Learning about Labor Markets

Jake Bradley[1]
Lukas Mann[2]

Abstract

We study a general equilibrium model of the labor market in which agents slowly learn about their suitability for jobs. Our model reproduces desirable features of the data, many of which standard models fail to replicate. We explore how, in such an environment, asymmetric information can lead to substantial misallocation. We calibrate our model to US data and quantify the welfare loss arising from misallocation due to informational frictions. The tractability of the model allows us to explore the responsiveness of wages and employment to an aggregate shock. We find that wage rigidity arises endogenously because of protracted learning, and in line with the data, the model is able to generate a larger and more persistent employment response.

Keywords: Learning, misallocation, labor markets, wage rigidity
JEL classification: D83, D61, J64

1. Introduction

Recent empirical literature has documented that workers' labor market environments often do not perfectly align with their beliefs about those environments. Survey studies have found evidence of biased beliefs about job finding rates that suggests that workers under-estimate high job finding rates and over-estimate low ones (see Mueller et al. (2021)). This suggests that workers frequently have incorrect beliefs when weighing their options in the labor market, such as deciding whether to switch locations or occupations or when assessing their outside option while negotiating with prospective employers. In this paper, we explore the allocative effects of such incorrect beliefs and quantify the welfare loss arising from missing information in labor markets. We do this by incorporating a simple informational friction into an otherwise standard labor market setting. Workers can choose among many markets but can only learn about a market's job finding rate by spending time in it. Workers are Bayesian, updating their beliefs and slowly building a complete picture of the market they inhabit. We show that, even though there is a large potential state space of individual labor market histories, workers in our model only need to keep track of two sufficient state variables. This sufficiently simplifies the model to allow us to embed our belief structure in a general equilibrium framework.

Our model can capture salient facts that models from the literature often fail to replicate. First, our model provides a natural interpretation for the finding in Mueller et al. (2021) that workers with high job finding rates underestimate their true job finding probabilities and, conversely, workers who have low job finding rates are overly optimistic. Second, since workers in

[1]University of Nottingham
[2]Princeton University

our model negatively adjust their beliefs during times without job offers, the model can account for much of the measured wage dependence on unemployment duration in the data. While the existing literature has often attributed this dependence to human capital depreciation, our results suggest that these papers possibly overstate the role of human capital depreciation in generating this dependence. Third, as is pointed out by Hornstein et al. (2011), without a substantial cost associated with unemployment, the prototypical search model is unable to generate the level of frictional wage dispersion present in the data. As it turns out, variation in worker's beliefs reconcile wage dispersion in a natural way. Fourth, following an aggregate shock to the economy, workers learn about the changing environment gradually. This creates sluggishness in the wage response and the model is able to generate wage rigidity as has been documented empirically.

Finally, we use our model to quantify the welfare loss arising from the informational friction and find that it is substantial. In comparison to a full information benchmark, welfare losses are approximately 34%. These losses arise as a result of missing information that affects workers in two ways. First, workers make ill-informed decisions because of missing information about their own market. Second, workers do not know the demand for their labor ahead of entering any other market. We find that this second margin is quantitatively important in generating welfare losses. It discourages workers from switching out of markets in which the demand for their labor is low. In equilibrium, workers thus congest markets they are unsuitable for which leads to lower aggregate job finding rates compared to a full information benchmark.

Related literature. In the model presented, workers learn about the state of their labor market through sequential sampling. The process of learning and search has been explored in previous work. The seminal paper is Rothschild (1974), where consumers update their beliefs over prices in a product market. In a labor market setting, Burdett and Vishwanath (1988) explore an environment in which workers update their beliefs regarding the distribution of wages that they sample from. As in our model, workers with worse realizations become more pessimistic and lower their reservation wage. Conlon et al. (2021) extend Burdett and Vishwanath's framework to include workers looking for jobs while in employment. The updating rule and parameters of the search model are carefully calibrated to U.S. data and the authors find that the social cost of information frictions is comparable to that of search frictions, which have dominated the macro-labor literature.

Our model also contributes to understanding the relationship between belief dispersion and wage rigidity. We show that slowly adjusting beliefs can be a major amplification mechanism in generating business cycle fluctuations. In a similar vein, Menzio (2022) considers a general equilibrium environment in which a share of workers' beliefs about the labor market evolves sluggishly. In contrast to Menzio (2022), agents in our model learn eventually about their true market conditions, but do so with a delay, triggering some of the same mechanisms emphasized in Menzio's work.

Unlike the models discussed, in our framework there is no uncertainty over the distribution of prices or wages, but rather over the frequency to which job offers arrive. Gonzalez and Shi (2010) consider a directed search model in which workers infer their type through the success or failure of the applications they make. In our setting, the uncertainty is not over a worker's type per se, but the labor demand of their services, measured as the frequency of job offers. In this sense, the most comparable paper to ours is Potter (2021). Relative to Potter (2021) and the preceding literature we make two novel theoretical contributions. First, the the model is derived in a general equilibrium setting. This allows us to make predictions over the welfare cost of uncertainty. Crucially, it also allows worker's information (or lack of) to be endogenous to the model, rather than imposed upon it. Second, we do not rely on agents optimizing under the "anticipated utility" framework of Kreps (1998). In our setting, an agent fully internalizes how their beliefs will change tomorrow when making decisions today.

