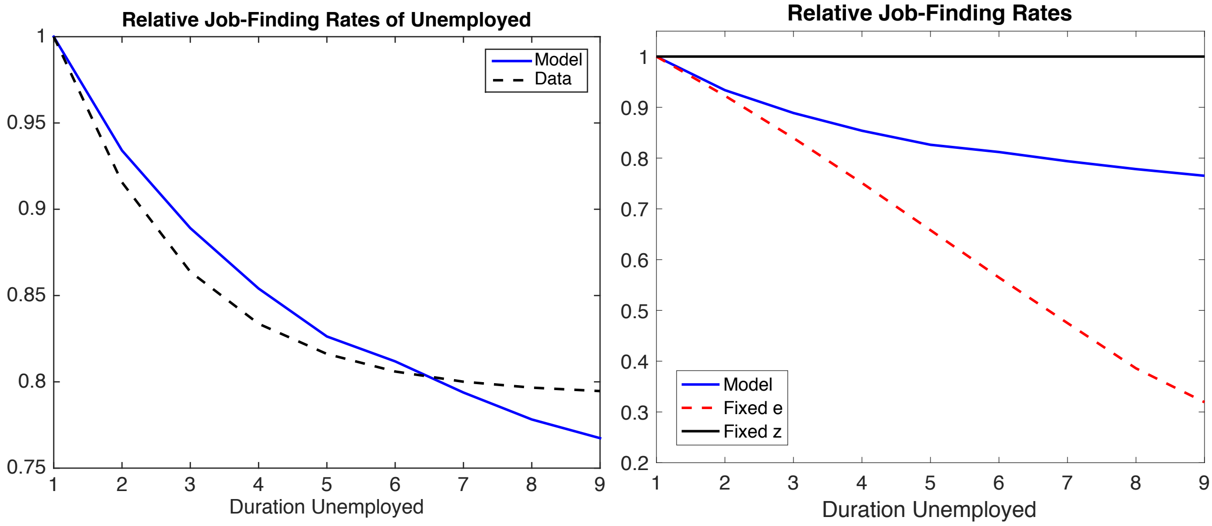
Externally calibrated

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	Petrongolo and Pissarides (2001)

Parameters estimated internally

Parameter	Description	Value
b	home production	0.2310
δ	exogenous separation rate	0.0190
κ	vacancy posting cost	0.0001
χ	channel capacity	0.0768
A_z	first shape parameter (z)	1.717
B_z	second shape parameter (z)	6.377
A_e	first shape parameter (e)	4.858
B_e	second shape parameter (e)	15.32

Overall, the model does a good job of matching moments in the data. The calibrated value of b gives rise to a home production value that is 71% of total output, close to our targeted 70%. Our parameterization for δ and κ generates an EU rate and unemployment rate of 3.1% and 6.01% respectively. Figure 6a also shows how well our model-implied relative job-finding rates replicates its empirical counterpart. Figure 6b shows that the empirical relative job-finding rates contain important information on the heterogeneity in e



(a) Model Fit

(b) Fixed e vs. fixed z

Figure 6: Relative job-finding rates

and z. With only heterogeneity in e, the solid black line in Figure 6b reveals that relative job-finding rates would be perfectly flat. Absent heterogeneity in e, the dashed red line in Figure 6b shows that relative job-finding rates would be strictly declining when every one draws the same fixed e.[29]

The shadow cost of processing information In steady state, the average shadow value of information, $\lambda = \sum_\tau g_\tau(\tau)\lambda(\tau) = 0.37$. 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. Recent work by Gavazza et al. (2017) document that the annual median spending on recruiting activities is about \$3,479 per worker or about 92% of median monthly earnings. In our model, the average λ is around 13% of annual earnings,[30] which is close to estimates found in the data. 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.

[29] We set the fixed level of $e = 0.33$, this ensures that at least half of the z types in our model are employable. If even the worst type of individual were employable with a known fixed e, then firms would accept all workers with probability 1 and there would be no difference in relative job-finding rates since all workers have the same probability of contacting a firm.

[30] In our model, households' annual earnings are pinned down by $b \times 12$.

5 Understanding the dynamics of hiring decisions

Since hiring decisions in our model are non-linear and state-dependent, we study the impulse response of key labor market variables to negative aggregate productivity shocks of different sizes. We also compare the response of the economy to a shock of the same magnitude, but starting from a different initial distribution of job-seekers.

We compare the responses of the model with RI firms to the responses when firms have full information. In doing so, we highlight how the absence of informationally constrained firms in an environment with permanent worker heterogeneity and endogenous separations gives rise to counterfactual outcomes when the economy is hit by a negative productivity shock. The RI model highlights how the changing shadow value of information over the business cycle can help rationalize the observed behavior of labor market variables, and in particular, measured matching efficiency. 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.[31]

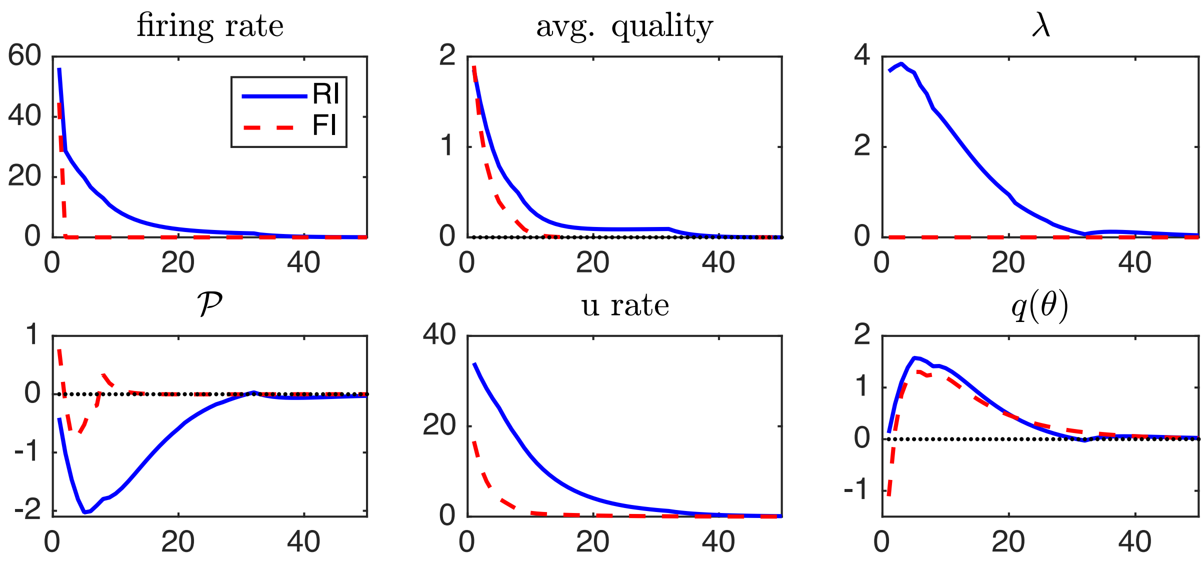


Figure 7: Response to a 5% shock

Figure 7 shows the response of key labor market variables to a 5% fall in aggregate productivity which then gradually reverts back to its mean.[32] The decline in a causes firing rates to spike[33] on impact in both the RI and FI models (blue solid and red dashed lines respectively in the top left panel of Figure 7) as firms release workers from matches that no longer have positive surplus. Unlike the RI model, the FI model only observes higher firing rates on impact. Since FI firms can perfectly identify and hire only suitable workers,

[31]In order to keep the comparison fair, we recalibrate the FI model but hold fixed the unconditional distribution of worker types $\Pi_z(z)$ and the unconditional distribution of match quality $\Pi_e(e)$ so that the two model economies have the same types of workers. See Appendix F for more details.

[32]Figure 15a in Appendix I shows how quickly productivity recovers over time.

[33]We define firing rates as the fraction of employed who are separated at the start of the period.

FI firms would not terminate these matches as they continue to have positive surplus as aggregate productivity recovers. Thus, only exogenous separations occur in the FI model after the first period. In contrast, both exogenous and endogenous separations occur in the RI model since RI firms can still mistakenly hire workers who bring negative surplus. As such, the RI model generates persistence in firing rates over the recession. While not targeted in our model, the persistence in firing rates is consistent with the empirical finding that the job tenure of newly employed workers is lower when the unemployment rate is high (see for example Bowlus (1995)).

Higher firing rates on impact change the composition of job-seekers - the average quality of job-seekers increases in both the RI and FI models as demonstrated by the top middle panel of Figure 7. Both previously employed low z type workers who had drawn high e's and high z type workers who had drawn low-to-middling e's are released into the pool of job-seekers when firms raise their retention standards in response to a fall in a. In response to the 5% fall in productivity, more of the latter type are released and this drives up the average quality of job-seekers.[34] While the initial increase in the average quality is of roughly the same magnitude, the rate at which it dissipates substantially differs between the two models. Crucially, the rate at which higher average quality of job-seekers decays depends on firms' ability to correctly identify and hire suitable workers.