Modern surveys have allowed researchers to assess whether worker's expectations are consequential for their realizations. Stephens (2004) finds that respondents reported beliefs over their opportunities are informative for future outcomes. Although workers beliefs have predictive power over realizations, as in our model, there exists systematic biases, see Spinnewijn (2015) and Mueller et al. (2021) for evidence from survey responses, Falk et al. (2006) from the laboratory, and Mueller and Spinnewijn (2022) for a comprehensive overview. In particular, Mueller et al. (2021) documents that unemployed workers with poor labor market prospects tend to overly optimistic, and the unemployed with better prospects tend to be overly pessimistic, a feature that the model naturally reconciles.

Outline. The rest of the paper is structured as follows. Section 2 presents motivating empirical evidence. Section 3 derives the baseline model. For the purpose of comparison, section 4 presents the full information case. Section 5 calibrates the model and section 6 presents the results and implications of the model. Section 7 examines the response of the economy to an aggregate shock and section 8 concludes.

2. Empirical Motivation

The model presented in the next section features workers updating their perception about their labor market prospects while in unemployment. There are several relatively recent labor market surveys that elicit expectations about job finding rates, wages, and income risk from respondents. For a comprehensive overview of the literature and related surveys see the handbook chapter Mueller and Spinnewijn (2022), and the references contained therein. To motivate the specific learning mechanism present in our framework we will rely on the Survey of Consumer Expectations (SCE), a data set that is maintained by the Federal Reserve Bank of New York and is a representative survey of approximately 1,300 individuals living across the United States. In the analysis presented, attention is restricted to unemployed workers between the ages of 20 and 65. The publicly available SCE data begins in June 2013 and our sample runs until February 2020. While there is the possibility of a longer time dimension, in order to avoid unwanted effects associated with COVID-19 restrictions we end our sample before the first implemented lockdown. Finally, since we are concerned with worker's beliefs, we further restrict the sample to ensure that respondents elicited beliefs are somewhat consistent. Full details of the restrictions imposed on the sample can be found in Appendix A.1.

The aim of this section is to present two pieces of supportive evidence for our mechanism. First, that the unemployed update their beliefs about their job finding probability within an unemployment spell, and do so in a way that is consistent with our model. Second, that after a change in labor market, an unemployed worker becomes more optimistic about their employment prospects. The increase in optimism, consistent with our model, is more pronounced following a less fruitful spell in unemployment.

Mueller et al. (2021) document that in the cross-section, elicited beliefs about job finding rates decline with the duration of unemployment. However, this decline is the result of dynamic

Table 1: Perceived 3-month job-finding probability and unemployment duration

	Full sample			Spells without job offers		
	(1)	(2)	(3)	(1)	(2)	(3)
Log unemployment duration months	-0.190*** (0.016)	-0.160*** (0.018)	-0.064 (0.051)	-0.174*** (0.030)	-0.144*** (0.031)	-0.149** (0.068)
Demographic controls		x			x	
Spell fixed effects			x			x
Observations	2,156	2,133	2,156	673	665	673
R^2	0.097	0.158	0.783	0.073	0.189	0.764

Note: The dependent variable in each case is the log of the elicited 3-month job-finding probability. Included in the controls are: gender and race dummies; age and age squared; dummies for household income; and dummies for level of education. Robust standard errors are reported in the parentheses, and the asterisks represent the conventional levels of significance. All regressions are weighted using appropriate sampling weights.

selection rather than *true* duration dependence, implying that as workers spend longer in unemployment, they do not revise down their beliefs about finding a job. In our model, workers learn about the likelihood of finding a job when unemployed. However, beliefs about one's employment opportunities can improve as well as deteriorate. If a worker receives many job offers, which are rejected, in our framework it is likely that they revise their belief about finding a job upward. On the contrary, if a worker receives no job offers, as time passes, they will update their perceived chances of finding a job downward. Respondents in the SCE are asked on a monthly basis, "*the percent chance that within the coming 3 months, you will find a job that you will accept*". To understand how the duration of the unemployment spell impacts beliefs we regress the log of the elicited probability on the log of unemployment duration. Results are presented in Table 1. The left hand panel considers all unemployed workers.

In addition to the baseline SCE we use information from the labor market survey, a supplementary survey that interviews respondents every four months. The first period in our sample is March 2014 and the final survey is conducted in November of 2019. Survey respondents are asked how many job offers they received in the previous four months. If respondents indicate zero offers, we can identify specific unemployment spells, or subsets of spells, to which we know the unemployed worker has not received an offer. Note, this approach will omit many spells without offers either from respondents not interviewed in the supplementary survey, or those interviewed while in employment or at the start of their unemployment spell. Full details of the construction of this subsample is provided in Appendix A.1. The results in which the sample is restricted to such spells are presented in the right hand panel of Table 1.

The results from the full sample are consistent with the findings of Mueller et al. (2021). Absent any controls the elasticity of the perceived job finding rate with respect to unemployment duration is -0.190, implying that someone with 10% longer time spent in unemployment will believe they have a 1.9% lower probability of finding a job in three months. The fact that those in unemployment longer have more pessimistic beliefs could be because they reduce their beliefs while in unemployment, i.e., *true* duration dependence. Alternatively, it could be because of dynamic selection, i.e., those who believe they have a high probability of finding a job are correct, and consequently leave unemployment faster. Comparing column one with column two and three, one can see that as more time invariant characteristics are controlled for, duration dependence diminishes. In the case of the full sample, once spell fixed effects are included the majority of the duration dependence has dissipated, and the coefficient on log duration is no longer significant at any conventional level.

As discussed, our model implies that a worker's perception of the job finding rate can rise or fall, depending on circumstance. Since good news comes in the form of job opportunities, and bad news is the absence of new information, restricting the sample to those who do not receive job offers provides an interesting test of the model's mechanism. Results presented in columns one and two look similar across the full and restricted sample. That is, those who have spent longer in unemployment are more pessimistic about their probability of finding a job. However, once the spell fixed effect is included in column three, the picture is quite different. This column represents *true* duration dependence since time invariant unobservable characteristics within a given unemployment spell are controlled for. When we restrict the sample to those for whom we can be sure have not received a job offer, there is a clear sign of *true* duration dependence. Unemployed workers without job offers reduce their perceived probability of finding a job by 1.56% following a 10% increase in the length of their unemployment spell.

Table 2: Perceived 3-month job-finding and labor market switching

	Full sample			Spells without job offers		
	(1)	(2)	(3)	(1)	(2)	(3)
Log unemployment duration months	-0.184***	-0.156***	-0.076	-0.168***	-0.139***	-0.156**
	(0.016)	(0.018)	(0.050)	(0.030)	(0.031)	(0.068)
Moved state last 3 months	0.418***	0.351***	0.343*	0.610***	0.576***	0.558***
	(0.076)	(0.091)	(0.189)	(0.148)	(0.175)	(0.116)
Demographic controls		x			x	
Spell fixed effects			x			x
Observations	2,156	2,133	2,156	673	665	673
R^2	0.103	0.162	0.784	0.079	0.194	0.765

Note: The specification is as in Table 1 with the inclusion of a dummy variable, which takes the value one if the respondent's primary residence changed state in the last three months.

One potential pitfall of segmenting the unemployed by those without job offers, is that it could reflect the unemployed who have disengaged from the search process. To explore this, we exploit another question in the labor market supplementary survey. Respondents are asked "*within the last seven days, about how many total hours did you spend on job search activities*". We define a worker to be searching with low intensity if their maximum hours searching in a spell is two hours or less, representing approximately 10% of respondents. Table A.3 interacts unemployment duration with this measure of search intensity. The result that beliefs deteriorate with unemployment duration are stronger if one focuses on the unemployed exerting greater effort in finding a job. This suggests that our results in Table 2 are not driven by unemployed workers exerting little effort in finding employment.

The second fact we document relates to workers switching labor markets. In the model, workers form beliefs over their opportunities. These beliefs manifest themselves into actions. For example, if beliefs about future employment become sufficiently pessimistic, a worker will switch between markets. The labor market can be interpreted as a sector, occupation or geographical location. In the data, the advantage of the latter over the former, is that the moment a geographical switch occurs in unemployment can be clearly identified. We define a change in labor market if a worker moves their primary residence across state. To identify this change, a respondent in the survey must: (i) respond as living in a different primary residence from the previous survey; (ii) report different states as their primary residence from one survey to the next.

Table 2 presents the same specification in Table 1, with the inclusion of a dummy variable to indicate whether an unemployed worker has moved state in the past three months. If the decision to move state is a result of the endogenous decision in the model, they will do so when their beliefs have sufficiently deteriorated. Thus, after changing labor markets, their beliefs about finding a job will be higher than they were previously. Across all specifications, the unemployed who have moved state in the previous three months, see a positive and significant increase in their perceived likelihood of finding a job. Acknowledged in the model is that not all changes in labor market are driven by beliefs. For example, people may change their geographical location as a result of external factors like family, amenities etc. If these external factors are exogenous to one's beliefs, then movements in the labor market need not be associated with increases in beliefs. It is further evidence of the model therefore that those with relatively worse unemployment spells, as measured by offers, see larger increases in their perception of the labor market following a change in market.

3. The Baseline Model

We now turn to our model setup. As the data clearly show, beliefs change endogenously with individual experiences in the labor market and play a substantial role in wage setting. While most models of the labor market incorporate only the latter feature, our model is designed to showcase the interaction between both mechanisms.

3.1. The Environment

Time is continuous and the economy consists of a continuum of infinitely lived workers and firms who discount the future at a constant rate r, and populate a continuum of identical markets $m \in [0, 1]$. In each market there is a potentially infinite supply of homogeneous firms that can enter and do so if it is profitable. Workers are *ex ante* homogeneous and can move between markets at a cost χ. In the model, we are relatively agnostic about what exactly such a market is. One might think of it as a region or as an occupation for example. Hence, χ can be thought of as a catchall cost that might embed retraining or moving. When taking our model to the data, however, we narrow our interpretation of a market to that of a geographic area, which is useful for the reasons outlined in section 2.

Each worker has a unique suitability $s(m) \in [0, \overline{s}]$ for every market, where $\overline{s} \leq 1$. The suitability of a particular market represents the share of jobs in that market a worker is able to do, the index s is unobservable to the worker. However, agents can learn about their suitability by spending time in the market. We assume that an individual's suitability s is random across market index m. Therefore, from an aggregate perspective and from the perspective of firms, all markets are identical. From the perspective of workers, markets are in principle heterogeneous according to their suitability parameter but *ex ante* identical, as suitability in a market is unobserved.

In addition to the informational friction described, the labor market is also characterized by search frictions. Every market m is modeled as in the canonical Diamond-Mortensen-Pissarides model (DMP) (see e.g. Mortensen and Pissarides (1994)) with heterogeneity in match productivity. Unemployed workers meet open vacancies in a market according to a Cobb-Douglas matching function. Job offers arrive to workers at a contact rate $\tilde{\lambda}_s := s\overline{\lambda}$. Where $\overline{\lambda}$ is determined by the search frictions in the economy and s, the share of suitable offers is imperfectly observable. Once matched, a worker and a firm draw productivity $z \sim \text{Pareto}(1, \alpha)$. Conditional on this draw, the worker and the firm engage in Nash bargaining, and, if this negotiation is successful, form a match that produces z units of output.