Lower aggregate productivity lowers $\mathbf{x}(a, z, e)$ for all applicants, magnifying the losses a firm incurs from hiring an unsuitable worker. As a firm can always guarantee itself a zero payoff by rejecting an applicant, the lower $\mathbf{x}(a, z, e)$ and higher likelihood of incurring losses makes rejecting applicants more attractive. In order to be willing to hire an applicant, the firm desires more information to gauge the suitability of the applicant. However, their finite processing capacity prevents them from doing so, as reflected by the higher shadow value of information λ. The combination of a higher λ and lower $\mathbf{x}(a, \cdot, \cdot)$ further pushes firms to err on the side of caution and reject applicants more often to avoid making losses.

Accordingly, measured matching efficiency, $\mathcal{P}$, in the RI model falls by 0.5% on impact (first panel, second row of Figure 7). In subsequent periods, $\mathcal{P}$ continues to decline and falls by as much as 2% below steady state as the increase in the average quality of job-seekers dissipates. The smaller initial decline in $\mathcal{P}$[35] is not just because higher average quality of job-seekers counteracts part of the fall in aggregate productivity on impact, but is also due to the fact that the change in the composition of job seekers over time makes it harder for firms to identify whether the applicant is suitable to hire. Higher rejection rates cause the pool of job-seekers to be more disparate and for this disparity to persist for an extended period of time (recall from Section 3 that a more 'uniform' distribution of job-seekers tends to make the firm's inference about an applicant's type harder). We explore this in greater

[34] Average quality need not always increase in response to negative shocks as we show later in our exercise with a larger fall in aggregate productivity.

[35] Note that if the distribution of job-seekers had instead remained the same as in steady state, then the fall in a would have resulted in $\mathcal{P}$ falling by close to 3% on impact (See Figure **??** in Appendix I). The smaller initial decline in Figure 7 is because the distribution of job-seekers change on impact.

detail in Section 5.1 and show that firms are further informationally constrained when they have more initial uncertainty over the job-seeker they meet.

In contrast, $\mathcal{P}$ in the FI model actually rises by close to 1% on impact. This counterfactual rise in $\mathcal{P}$ stems from the improvement in the average quality of job-seekers. Just as in the RI model, a fall in aggregate productivity lowers the payoffs from hiring a worker, $\mathbf{x}(a, z, e)$, and is a force towards depressing $\mathcal{P}$. Improvements in the average quality of job-seekers instead cause $\mathcal{P}$ to rise. Unlike the RI model, FI firms are not informationally constrained and do not worry about mistakenly hiring an unsuitable worker. In other words, the shadow value of information, λ, remains at zero. As such, higher λ is not a force towards driving lower $\mathcal{P}$ in the FI model as it is in the RI model. When aggregate productivity falls by 5%, FI firms' higher retention standards causes newly separated workers at the time of the shock to be comprised of higher z types. The shift in the distribution of job-seekers towards these higher z type workers overwhelms the fall in a, causing FI firms to accept job-seekers more often on impact.

In the same vein, despite the fall in aggregate productivity, the higher average quality of job-seekers makes it more attractive for FI firms to post a vacancy. This can be seen in the the bottom right panel of Figure 7, where q, the rate at which firms contact applicants, declines on impact in the FI model due to the increased number of vacancies posted. Consequently, the bottom middle panel of Figure 7 shows the unemployment rate increasing by less in the FI model, rising only 20% on impact compared to the 35% increase observed in the RI model.[36]

After the initial shock, as FI firms continue to correctly identify and hire higher productivity workers out of the pool of job-seekers, the higher average quality of job-seekers decays rapidly. In periods following the shock, $\mathcal{P}$ falls below steady state as the rise in the average quality of job-seekers diminishes and fails to fully counteract the fall in aggregate productivity. The fall in $\mathcal{P}$ is, however, short-lived as higher average quality of job-seekers together with recovering aggregate productivity leads FI firms to raise their average acceptance probabilities over time, causing $\mathcal{P}$ to rebound quickly.

In contrast, $\mathcal{P}$ continues to decline and is slow to recover in the RI model - the impulse response is *hump*-shaped. The elevated shadow value of information, λ, compounds the effect of lower a and mitigates the effect of higher average quality of job-seekers, causing acceptance rates to remain persistently low. These persistently lower acceptance rates reduce the rate at which high z types leave the pool of job-seekers, causing the high average quality in the pool of job-seekers to dissipate only gradually. Overall, $\mathcal{P}$ in the RI model remains below its steady state level for about 30 months following the shock. Moreover, persistently low $\mathcal{P}$ in the RI model in turn drives persistently high unemployment rates. The unemployment rate falls to half its initial increase in the FI model by the third month, while the increase in the unemployment rate in the RI model falls to half its value by the ninth month.

[36] This feature of the FI model reaffirms the findings of Mueller (2015) who argues that the higher average quality of job-seekers in recessions can further strengthen the Shimer puzzle.

5.1 The role of $H(G)$ in keeping $\mathcal{P}$ depressed

In our model, there are two forces that affect matching efficiency. Both aggregate productivity, a, and the distribution of job-seekers affect the extent to which a firm is informationally constrained and this in turn affects the firm's average probability of accepting a worker. Both these forces were at play in our exposition above but in order to isolate the role of higher uncertainty, we study the case where aggregate productivity a falls and remains 5% below steady state for 6 consecutive months. In this experiment, while a is fixed over these 6 months, the composition of job-seekers continues to change and hence any further change in λ for these 6 months solely reflects how changes in the distribution of job-seekers affect the extent to which firms are informationally constrained.

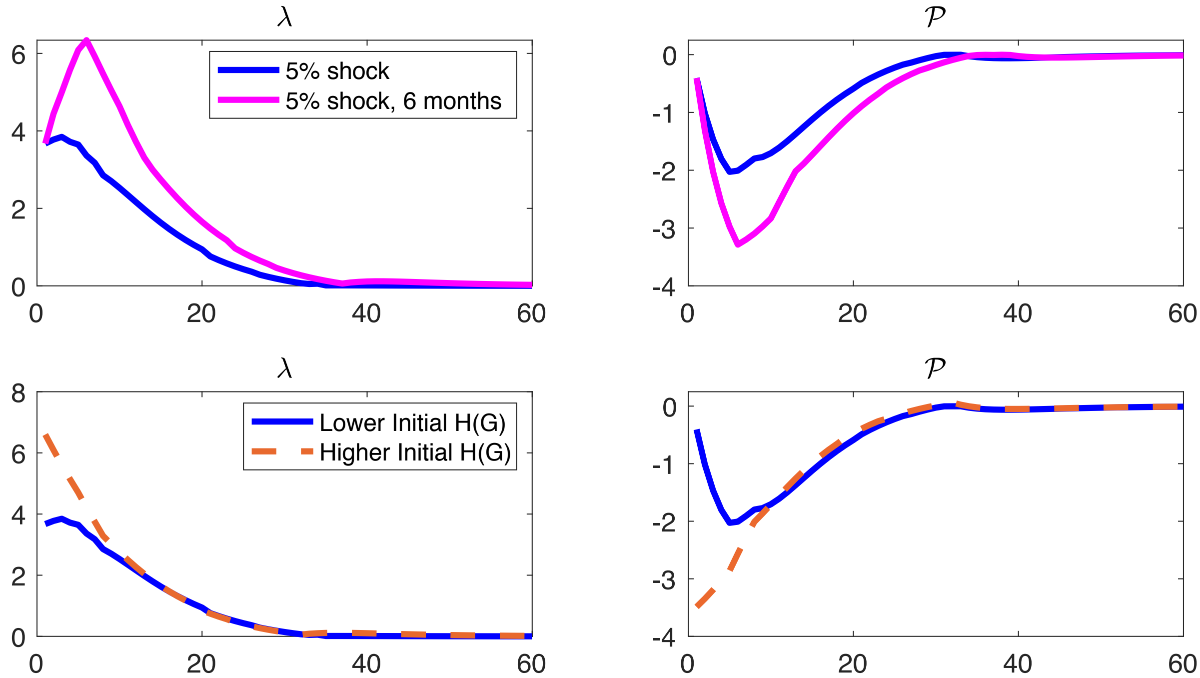


Figure 8: Impact of distribution of job-seekers

The top left panel of Figure 8 compares the response of λ when a is held fixed at 5% below steady state for 6 consecutive periods (pink dashed line) against the benchmark case where a falls on impact but recovers from the second period onwards (blue solid line). During these 6 months, the firm's initial uncertainty, as measured by $H(\mathcal{G})$, regarding the job-seeker it meets rises by 1%. At the time of the shock, the initial spike in hiring and firms' lower acceptance rates cause the distribution of job-seekers to become more varied.[37] As described in Section 3, increased initial uncertainty due to a change in distribution causes information constraints to tighten, making it harder for firms to further distinguish between applicants and causing λ to rise an additional 2% above its steady state level during the 6 months and $\mathcal{P}$ in the top right panel of Figure 8 to fall an additional 1%.

[37] In terms of the nomenclature used in the static example in Section 3, this corresponds to the distribution becoming more 'uniform'.

Separately, we consider an alternate exercise where firms have higher initial uncertainty over the applicant they meet i.e. higher $H(\mathcal{G})$ relative to steady state.[38] We then simulate a 5% fall in aggregate productivity below steady state that recovers over time. The bottom left panel of Figure 8 shows that when $H(\mathcal{G})$ is higher, firms are more informationally constrained and λ rises by more (orange dashed line) in response to the same sized fall in a. Consequently, firms reject workers more often as demonstrated by the larger fall in $\mathcal{P}$ in the bottom right panel of Figure 8. This highlights the state dependence in hiring decisions and hence the behavior of measured matching efficiency.

5.2 The role of average quality as a counteracting force

The above discussion made clear that a fall in aggregate productivity triggers a change in the composition of the pool of job-seekers. Importantly, how much measured matching efficiency and job creation declines depends on the extent an improvement in the average quality of job-seekers counteracts the fall in a. To highlight this, we next describe the response of the economy when the initial fall in aggregate productivity is larger and the induced change in average quality fails to compensate for this decline. Figure 9 shows the response of key labor market outcomes in the two models in response to a fall in aggregate productivity by 10%. In this case, average quality in the FI model actually falls slightly, as the set of previously employed low to middling z type workers who had drawn high match quality e now swamp the high productivity z workers who drew low match quality e.[39]

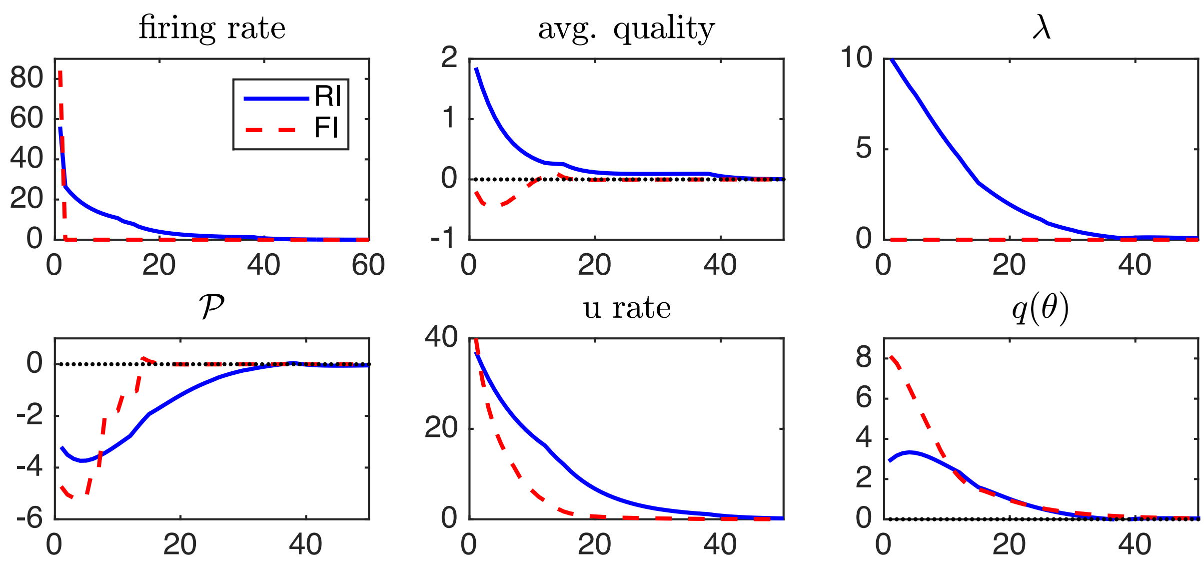


Figure 9: Response to a 10% shock

The larger decline in a magnifies the losses a firm incurs if it hires an unsuitable worker.

[38]In this exercise, we impose that the distribution of job-seekers is exactly equal to that observed at the end of six months in the example above.

[39]Note that since firms' hiring strategies are different, the FI and RI model do not necessarily share the same steady state distribution of workers.

Because losses from hiring an unsuitable applicant are even larger, RI firms want to be even more selective in their hiring decisions, causing informational constraints to tighten further as reflected by the 10% increase in λ. Larger losses from hiring unsuitable workers and the higher rise in λ cause RI firms to reject applicants more often, causing $\mathcal{P}$ to fall 3.5% on impact and reaching a low of 4% below steady state.

$\mathcal{P}$ in the FI model also *falls* 5% on impact under the large shock as the average quality of job-seekers in this case actually declines. In contrast to the response of the FI model under a smaller shock, the change in average quality of job-seekers fails to mitigate the fall in aggregate productivity. Initially, the decline in average quality compounds the effect of a lower a. Consequently, firms in the FI model reject workers more often, accounting for the larger initial decline in $\mathcal{P}$ observed in the FI model. However, as the decline in the average quality of job-seekers dissipates, acceptance rates in the FI model rapidly return to their steady state level as firms are willing to hire workers again when both average quality and aggregate productivity improve. While $\mathcal{P}$ in the FI model returns back to steady state after a year, $\mathcal{P}$ in the RI model remains 2% lower. The persistence in $\mathcal{P}$ in turn affects how quickly unemployment rates recover in both models. While the unemployment rate initially rises to the same extent in both models, the FI model observe the unemployment rate returning to its steady state level by the second year. In contrast, unemployment rates in the RI model remain elevated and only return to steady state 40 months after the shock.

Overall, our exercises suggests that the FI model has difficulty rationalizing the behavior of both measured matching efficiency and unemployment rates when the change in average quality mitigates the decline in aggregate productivity. The FI model correctly predicts a decline in $\mathcal{P}$ whenever average quality declines with aggregate productivity but these declines are short-lived as firms' acceptance rates rebound whenever both quality of the unemployed and aggregate productivity recover. In contrast, the RI model observes measured matching efficiency decline both on impact and through the course of the recession.

With a better understanding of the dynamic response in the RI and FI models, we proceed to test which model is better equipped to describe the joint behavior unemployment rates and measured matching efficiency in the data.

6 Model vs. data

Recall that the goal in this paper has been to explore the hypothesis that the same forces which affect the unemployment rate over the business cycle also affect the level of measured matching efficiency. This necessitates the study of joint dynamics of measured matching efficiency and unemployment in the data and in the model.

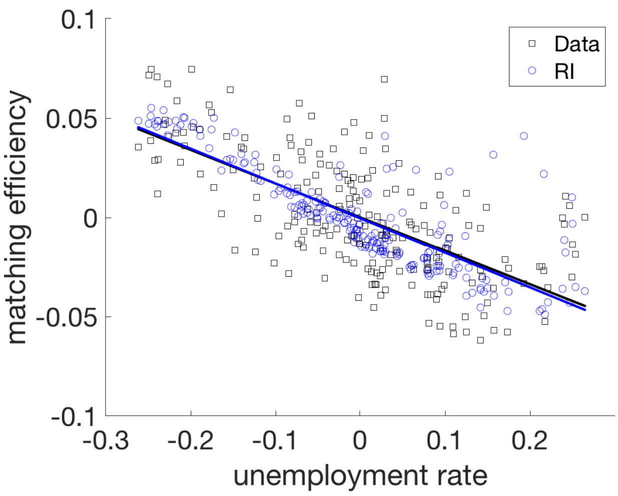
We first construct the empirically observed series of matching efficiency. Using the same matching function as in our model, we compute matching efficiency ξ_t as:

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

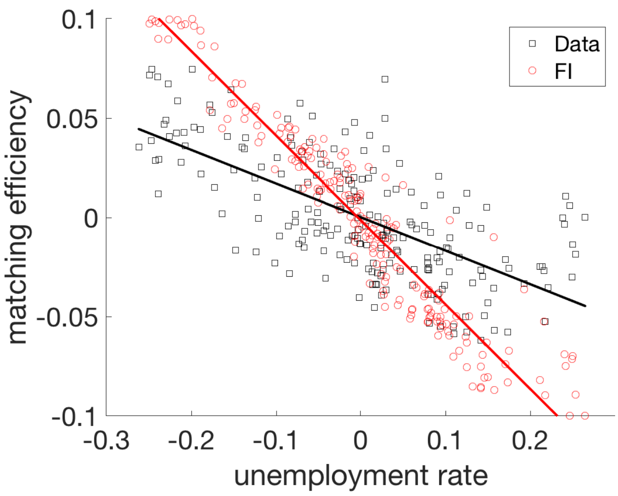
where we use data on total non-farm hires from JOLTS as our measure of matches, m, and data on the total non-farm job postings and total unemployed for our measures of v and u respectively. As per our calibration, we set $\iota = 0.5$.