Two distinct shocks can dissolve a worker-firm match. First, with Poisson rate δ, workers are hit by a match destruction shock that forces them out of their match if employed but leaves them in their labor market (meaning workers retain all information they have collected on the market). Secondly, with Poisson rate η, a reallocation shock forces the worker to reallocate to a new market. If such a worker is in employment, the match is dissolved. This entails a full loss of information on the market environment. We interpret this shock as a reallocation to a new market that is necessary or desirable for reasons outside the model, such as moving locations for family reasons or because of preference shocks.

3.2. Worker Beliefs

For a worker in a market in which they are suitable for a share s of jobs, the *true* Poisson arrival rate is given by $\tilde{\lambda}_s := s\overline{\lambda}$. The contact rate $\overline{\lambda}$ is finite and will be endogenized later. At any time during a worker's career, if they so choose, a worker has the option to pay a cost χ and allocate to a new market. Since *ex ante* all markets are indistinguishable from a worker's perspective, they will move to their new market in a random fashion.

Assumption 1. *For any worker, the distribution of suitability $f(s)$ is idiosyncratic across markets and follows a truncated gamma distribution taking the form*

$$f(s|k_0, \theta_0) = \frac{1}{\gamma\left(k_0, \frac{\overline{s}}{\theta_0}\right)\theta_0^{k_0}} s^{k_0-1} e^{-\frac{s}{\theta_0}} \quad \textit{for } s \in [0, \overline{s}]. \tag{As. 1}$$

where $\theta_0 > 0$, $k_0 > 0$ and $\overline{s} \in (0, 1]$ are primitive parameters and $\gamma(\cdot)$ is the lower incomplete gamma function[3].

The truncated gamma distribution specified in assumption 1 nests a broad class of distributions for different parameter values of k_0 and θ_0. Commonly used distributions that are special

[3]The lower incomplete gamma function is defined as

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt.$$

cases of (As. 1) are reported in Table A.4 in the appendix. Given rational expectations of workers, the prior distribution of offer arrival rate also follows a truncated gamma distribution with shape parameter k_0, but scale parameter $\theta_0\overline{\lambda}$.

Result 1. *Given the distribution of suitability in* (As. 1) *and rational expectations of workers, it follows that a worker's prior belief over the offer arrival rate is also a truncated gamma distribution, and given by*

$$\pi(\lambda) = \frac{1}{\gamma\left(k_0, \frac{\overline{s}\overline{\lambda}}{\theta_0\overline{\lambda}}\right)\left(\theta_0\overline{\lambda}\right)^{k_0}} \lambda^{k_0-1} e^{-\frac{\lambda}{\overline{\lambda}\theta_0}} \quad \text{for } \lambda \in [0, \overline{s}\overline{\lambda}]. \tag{Re. 1}$$

A proof of Result 1 is provided in Appendix A.4. Once a worker is established in a particular market they receive job offers, and the frequency of these offers will help inform the worker regarding the underlying arrival rate in the market. Since workers only search for jobs in unemployment the information available to them is a series of lapsed time between job offers. As we show in Appendix A.5, Bayesian updating yields a truncated Gamma posterior over arrival rates which is fully determined by two sufficient statistics: The length of time spent in unemployment in a given market and the total number of job offers received during that time. Although every worker has a rich history of employment and unemployment spells, only two statistics are sufficient to summarize a worker's beliefs, which simplifies the model enormously. Result 2 formalizes this insight.

Result 2. *For a given worker, let τ denote the time spent in unemployment in a given market and let n denote the number of encounters during that time span. Given the worker's prior over the encounter rate is as defined by* (As. 1)*, then the worker's posterior over the encounter rate is given by*

$$f(\lambda|n,\tau) = \frac{1}{\gamma\left(n+k_0, \overline{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)} \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{n+k_0} \lambda^{n+k_0-1} e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)} \quad \text{for } \lambda \in [0, \overline{s}\overline{\lambda}] \tag{Re. 2}$$

The posterior density is a truncated gamma distribution with an upper bound $\overline{s}\overline{\lambda}$, scale parameter $1\Big/\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)$, and shape parameter $n+k_0$.

In the expression (Re. 2) and throughout the paper, we drop the tilde notation on λ to distinguish it from the actual encounter rate with suitability s, given by $\tilde{\lambda}_s$. One can derive an expression for the mean of this distribution, which is crucial for the worker value functions:

Result 3. *For a worker with labor market history summarized by n and τ, the instantaneous expectation of the offer arrival rate is given by*

$$\lambda(n,\tau) = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-1} \frac{\gamma\left(n+k_0+1, \overline{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)}{\gamma\left(n+k_0, \overline{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)} \tag{Re. 3}$$

where the second term is the ratio of two recurrent incomplete gamma functions.

Result 3 is derived in Appendix A.6. Figure 1 shows an example of how the posterior density, defined by (Re. 2) evolves over time for a hypothetical worker's labor market history. The distribution of suitability is $f(s)$ is as is calibrated later, which corresponds to $\overline{s} = 1$, $k_0 = 1.1$ and $\theta_0 \to \infty$, see Table 3. The figure is initialized with a worker with $\tau = 0$ and $n = 0$. Thus the initial distribution over λ is the prior reflected in Result 1. The sawtooth pattern of the mean of the posterior represents a worker becoming more pessimistic as time goes on without an offer, where τ increases and n stay constant with intermittent discrete upward jumps when offers arrive and n increases by one. As the worker spends more time in the market, the credible interval narrows and the posterior becomes more and more precise. While the specifics of the second panel of Figure 1 are unique to this arbitrary history, the general pattern in the figure holds universally. Workers' expectations over the arrival rate worsen in a continuous fashion with increasing τ and improve discontinuously with discrete jumps in n.

3.3. Value Functions

Simplifying the posterior in the fashion described in Result 2 allows us to write down the workers value functions. Let V_e and V_u denote the value functions of an employed and an unemployed worker respectively. An employed worker's value satisfies the following Hamilton-Jacobi-Bellman (HJB) equation:

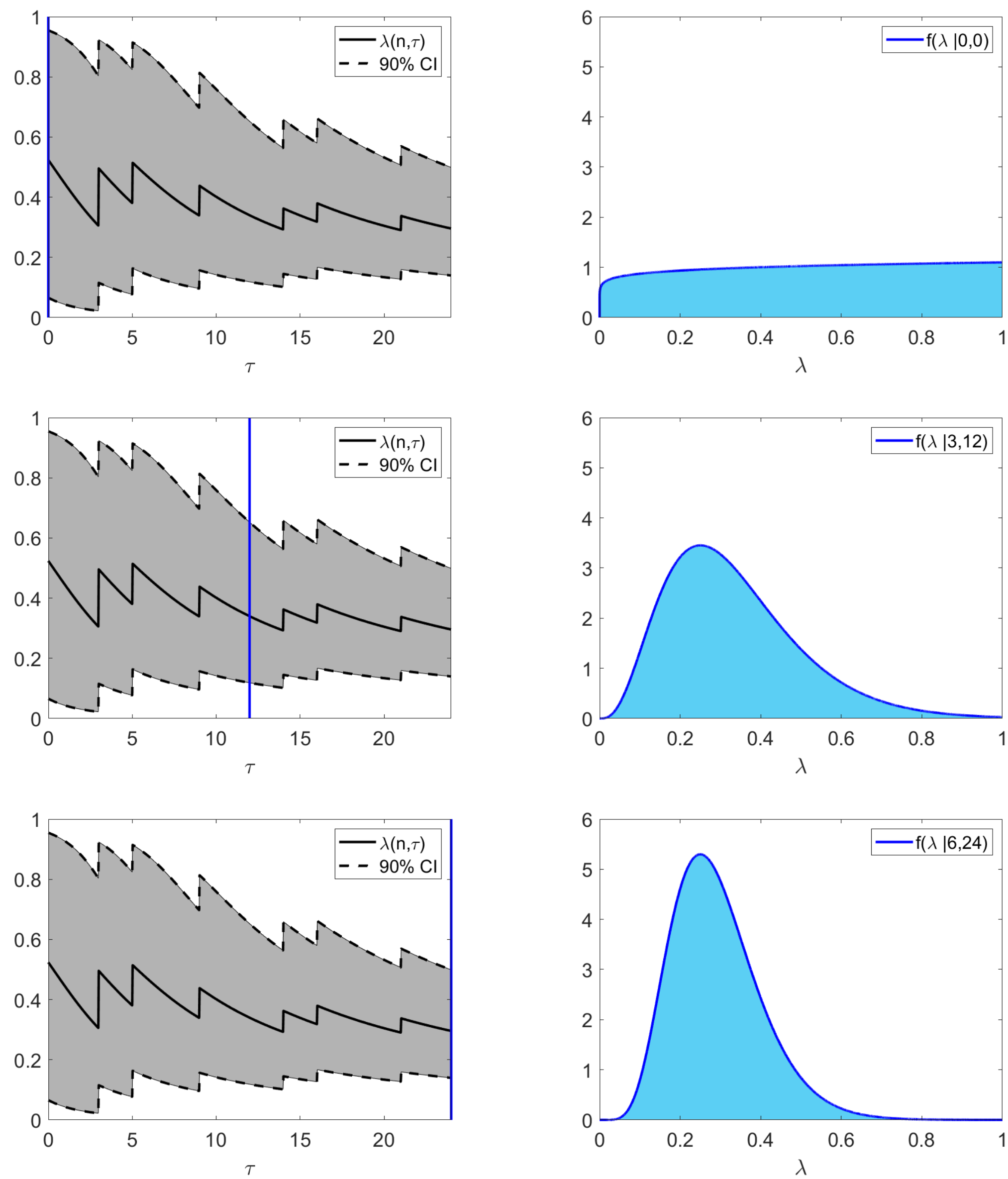
$$(r + \delta + \eta)V_e(z, n, \tau) = w(z, n, \tau) + \delta V_u(n, \tau) + \eta \left[V_u(0, 0) - \chi\right] \tag{1}$$

Employed workers receive their wage until the match is terminated. If it is terminated by a δ-shock, they re-enter unemployment, retaining their information (n, τ). If the match is terminated by an η-shock, they also enter unemployment but in addition lose all of their information ($n = \tau = 0$), as they reallocate to an unknown market. Moreover, a market transition, voluntary or not, always incurs the mobility cost χ. Contained in equation (1) is no option for a worker to voluntarily switch markets. Although a worker can always decide to switch markets voluntarily at cost χ, they will never do so in employment, since their information set and the expected value of switching remains unchanged. Absent on-the-job-search no new information regarding the market is obtained while in employment, hence n and τ are fixed during an employment spell. A simple revealed preference argument shows that it would never be optimal for an employed agent to voluntarily switch markets. If a worker accepts a job at wage $w(z, n, \tau)$ given their information (n, τ), they must value employment at wage $w(z, n, \tau)$ over the option to switch, as otherwise they would value switching over staying unemployed under information (n, τ) and would have left their current market prior to meeting their employer.[4] Finally, in expression (1) no terms contain a tilde implying arguments are in expectation. Hence, the *true* suitability of a worker to the market plays no direct role in the value function.