We start by comparing the joint distribution of the unemployment rate and measured matching efficiency in the data with the joint distribution in the RI and FI models. To facilitate this comparison, we first construct a time series for the level of aggregate productivity a which allows the RI model to generate the path of the unemployment rate observed in the data from 2000m12 to 2016m12.[40] We repeat the exercise for the FI model. Since our framework is suited to evaluate cyclical fluctuations and not changes in trends, we follow Fujita and Moscarini (2013) and use the HP-filtered unemployment rate over this period.[41]

Next, feeding the respective series of aggregate productivity shocks into the calibrated RI and FI models, we compute the implied measured matching efficiency in both models. It is important to remember that while both models match the observed unemployment rate series by construction, the path of matching efficiency is not targeted.



(a) RI vs data



(b) FI vs data

Figure 10: Model vs data: Joint distribution - unemp. rate and matching efficiency.

In the data, the unemployment rate and matching efficiency share a negative relationship.

[40]We construct the time series of aggregate productivity using the implementation of the particle filter in Fernandez-Villaverde et al. (2016). We assume that the unemployment rate takes the form of $u_t^{data} = u_t^{model} + e_t$ where we treat $e_t \sim N(0, \sigma_e^2)$ as measurement error. Conditioning on the unemployment rate series, 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.

[41]Recent work by Hyatt and Spletzer (2013) and Davis and Haltiwanger (2014) show that job creation, job destruction and worker reallocation rates have trended downward over time. As these low-frequency movements in the flows of jobs and workers can impact the unemployment rate, we filter the data to extract only its cyclical component.

As Figure 10 shows, both the FI and RI model also exhibit this negative relationship. Figure 10a shows that the joint distribution of matching efficiency and the unemployment rate generated by the RI model strongly resembles the joint distribution in the data.[42] In fact, Figure 10a shows that the best linear predictor of matching efficiency given unemployment rates generated by the RI model is coincident with its empirical counterpart. Importantly, this feature of the model was not targeted in our exercise and thus shows how well the RI model explains the data. Formally, we further validate the RI model's ability in explaining the data by using the Kolmogorov-Smirnov test to test the hypothesis that measured matching efficiency generated by the RI model and measured matching efficiency in data are drawn from the same distribution. We cannot reject this hypothesis - the p-value is 0.48 implying that the null hypothesis cannot be rejected.

In contrast, the FI model performs noticeably worse at matching the joint distribution of unemployment rates and matching efficiency as in the data. As can be seen in Figure 10b, the FI model, relative to data, tends to systematically over-predict the level of matching efficiency when the unemployment rate is low and also predicts a much lower level of matching efficiency when the unemployment rate is high. Formally, the Kolmogrov-Smirnov test rejects the hypothesis (at 1% significance) that matching efficiency in the FI model is drawn from the same distribution as in the data.[43]

To understand why the FI model underperforms in fitting the joint distribution, our next exercise focuses on the Great Recession period and the protracted recovery in unemployment rates and matching efficiency that followed thereafter. In examining each model's ability to match the decline in matching efficiency, we take January 2008 as our starting point and calculate the log difference in unemployment rates and matching efficiency from its January 2008 value. Figure 11 plots the evolution of the labor market variables of interest. As discussed in Section 5, negative shocks which drive an improvement in the average quality of the unemployed lead matching efficiency in the FI model to increase, weakening the desire to curb job creation and causing unemployment rates to rise by less. Thus in order to rationalize such high unemployment rates, aggregate productivity needs to fall even more in the FI model. This can be seen in in Figure 12b which shows how average quality improved in the FI model over the Great Recession. Consequently, Figure 12a shows that the FI model requires a drop in aggregate productivity by close to 23% (relative to its pre-recession level) to generate the same rise in unemployment rates (as shown in Figure 11) as in the data. This severe decline in aggregate productivity in turn causes matching efficiency in the FI model to decline by 20%, almost twice the size of the fall in its empirical counterpart. In summary, the FI model cannot replicate the joint behavior of matching efficiency and unemployment

[42] While our model implied matching efficiency is somewhat less volatile than in the data, it should be noted that we focus exclusively on fluctuations in the matching efficiency due to changes in firm's hiring decisions. To the extent that matching efficiency in the data is also affected by workers' search intensity (see for example Hornstein and Kudlyak (2016) and Hall and Schulhofer-Wohl (2015)), our model abstracts from this margin and hence can only partially account for the volatility present in the data.

[43] The p-value for the full information model is given by 6.5e-5.