We now turn to the value function of the unemployed which features learning. Since at any moment workers have the option to switch markets, the dynamic programming problem takes the form of an optimal stopping time problem. Let $\lambda(n, \tau)$ denote the posterior mean of the encounter rate distribution. Then the value of an unemployed worker $V_u(\cdot)$ satisfies (2). A

[4]Formally, this argument is implied by the continuity of V_u in both arguments which can be shown from the Bellman equation using standard arguments.

Figure 1: Illustration of a worker's updating rule



Note: The left hand panel displays the evolution of the mean of the posterior density, with associated 90% credible intervals. Where the parameter $\overline{\lambda}$ has been fixed at unity and the prior density for suitability is as described in Assumption 1 with upper-bound $\overline{s} = 1$, scale parameter $\theta_0 \to \infty$ and shape parameter $k_0 = 1.1$, as in the calibration. The right hand panel shows the evolution of the entire posterior density.

formal derivation of the value function is performed in Appendix A.7.

$$rV_u(n,\tau) = \max\left\{\left(b + \overbrace{\lambda(n,\tau)\mathbb{E}_z\left[\max\left(V_e(z,n+1,\tau) - V_u(n+1,\tau),0\right)\right]}^{(1)} + \underbrace{\eta\left(V_u(0,0) - \chi - V_u(n,\tau)\right)}_{(2)} + \underbrace{\lambda(n,\tau)\left(V_u(n+1,\tau) - V_u(n,\tau)\right)}_{(3)} + \underbrace{\frac{\partial V_u(n,\tau)}{\partial \tau}}_{(4)}\right), r(V_u(0,0) - \chi)\right\} \quad (2)$$

Consider first the subset of the domain $\mathbb{N}_0 \times \mathbb{R}_+$ on which the worker does not voluntarily switch markets. In this case, the max operator in equation 2 selects the first of its two arguments. It is comprised of five components: A worker receives a flow value of b while unemployed. At an unknown rate $\tilde{\lambda}_s$, the worker encounters a firm and has the option to become employed. Since the true encounter rate $\tilde{\lambda}_s$ is unknown, the worker uses their posterior f to determine their expected value, integrating over all possible values of λ. This turns out to be identical to simply using the posterior mean as the arrival rate when computing the value function. Term (1) therefore captures the option value of employment. At the time of a job offer, the worker's information state changes and therefore so do the worker's beliefs. Therefore, both the value of employment and unemployment are compared at the new information state with $n+1$ encounters. Unemployed workers are subject to forced reallocation shocks, captured by term (2). Terms (3) and (4) correspond to the Bayesian updating rule. Every encounter causes the worker to positively update their posterior (term (3)) but time spent in unemployment without receiving offers leads to negative updating of the posterior (term (4)).

Now consider the subset of the domain on which the max operator selects the second argument, i.e. states in which the worker chooses to leave their market. Since suitability is idiosyncratic, each market looks identical to one another from the perspective of an uninformed worker. A worker thus reallocates to a randomly drawn new market with a prior over suitability defined by Assumption 1. The value associated with this option is $V_u(0,0) - \chi$.

Although the posterior distribution enters equation 2 only through the posterior mean, $\lambda(n,\tau)$ is by no means a sufficient statistic for the value function at (n,τ). This is because of the learning terms (3) and (4), which will generally play a larger role when n and τ are small and disappear as n and τ go to infinity. The presence of these learning terms means that the worker takes into account how their beliefs will change in the future given any possible labor market history they might experience. This sets our model apart from models using the "anticipated utility" framework (see e.g. Potter (2021)), a popular assumption that rules out such an internalization of belief changes to simplify computation.

As we show in Appendix A.8, we can use the Nash bargaining assumption detailed in section 3.4 to write the right hand side of equation 2 solely as a function of V_u, n and τ. Appendix A.9.1 outlines how the resulting HJBVI can be written as the special case of a linear complementarity problem (LCP) for which fast solution methods are known. The corresponding solution yields optimal stopping times $T^\star(n)$ for each $n \in \mathbb{N}_0$, i.e. the total market unemployment duration at which a worker with n encounters decides to leave the market.

Consider now the firm's problem. A matched firm receives the flow output of a match minus the wage until termination. By free entry, the value to the firm of any match is zero after termination (as the value of a vacancy is zero). Therefore, the value of a filled match is given by

(3). Notice again the *true* suitability of the worker for the market does not feature in the value function. We assume that the firm also does not observe the true s, just whether the worker is suitable for this particular job. However, since wages depend on the worker's beliefs, even if the firm were to know s it would not change the value of the match.

$$(r + \delta + \eta)J(z, n, \tau) = z - w(z, n, \tau) \tag{3}$$

Given V_u, we can therefore write the joint surplus $S(z, n, \tau) := V_e(z, n, \tau) + J(z, n, \tau) - V_u(n, \tau)$ of a match as

$$(r + \delta + \eta)S(z, n, \tau) = z - (r + \eta)V_u(n, \tau) + \eta\,[V_u(0, 0) - \chi]$$

Hence, surplus is positive when

$$z > z^\star(n, \tau) := (r + \eta)V_u(n, \tau) - \eta\,[V_u(0, 0) - \chi] \tag{4}$$

$z^\star(n, \tau)$ is the minimum productivity draw necessary for an encounter to become a match. The formula determining this threshold is intuitive - at the threshold, the value of unemployment in flow units must equal z adjusted for the possibility of forced reallocation shocks. $z^\star$ turns out to be an important statistic. For one, it is a sufficient statistic for computing wages. Moreover, since we assume that $z \sim \text{Pareto}(1, \alpha)$, the probability of a match given an encounter is given by $\max\{1, z^\star(n, \tau)\}^{-\alpha}$.