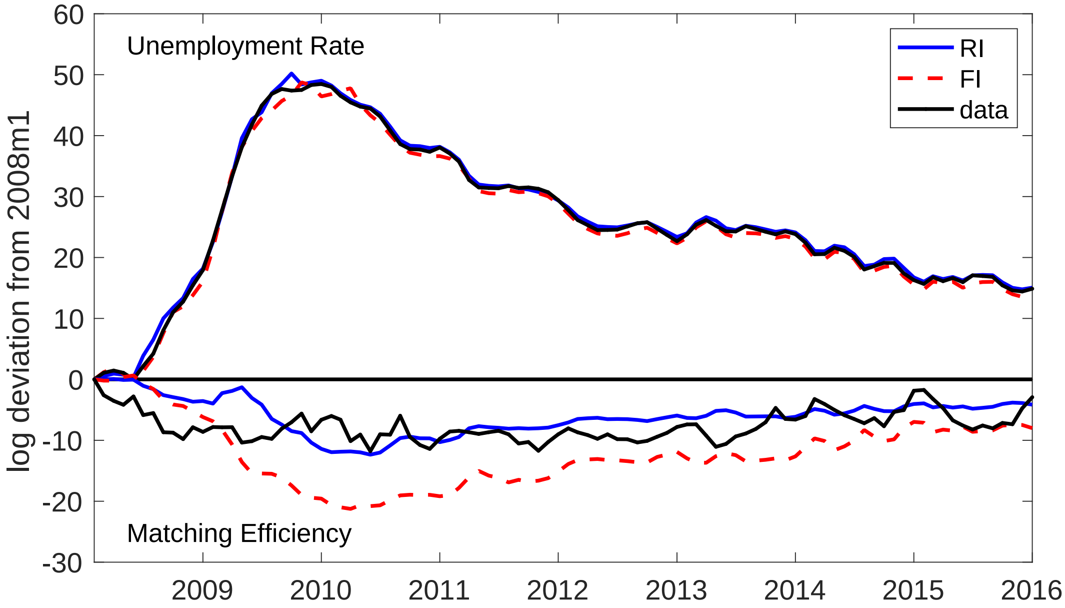


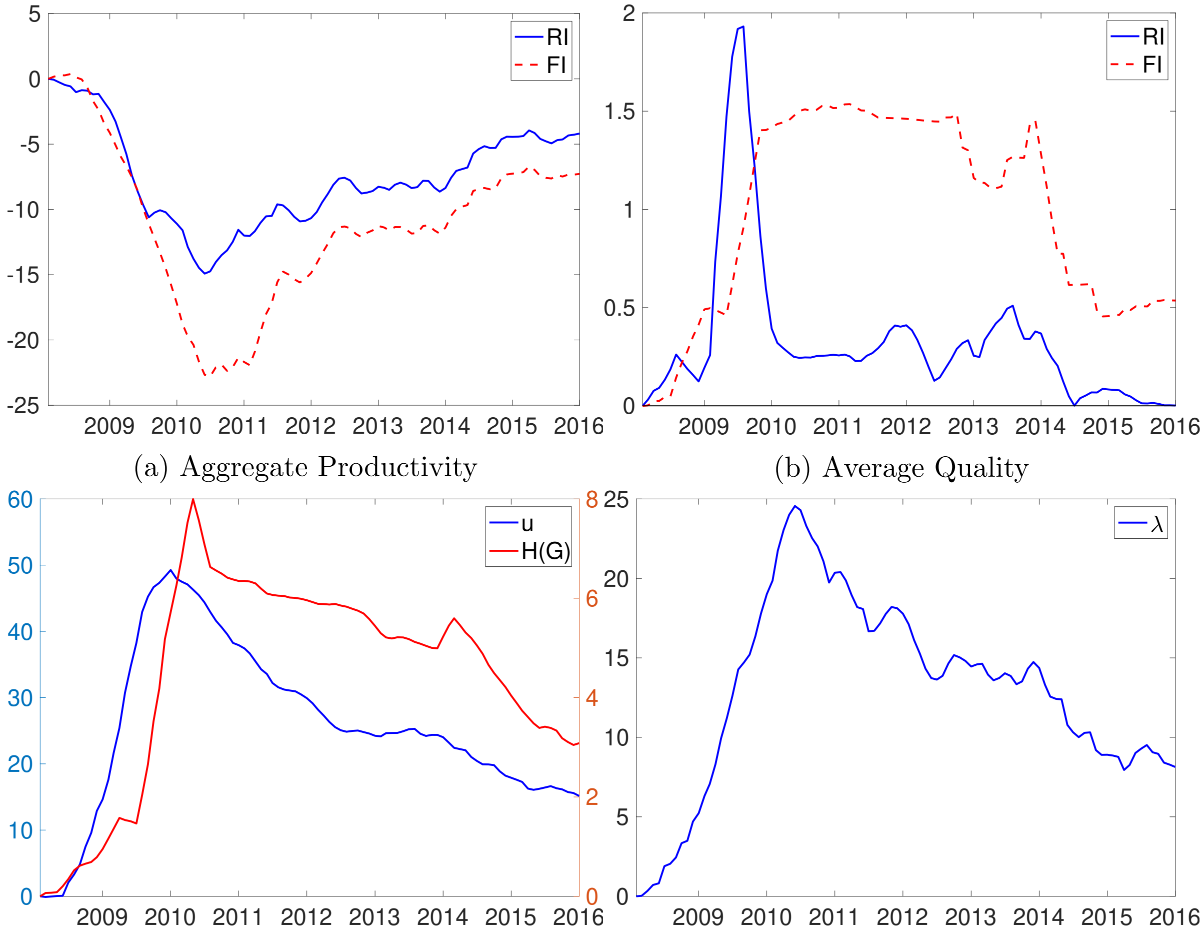
Figure 11: Unemployment rate and matching efficiency during the Great Recession

Notes: (i) The figure plots the unemployment rate and matching efficiency in the data (solid black curve), the RI model (the solid blue curves) and the FI model (the dashed red curves). (ii) All data series are HP filtered and the figure plots the (log) change relative to their respective levels in 2008m1. All series are plotted as 3-month moving average.

rates.

Notably, over the same Great Recession period, the RI model does a good job of matching the joint behavior of both unemployment rates and matching efficiency. While average quality also rises in the RI model, the initial uncertainty that RI firms have over the job-seeker that they meet, $H(\mathcal{G})$, is also rising over the same period. Figure 12c shows that initial uncertainty actually increases with a lag relative to the spike in unemployment rates, reflecting that more indiscriminate rejection of job-seekers today causes firms in the future to have more uncertainty about the type of job-seeker they meet. As previously explained, this higher $H(\mathcal{G})$, coupled with lower aggregate productivity raises the firm's desire to acquire more information to guide its hiring decisions, causing the shadow value of information, λ, to rise as in Figure 12d. The higher shadow value of information in turn lowers the acceptance rates, $\mathcal{P}_t$, and hence measured matching efficiency in the RI model falls without requiring as large a fall in aggregate productivity as that observed in the FI model. Overall, these exercises show how the RI model can quantitatively replicate the joint behavior of the unemployment rate and matching efficiency observed in the data.

.



(a) Aggregate Productivity (b) Average Quality

(c) Initial Uncertainty (d) Shadow value of information

Notes: (i) The top left panel plots the percent decline aggregate productivity a consistent with the observed unemployment rate in the RI model (solid blue curve) and FI models (dashed red curve). (ii) The top right panel plots the percentage increase in the model generated average quality of job-seekers for the RI (solid blue curve) and FI (dashed red curve) models (iii) The bottom left panel plots the evolution of $H(G)$ (red curve,right axis) in the RI model against the evolution of the unemployment rate (blue curve, left axis) (iv) the bottom right panel plots the evolution of λ in the RI model (v) All data series plot the (log) change relative to their respective levels in 2008m1. All series are plotted as 3-month moving average.

Figure 12: Labor Market Dynamics during the Great Recession

7 Discussion

7.1 Alternative cost structures of processing information

Fixed costs 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.[44] 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 matching efficiency behaves as in the full information model. In fact, fixed

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

costs of acquiring information, when incurred, 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. This may rationalize lower meeting rates, since fewer vacancies are posted, but not why there is a larger wedge between meeting and hiring rates. Thus, much like the FI model, such a cost structure would be unable to explain the large changes in matching efficiency.[45]

Noisy Information 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 not optimal. Signals of this form are only optimal if agents have a quadratic objective which is not the case in our model.[46] Restricting the information structure to be normal imposes additional costs on firms' processing capacity and prevents firms from designing more cost-effective information strategies.

To show how restricting the information structure to be that of Gaussian signals can lead to suboptimal signals, we compute the error probabilities that arise in a noisy information model relative to that observed in the RI model. In our exercise, the variance of the noise is set so that the amount of information processed (mutual information) in the noisy information model is exactly equal to the fixed channel capacity, χ, in the RI model. We assume all other parameters are the same as in the RI model. Appendix G details the noisy information model we compute.

Table 2: Error Probabilties in RI vs. Noisy Information Model

Error Probabilities		
Relative to RI	Type 1	Type 2
Error probability (ratio)	1.11	1.01
Losses from hiring	18%	11%

[45]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.

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

Table 2 summarizes the error probabilities and output losses relative to that achieved in the RI model. To make our comparisons fair, we impose that both the noisy information model and the RI model face the same distribution of job-seekers, and examine the average acceptance probabilities that arise under the two models. In general, firms in the noisy information model are more likely to commit mistakes in hiring. Firms in the noisy information model are 11% more likely to make a Type 1 error and 1% more to make a Type 2 error. While these differences may not seem very large, the losses they generate are substantial. Noisy information firms observe close to a 20% lower profit due to hiring the wrong worker (Type 1 error), as well as 11% lower profits from failing to hire suitable workers (Type 2 error).

Moreover, the choice of rational inattention over simply assuming a linear Gaussian signal structure can be justified by appealing to the Lucas critique. Under RI, firms endogenously change their information strategy based on the state of the economy and this has important implications for the behavior of matching efficiency. An exogenously specified signal structure would miss out on these endogenous mechanisms.[47]

7.2 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.

The other important assumption about wage setting is that in the model, workers are always paid b. We make this assumption because in our setting with asymmetric information, the standard Nash bargaining protocol is not well defined and having firms make take-it-or-leave-it offers avoids the complications associated with bargaining under asymmetric information.

[47] Another point in favor of modeling information constraints as rational inattention is that it is much easier and more elegant than the linear Gaussian setup in the context of our model since our model is not linear. Contrary to the common perception, solving and interpreting our model under rational inattention is easier than under noisy information.

A natural question that arises is whether the forces we described in our model that affect measured matching efficiency would still be at play if workers received a larger share of the surplus as opposed to just their outside option. Suppose that this higher share of surplus to workers is delivered via a higher constant wage $w > b$. This higher wage lowers the firms' lifetime profits $\mathbf{x}(a, \cdot, \cdot)$ which would cause them to lower the average probability of accepting a worker. While this would lower the average matching efficiency, the rest of the dynamics of the model would be qualitatively unchanged. Thus, simply raising the share of surplus accruing to the worker does not qualitatively change the results.

However, an important consequence of the assumption that workers have no bargaining power is that regardless of the level of aggregate productivity, wages are fixed at b. If wages fell when aggregate productivity declined, this would mitigate the effect of a negative productivity shock on $\mathbf{x}(a, \cdot, \cdot)$. A smaller fall in $\mathbf{x}$ in response to falls in a would lead firms to reduce their acceptance probabilities by less and mitigate the decline in $\mathcal{P}$ during a recession. However, there is also a new countervailing force which would tend to increase the responsiveness of $\mathcal{P}$ to falls in a. More generally, suppose that wages paid to all workers were given by by some function $\omega(a)$ where $\omega'(a) \in (0, 1)$. Since b is the lowest wage a worker is willing to accept, we also require that $\omega(a)$ must be at least as large as b in all states. Notice that $\omega > b$ in all states means that the firm's gain to matching $\mathbf{x}$ is lower. As noted by Hagedorn and Manovskii (2008), when the firm's gain to matching is small, even small declines in productivity can cause large percentage declines in its payoffs which in turn greatly reduce vacancy creation. In our model, a reduction in vacancy posting during a recession would raise the probability that all job-seekers regardless of their type remain unemployed for longer, causing the pool of job-seekers to be more varied and for the firm to have higher initial uncertainty regarding the type of job-seeker it meets. This in turn would cause the firm to face tighter informational constraints, increasing the fall in acceptance rates and measured matching efficiency during recessions. Thus, even when wages co-move with aggregate productivity, the change in the composition of job-seekers would still tighten informational constraints in recessions and act towards lowering matching efficiency.

7.3 The role of duration

Resume audit studies (See Kroft et al. (2013) and Eriksson and Rooth (2014) for example) have found evidence of firms using unemployment duration to filter applicants. To highlight how conditioning on τ substantially reduces the information burden on firms in terms of identifying the correct workers to hire, Figure 14a in Appendix H shows how the relative job-finding rates of unemployed applicants would differ if information on the unemployment duration of the worker were not freely available. Notably, the negative duration dependence associated with long-term unemployed workers weakens. Intuitively, when firms are unable to condition on τ and all applicants are pooled into one group, the firm has greater initial uncertainty over the job-seeker it meets, i.e. $H(G) > H(G \mid \tau)$. This greater initial uncertainty makes it more likely that the firm is informationally constrained and less able to distinguish between applicants with different (z, e). Since firms are less able to distinguish between workers, they are less likely to correctly identify and leave behind unsuitable applicants in the unemployment pool. As such, the pool of unemployed of each duration type is more varied and relative job-finding rates decline less rapidly with duration.[48]

While duration provides information about the worker's type, our model suggests that duration becomes a weaker signal of the individual's true (z, e) during a recession. As job creation falls during recessions, 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, long unemployment duration spells provides less information about an applicant's true (z, e) during a recession. Figure 14b in Appendix H shows how relative job-finding rates decline at a gentler rate during a recession.

7.4 Unemployment Volatility

While not the focus of our paper, a natural question that arises is whether fluctuations in matching efficiency can better explain the amount of unemployment volatility observed in the data. As seen from Figure 12a, our model still relies on relatively large TFP shocks to replicate the rise in unemployment rate during the Great Recession. More generally, Table 3 compares the volatility of key labor market variables in the data with their corresponding values from our model and from Shimer (2005). While our model gives rise to greater volatility in unemployment rates (first column) and job-finding rates (last column) relative to Shimer (2005), our model still cannot replicate the volatility in unemployment observed in the data. Unlike Shimer (2005), the more subdued levels of volatility in our model stem from the fact that the composition of job-seekers improves during a recession when firms fail to identify and hire suitable workers. This improvement partially mitigates the firms' desire to create fewer vacancies in a downturn.

Table 3: Volatility of Labor Market Variables

Standard Deviations	u	v	$\mathcal{P} \times p$
Data (from Shimer (2005))	0.190	0.202	0.118
RI Model	0.070	0.042	0.020
Shimer (2005)	0.009	0.027	0.010

[48]In the knife edge case where the firm is completely unable to distinguish between workers and applies the same hiring probabilities $\gamma(z, e \mid \sigma) = \gamma(\sigma)$ for all (z, e) types, the relative job-finding rate becomes flat and equal to 1 for all duration spells.

8 Conclusion

We present a novel channel through which firms' hiring standards affect fluctuations in measured matching efficiency. The key insight is the presence of a tight link between matching 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 matching efficiency. Further, lack of adequate screening by firms in filtering out suitable applicants today adds to the uncertainty firms tomorrow face regarding the type of job-seeker they meet, 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 matching efficiency to remain persistently low.

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.

Bowlus, Audra J, "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality," *Journal of Labor Economics*, April 1995, *13* (2), 335–350.

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 and John Haltiwanger, "Labor Market Fluidity and Economic Performance," Working Paper 20479, National Bureau of Economic Research September 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.

Doppelt, Ross, "The Hazards of Unemployment: A Macroeconomic Model of Job Search and Resume Dynamics," Technical Report 2016.

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.

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

Gavazza, Alessandro, Simon Mongey, and Giovanni L. Violante, "Aggregate Recruiting Intensity," Staff Report 553, Federal Reserve Bank of Minneapolis August 2017.

Hagedorn, Marcus and Iourii Manovskii, "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, September 2008, *98* (4), 1692–1706.

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.

Hyatt, Henry R. and James Spletzer, "The Recent Decline in Employment Dynamics," Working Papers 13-03, Center for Economic Studies, U.S. Census Bureau March 2013.

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.

Lise, Jeremy and Jean-Marc Robin, "The Macrodynamics of Sorting between Workers and Firms," *American Economic Review*, April 2017, *107* (4), 1104–1135.

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.

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.

Appendix

A Proof of Proposition 1

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:

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

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:

$$
\gamma(z,e \mid \sigma) = \frac{\mathcal{P}(\sigma)e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma)}}}{1-\mathcal{P}(\sigma)\left[1-e^{\frac{\mathbf{x}(a,z,e)}{\lambda(\sigma)}}\right]} \tag{20}
$$

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

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

B Full Information Hiring Decisions

For $\chi \to \infty$, the shadow value of an additional nat, $\lambda = 0$. Evaluating (8) for (z, e) combinations such that $\mathbf{x}(a, z, e) < 0$ in the limit as $\lambda \to 0$ yields:

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

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:

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

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

C Proof of Lemma 2

Let $\gamma_H(a)$ and $\gamma_L(a)$ denote the optimal choices when aggregate productivity is given by a. Differentiating (18) with respect to aggregate productivity a yields:[49]

$$[\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_H)]\, \alpha \gamma_H'(a) + [\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_H)]\,(1 - \alpha)\, \gamma_H'(a) = 0 \tag{22}$$

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:

$$\mathbf{x}(a, z_i) = \lambda(a)\,[\mathcal{H}'(\mathcal{P}) - \mathcal{H}'(\gamma_i(a))]\ ,\ i \in \{H, L\} \tag{23}$$

where $\lambda(a)$ is the multiplier on the constraint (18). Using this, rewrite equation (22) as:

$$\frac{\gamma_H'(a)}{\gamma_L'(a)} = -\frac{(1 - \alpha)\, \mathbf{x}(z_L, a)}{\alpha \mathbf{x}(z_H, a)} \tag{24}$$

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, (24) 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 (19). For $a \in [{}^b/_{z_H}, {}^b/_{z_L}]$, the LHS of (19) 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,

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

$\gamma_H(a)$ 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.

Online Appendix (not for publication)

D Proof of Lemma 1

This proof is similar to Appendix A in Matejka and McKay (2015). The main purpose of this appendix is to specialize it to our particular environment. For ease of exposition, we re-write equation (4) and suppress the dependence on σ and τ.

$$\mathbf{V} = \max_{\Gamma \in \Delta} \sum_z \sum_e \int_s J\left[\Gamma(z, e \mid \mathbf{s})\right] d\Gamma(\mathbf{s} \mid z, e) g(z, e)$$

Conditional on observing a signal $s \in \mathcal{S}$, the firm chooses to hire, implying that its payoff is $\mathbf{x}(a, z, e)$, which is constant for all $s \in \mathcal{S}$. Hence, we can re-write the above as:

$$\begin{aligned}
\mathbf{V} &= \max_{\Gamma \in \Delta} \sum_z \sum_e \int_s J\left[\Gamma(z, e \mid \mathbf{s})\right] d\Gamma(\mathbf{s} \mid z, e) g(z, e) \\
&= \max_{\Gamma \in \Delta} \sum_z \sum_e \gamma(z, e) \mathbf{x}(z, e) g(z, e)
\end{aligned}$$

where we have suppressed the dependence on a and used the definition $\gamma(z, e \mid \sigma, \tau) = \int_{s \in \mathcal{S}} d\Gamma(s \mid z, e, \sigma, \tau)$. The above is equivalent to 6. Next, we move to the constraint. The information processing constraint in its original form is given by (2):

$$H\left(\mathcal{G}(z, e)\right) - \mathbb{E}_s H\left(\Gamma(z, e \mid s)\right) \leq \chi$$