3.4. Wages

We assume that wages are set by Nash bargaining. In our environment, neither the firm nor the worker know the true market environment for the worker. Instead the information set is limited to (n, τ) which we assume is observable to both parties. While the assumption that (n, τ) is observable to the firm is not completely innocuous, we argue that two features of the environment are sufficient for this information being revealed even in a private information setting. First, τ has to be public information. We deem this assumption reasonable, as in reality unemployment duration in a given market is generally observable during the job application process. Second, we assume that a worker has access to a technology that allows them to credibly prove any past encounter to the firm. Since a higher n strengthens the worker's bargaining position, they will always have an incentive to communicate all encounters to any prospective employer. However, the firm will require proof of every encounter, since a higher n deteriorates their position. In equilibrium, the worker will communicate and submit proof for every true encounter. Under this assumption, the co-operative Nash bargaining game therefore mirrors the standard, full information cases with disagreement points $(V_u(n, \tau), 0)$. In Appendix A.11 we show that, given a solution for $z^\star$, wages are affine in productivity z.

$$w(z, n, \tau) = \beta z + (1 - \beta)z^\star(n, \tau) \tag{5}$$

3.5. Worker Distributions

In Appendix A.10, we show that given the matching thresholds $z^\star(n, \tau)$, the distributions $u(n, \tau, s)$ and $e(n, \tau, s)$ can be found as the solution to a system of differential equations which can be solved numerically by iterating over values of n: First, for $n = 0$ and $\tau < T^*(0)$,

$$\partial_\tau u(0,\tau,s) = -(\tilde{\lambda}_s + \eta)u(0,\tau,s)$$
$$\implies u(0,\tau,s) = u(0,0,s)\cdot\exp(-(\tilde{\lambda}_s+\eta)\tau) = \mu\cdot f(s)\cdot\exp(-(\tilde{\lambda}_s+\eta)\tau) \tag{6}$$

where μ is the number of entrants in each market (which can be found using the condition that population must integrate to one). Then, for $n \geq 1$, $\tau < T^*(n)$, the employment distribution can be characterized as a solution to a set of linear non-homogeneous ODEs given the solution for $n-1$:

$$e(n,\tau,s) = \mathbb{I}(\tau < T^*(n-1))P(z \geq z^*(n,\tau))\frac{\tilde{\lambda}_s}{\delta+\eta}u(n-1,\tau,s) \tag{7}$$

$$\partial_\tau u(n,\tau,s) = \Bigg[\mathbb{I}(\tau < T^*(n-1))\left(P(z < z^*(n,\tau)) + P(z \geq z^*(n,\tau))\frac{\delta}{\delta+\eta}\right)\cdot \tilde{\lambda}_s u(n-1,\tau,s)\Bigg] - (\tilde{\lambda}_s+\eta)u(n,\tau,s) \tag{8}$$

with boundary conditions $u(n,0,s) = 0\,\forall n \geq 1$. After solving the system of differential equations, the aggregate unemployment rate is computed as

$$\overline{u} = \sum_{n'=0}^{\infty}\int_0^{T^\star(n)}\int_0^1 u(n',\tau',s')d\tau' ds'.$$

3.6. Free Entry

To close the model we now consider the free entry condition. Using equations 3 and 5, the value of a match to a firm is given by

$$(r+\delta+\eta)J(z,n,\tau) = (1-\beta)z - (1-\beta)z^\star(n,\tau) \tag{9}$$

The value of posting a vacancy to a firm is given by V_v.

$$rV_v = -\kappa + \overline{\lambda}_f\left(\frac{1}{\overline{u}}\right)\sum_{n=0}^{\infty}\int_0^{T^\star(n)}\int_0^1\left(s'u(n,\tau',s')\left(\int_{\max\{1,z^\star(n+1,\tau)\}}^{\infty} J(z,n+1,\tau)d\Gamma(z) - V_v\right)\right)d\tau' ds'$$

To keep a vacant position open requires a flow cost κ. The aggregate contact rate of workers to firms — both the suitable and unsuitable potential employees, is defined as $\overline{\lambda}_f$. Upon meeting a worker, they may or may not be suitable, and will vary in their labor market history summarized by n and τ. The unsuitable workers provide no value to the firm. The labor market history is of consequence to the firm as it will determine the bargained wage, given by (5). The ideal candidate from a firm's perspective would be a suitable worker who believes, with some precision, that they are in a poor market. This worker will be more likely hired, having a lower threshold productivity, and for a given productivity will command a lower wage.

Free entry of firms implies that firms continue to enter until the value of doing so is zero, $V_v = 0$. Substituting the free entry condition into the value of a vacancy and following analogous steps to those of Appendix A.8 for the value of a match to the firm yields expression (10).

$$\kappa = \overline{\lambda}_f \left(\frac{1}{\overline{u}}\right) \sum_{n=0}^{\infty} \int_0^{T^\star(n)} \int_0^1 \left(s' u(n, \tau', s') \int_{\max\{1, z^\star(n+1,\tau)\}}^{\infty} J(z, n+1, \tau) d\Gamma(z) \right) d\tau' ds' \tag{10}$$

$$\text{where } \int_{\max\{1, z^\star(n+1,\tau)\}}^{\infty} J(z, n+1, \tau) d\Gamma(z) = \begin{cases} \frac{1-\beta}{r+\delta+\eta}\left(\frac{\alpha}{\alpha-1} - z^\star(n,\tau)\right) & \text{for } z^\star(n,\tau) < 1 \\ \frac{1-\beta}{r+\delta+\eta}\left(\frac{1}{\alpha-1}\right) z^\star(n,\tau)^{1-\alpha} & \text{for } z^\star(n,\tau) \geq 1 \end{cases}$$

The suitability parameter directly enters in the free entry condition. This is because workers with low suitability will still encounter firms and thereby congest the market without producing matches. This leads to an important source of misallocation: Without *ex ante* information on which markets they are suitable for, workers spend a lot of time in unsuitable markets. This drives down match efficiency for all workers, causing a decline in job posting incentives that leads to a lower job encounter rate $\overline{\lambda}$ in general equilibrium.