Using the fact that mutual information is symmetric, we know that the LHS of the above inequality can be written as:

$$H\left(\mathcal{G}(z, e)\right) - \mathbb{E}_s H\left(\Gamma(z, e \mid s)\right) = H\left(\Gamma_s(s)\right) - \mathbb{E}_{z,e} H(\Gamma(s \mid z, e)) \leq \chi \tag{25}$$

where $\Gamma_s(s)$ denotes the unconditional distribution of signals while $\Gamma(s \mid z, e)$ denotes the distribution of signals conditional on a particular (z, e). Furthermore, $\mathbb{E}_{z,e}[\cdot]$ denotes the expectation w.r.t. the random variables z and e. To show that the expression above is equivalent to (7), we start by manipulating the first term $H\left(\Gamma_s(s)\right)$:

$$\begin{aligned}
H\left(\Gamma_s(s)\right) &= -\left(\int_{s \in \mathcal{S}} d\Gamma_s(s)\right) \ln\left(\int_{s \in \mathcal{S}} d\Gamma_s(s)\right) - \left(\int_{s \notin \mathcal{S}} d\Gamma_s(s)\right) \ln\left(\int_{s \notin \mathcal{S}} d\Gamma_s(s)\right) \\
&= -\left[\sum_z \sum_e \int_{s \in \mathcal{S}} d\Gamma(s \mid z, e)\, g(z, e)\right] \ln\left[\sum_z \sum_e \int_{s \in \mathcal{S}} d\Gamma(s \mid z, e)\, g(z, e)\right] \\
&\quad - \left[\sum_z \sum_e \int_{s \notin \mathcal{S}} d\Gamma(s \mid z, e)\, g(z, e)\right] \ln\left[\sum_z \sum_e \int_{s \notin \mathcal{S}} d\Gamma(s \mid z, e)\, g(z, e)\right]
\end{aligned}$$

where we have used the fact that $\int_{s\in\mathcal{S}} d\Gamma_s(s) = \sum_z \sum_e \int_{s\in\mathcal{S}} d\Gamma(s \mid z,e) g(z,e)$. Next, using the definition $\gamma(z,e \mid \sigma,\tau) = \int_{s\in\mathcal{S}} d\Gamma(s \mid z,e,\sigma,\tau)$ we can rewrite the above expression as:

$$
\begin{aligned}
H\left(\Gamma_s(s)\right) &= -\left[\sum_z \sum_e \gamma\left(z,e\right) g\left(z,e\right)\right] \ln \left[\sum_z \sum_e \gamma\left(z,e\right) g\left(z,e\right)\right] \\
&\quad -\left[\sum_z \sum_e \left(1-\gamma\left(z,e\right)\right) g\left(z,e\right)\right] \ln \left[\sum_z \sum_e \left(1-\gamma\left(z,e\right)\right) g\left(z,e\right)\right] \\
&= -\mathcal{P} \ln \mathcal{P} - (1-\mathcal{P}) \ln(1-\mathcal{P}) \\
&= H(\mathcal{P}) \qquad (26)
\end{aligned}
$$

where the final equality follows from the definition of $\mathcal{P}(\sigma,\tau)$.

Next, we manipulate the second term $\mathbb{E}_{z,e} H\left(\Gamma(s \mid z,e)\right)$:

$$
\begin{aligned}
-\mathbb{E}_{z,e} H\left(\Gamma(\cdot \mid z,e)\right) &= -\sum_z \sum_e H\left(\Gamma\left(s \mid z,e\right)\right) g\left(z,e\right) \\
&= \sum_z \sum_e \left(\int_{s\in\mathcal{S}} d\Gamma\left(s \mid z,e\right)\right) \ln \left(\int_{s\in\mathcal{S}} d\Gamma\left(s \mid z,e\right)\right) g(z,e) \\
&\quad + \sum_z \sum_e \left(\int_{s\notin\mathcal{S}} d\Gamma\left(s \mid z,e\right)\right) \ln \left(\int_{s\notin\mathcal{S}} d\Gamma\left(s \mid z,e\right)\right) g\left(z,e\right) \\
&= \sum_z \sum_e \Big\{\gamma\left(z,e\right) \ln \gamma\left(z,e\right) + \left(1-\gamma(z,e)\right) \ln\left(1-\gamma(z,e)\right)\Big\} g\left(z,e\right) \\
&= -\sum_z \sum_e H\left(\gamma(z,e)\right) g\left(z,e\right) \qquad (27)
\end{aligned}
$$

Thus, from (25), (26) and (27), it is clear that the constraint (2) is equivalent to (7).

E Numerical Implementation

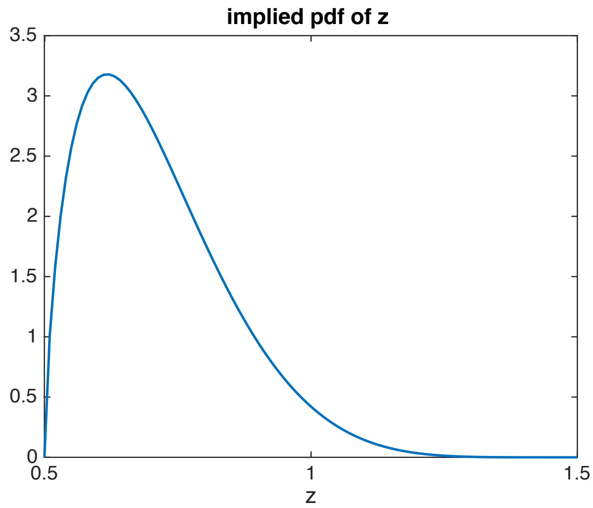
E.1 Constructing the sets $\mathcal{Z}$ and $\mathcal{E}$

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 $\mathcal{F}(\cdot; A, B)$.

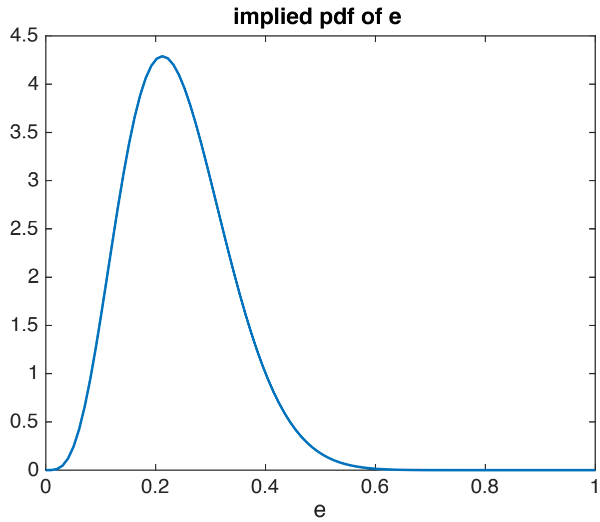
1. Construct a sequence $\{p_i\}_{i=1}^{n_z+1}$ where each $p_1 = 0$, $p_{n_z+1} = 1$ and for $i \in \{2, \cdots, n_z\}$ p_i is given by:

   $$p_i = i/n_z$$

   where n_z denotes the cardinality of the set $\mathcal{Z}$.



(a) z



(b) e

Figure 13: Implied probability distributions

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 $1/n_z$.

Implied Probability Distributions

Given our calibrated parameters for $\{A_z, B_z, A_e, B_e\}$, Figure 13a and 13b plot the implied probability density functions under these given parameters.

E.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 $0 \le \tau < \overline{\tau}$. 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:

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

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:

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

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.

F 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. Given the parameters governing the heterogeneity of workers, this leaves us with three parameters $\{b, \delta, \kappa\}$ to recalibrate for the full information model. We target the unemployment rate, exit probability and 70% UI ratio to recalibrate these three parameters. Table 4 details used in the full information model.

A fully re-calibrated FI model would make comparison difficult as the two economies would not have the same unconditional distribution of workers (the distribution of z) and the distribution of match quality e. This would change the surplus of a match and hence the hiring decisions. As a result it would be hard to identify if the differences in hiring decisions across the two models arose because of differences in information or because the differences in surplus induced by a different (z, e) distribution.

Table 4: Model Parameters for Full Information

Inferred Parameters		
Parameter	Description	Value
b	home production	0.1429
δ	exog. separation rate	0.0330
κ	vacancy posting cost	0.0850

On the other hand, if one were to keep all parameters in the FI model the same as in the RI model except that now $\chi = \infty$, then the FI model would not not have the same average unemployment rates or job-finding rates. This too complicates the comparison.

Thus, we choose to partially re-calibrate the FI model so that it has the same distribution of z and e as the RI model while also featuring the same average unemployment rate and job-finding rates.

G Noisy Information

G.1 Model

Consider the following noisy information model where upon meeting a job-seeker, firms obtain a signal of the form Firm gets a signal of form

$$s = ze + \varsigma\varepsilon$$

where ε is an iid draw from a normal $\mathcal{N}(0,1)$ distribution. Let ς be the parameter that governs the noisiness of the signal s. For a given aggregate state σ and unemployment duration τ of the job-seeker, the firm chooses to hire whenever he receives a signal $s \geq s^*(\sigma,\tau)$.

Then the firm's problem can then be re-written as choosing the cut-off for signal s so as to maximize ex-ante firm surplus.

$$\max_{s^*} \sum_z \sum_e \left[1 - \Phi\left(\frac{s^* - ze}{\varsigma}\right)\right] \mathbf{x}(a,z,e)\, g(z,e \mid \sigma,\tau)$$

Taking first order conditions with respect to s, we have:

$$\sum_z \sum_e \phi\left(\frac{s^* - ze}{\varsigma}\right) \mathbf{x}(a,z,e)\, g(z,e \mid \sigma,\tau) \begin{cases} \leq 0 & \text{if } s^* = \infty \\ = 0 & \text{if } s^* \in (-\infty,\infty)] \\ \geq 0 & \text{else} \end{cases}$$

where s^* is the cut-off for which the firm will always hire for signal values above this threshold. Observe that the probability of hiring a job-seeker of type (z, e, τ) in aggregate state σ is given by:

$$\gamma^{noisy}(z, e \mid \sigma, \tau) = 1 - \Phi\left(\frac{s^*(\sigma, \tau) - ze}{\varsigma}\right) \tag{28}$$

and the average probability of hiring a worker of duration τ is given by:

$$\mathcal{P}^{noisy}(\sigma, \tau) = \sum_z \sum_e \gamma^{noisy}(z, e \mid \sigma, \tau) g(z, e \mid \sigma, \tau) \tag{29}$$

Finally, match efficiency or the average probability of hiring any worker in state σ is given by:

$$\mathcal{P}^{noisy}(\sigma) = \sum_\tau g_\tau(\tau \mid \sigma) \mathcal{P}^{noisy}(\sigma, \tau)$$

G.2 Parameterizing ς

To figure out the noise in the signal, we use the acceptance probabilities from the model, i.e. equations (28) and (29), and set ς such that the amount of mutual information is equal to the fixed processing capacity χ in the rational inattention model in steady state. In other words, we choose ς such that the following equation is satisfied:

$$\chi = \mathcal{H}\left[\mathcal{P}^{noisy}(\sigma, \tau)\right] - \sum_z \sum_e H\left[\gamma^{noisy}(z, e \mid \sigma, \tau)\right] g^{noisy}(z, e \mid \sigma, \tau)$$

We keep all other parameters in the model the same as that for the RI model. Our calibration gives us $\varsigma = 0.1763$.

G.3 Computing Error Probabilities

To compute the error probabilities in the RI model vs. the noisy information model. We first calculate the probability that for a job-seeker of (z, e, τ) type, the firm incorrectly chooses to accept as opposed to reject the job-seeker. In other words, we calculate the probability of making a Type 1 error for a job-seeker of type (z, e, τ) as:

$$Pr(\text{Type 1 error} \mid z, e, \tau, \sigma) = \gamma^j(z, e \mid \sigma\tau) Pr(\gamma^{FI}(z, e \mid \sigma, \tau) == 0)$$

where $j \in \{\text{noisy}, \text{RI}\}$ indicates the model of interest. The average probability of making a Type 1 error is then given by:

$$Pr(\text{Type 1 error} \mid \sigma) = \sum_\tau \sum_z \sum_e Pr(\text{Type 1 error} \mid z, e, \tau, \sigma) g^j(z, e \mid \tau, \sigma) g^j_\tau(\tau)$$

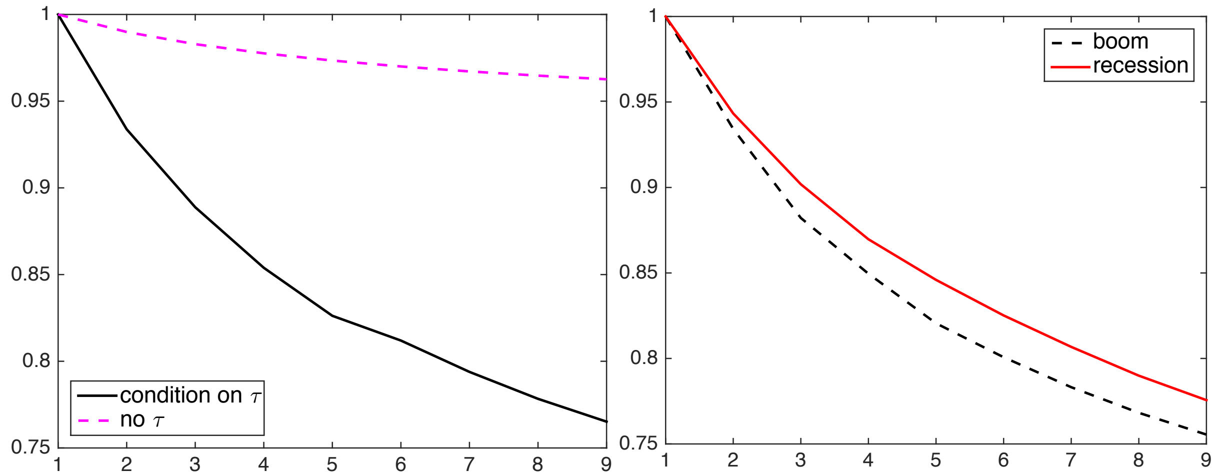
Similarly, we define the probability that a firm incorrectly chooses to reject a job-seeker of type (z, e, τ) that is suitable for hiring as the following:

$$Pr(\text{Type 2 error} \mid z, e, \tau, \sigma) = \left[1 - \gamma^j(z, e \mid \sigma\tau)\right] Pr(\gamma^{FI}(z, e \mid \sigma, \tau) == 1)$$

where $j \in \{\text{noisy}, \text{RI}\}$ again represents the model of interest. The average probability of making a Type 2 error is then given by:

$$Pr(\text{Type 2 error} \mid \sigma) = \sum_\tau \sum_z \sum_e Pr(\text{Type 2 error} \mid z, e, \tau, \sigma) g^j(z, e \mid \tau, \sigma) g^j_\tau(\tau)$$

H Duration



(a) Conditioning on τ

(b) Recessions vs. booms

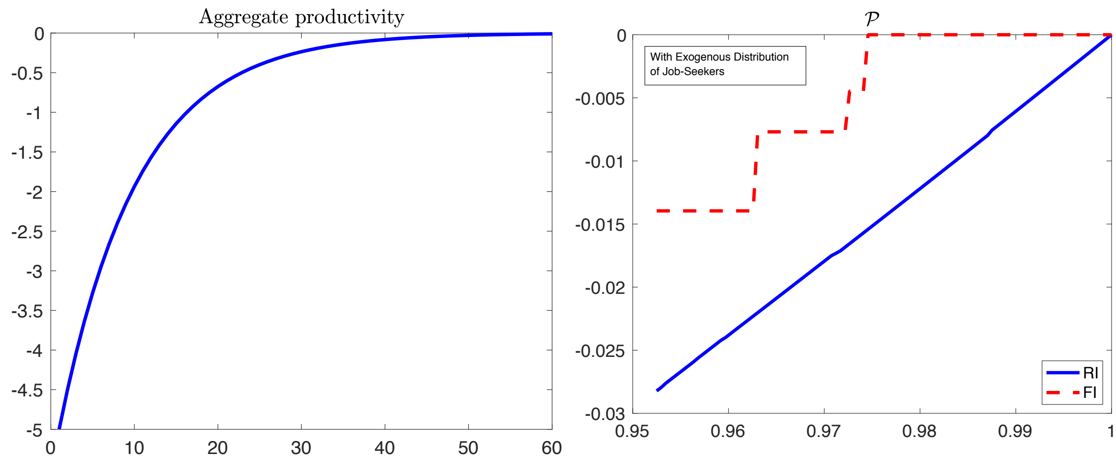
Notes: (i) The left panel plots the relative job-finding rates when firms can condition on duration (black solid line) and when firms cannot condition on τ (pink dashed line). (ii) The right panel plots the relative job-finding rates in a boom (aggregate productivity is 10% above steady state, black dashed line) vs a severe recession (aggregate productivity is 10% below steady state, red solid line).

Figure 14: Relative job-finding rates

I Additional Graphs

Recovery in aggregate productivity Figure 15a shows how quickly aggregate productivity recovers after a 5% shock.

Isolating the response of $\mathcal{P}$ with respect to a Because the shock to aggregate productivity causes the distribution of job-seekers to also change on impact, we conduct a separate exercise here to isolate the impact of productivity shocks on $\mathcal{P}$ when the distribution is held constant. In this exercise, we treat the distribution of job-seekers as exogenous and impose the steady state distribution of job-seekers. Holding constant the distribution of job-seekers, we show how $\mathcal{P}$ would change with respect to a in both the RI and FI models by plotting



(a) A 5% shock to TFP

(b) $\frac{\partial \mathcal{P}}{\partial a}$ when distribution fixed

the deviation of $\mathcal{P}$ from its value when aggregate productivity is equal to 1. Figure 15b shows how much lower measured matching efficiency would be in percentage terms relative to steady state (y-axis) with respect to different values of a (x-axis) when the exogenous distribution of job-seekers is given by our steady state distribution of job-seekers under the RI model. Note that without any change in the distribution of job-seekers, the FI model predicts that $\mathcal{P}$ would fall by about 1.5% with respect to a 5% productivity shock while the RI model observes a larger decline of about 3%. Contrasting Figure 15b with the decline in $\mathcal{P}$ (on impact) in Figure 7 makes clear how the changing distribution and overall higher average quality of job-seekers mitigates the decline in measured matching efficiency due to a fall in aggregate productivity.