Finally, the worker and firm contact rates are determined by a constant returns to scale Cobb-Douglas matching function $m(\overline{u}, \overline{v}) = Au^{\omega}\overline{v}^{(1-\omega)}$, where $\overline{v}$ is the aggregate measure of vacancies. Therefore, the maximum worker contact rate is computed as $\overline{\lambda} = m(\overline{u}, \overline{v})/\overline{u}$ and given by (11). Given constant returns to scale and identical markets, it is isomorphic to think of the matching function as defined at the market rather than aggregate level.

$$\overline{\lambda} = A^{\frac{1}{\omega}} \overline{\lambda}_f^{\frac{\omega-1}{\omega}}. \tag{11}$$

3.7. *Equilibrium*

We are now ready to define an equilibrium in the baseline model:

Definition 1. *A* ***baseline equilibrium*** *is a tuple* $(V_e(\cdot), V_u(\cdot), J(\cdot), z^\star(\cdot), w(\cdot), u(\cdot), e(\cdot), \overline{\lambda}, \overline{\lambda}_f)$ *such that*

1. *The value functions solve their Bellman equations* (V_e, J) *or the HJBVI* (V_u) *(equations 1, 2 and 3)*
2. *Productivity thresholds are set according to equation 4*
3. *Wages are set according to Nash bargaining (equation 5)*
4. *The employment distributions* $u(\cdot)$ *and* $e(\cdot)$ *solve the Kolmogorov Forward Equations (equations 6, 7 and 8)*
5. *Free entry pins down* $\overline{\lambda}$ *and* $\overline{\lambda}_f$ *(equations 10 and 11)*

4. The Full Information Model

To assess the effects of the informational friction, we compare our economy to one with full information. To this end, we construct what we term the "full information model". This model is set up to fully mirror the baseline model with one important difference: We assume that at any point every worker is fully informed about their suitability in every market. There is no uncertainty about the underlying arrival rate of jobs, only about the realization of the Poisson process. In this environment, there is no need for workers to keep track of the information state (n, τ). Under full information, a worker will always choose the market they are most suitable for, $s = \overline{s}$. In equilibrium, every worker will thus be suitable for $\overline{s}$ jobs in their market and the Poisson arrival rate of jobs to unemployed workers is $\overline{s}\bar{\lambda}$.

The resulting model turns out to be the baseline DMP model, replicated infinitely many times along an irrelevant dimension m, which we can drop from the notation. This is an elegant feature of our baseline model: When the informational friction is removed, the resulting model is simply the standard search model under free entry. The resulting HJB equations are

$$(r + \overline{s}\bar{\lambda} + \eta)V_u = b + \overline{s}\bar{\lambda}\mathbb{E}_z[\max\{V_e(z), V_u)\}] + \eta(V_u - \chi) \tag{12}$$

$$(r + \delta + \eta)V_e(z) = w(z) + \delta V_u + \eta(V_u - \chi) \tag{13}$$

$$(r + \delta + \eta)J(z) = z - w(z) \tag{14}$$

Since the full information model is well understood. For brevity, we simply present the equilibrium in Definition 2.

Definition 2. *A* ***full information equilibrium*** *is a tuple* $(V_e(\cdot), V_u(\cdot), J(\cdot), z^\star(\cdot), w(\cdot), \overline{u}, e(\cdot), \bar{\lambda}, \bar{\lambda}_f)$ *such that*

1. *The value functions solve their Bellman equations (equations 12, 13, and 14)*
2. *Productivity threshold $z^\star$ is set such that $V_e(z^\star) = V_u$.*
3. *Wages are set according to Nash bargaining.*
4. *The employment rate e and unemployment rate u are given by a steady-stare condition.*
5. *Free entry and a constant returns to scale matching function pin down the arrival rate of jobs to workers and firms, $\bar{\lambda}$ and $\bar{\lambda}_f$.*

5. Calibration and identification

5.1. Parameterization

We now turn to the calibration of the model. For the purpose of computing welfare gains, we use the same calibration for both the baseline and the full information model. This calibration is such that we match targets in the baseline model. For the purpose of comparing implications for frictional wage dispersion (section 6.2) and when studying the aggregate shock (section 7) we re-calibrate the full information model to hit the same targets as the baseline model. Appendix A.12 displays the corresponding calibration table for the full information model.

The distribution of suitability was left as general as possible in the exposition of the model. We parameterize the distribution such that the scale parameter $\theta_0 \to \infty$ and the upper bound of suitability $\overline{s} = 1$. The latter implies that a worker in their best market is suitable for every job in that market. Under these restrictions, the primitive distribution of suitability follows a beta distribution with parameters k_0 and one. Assumption 1 can be re-written as (As. 1').

$$f(s) = k_0 s^{k_0 - 1} \quad \text{for } s \in [0, 1] \tag{As. 1'}$$

Parameterizing $\overline{s} = 1$ is inoccuous: In the model, agents care about the distribution of offers $\lambda := \bar{\lambda}s$. Thus, from a worker's perspective, it is the product of $\bar{\lambda}$ and suitability s that is important. As is standard in search models, the potential frequency of offers $\bar{\lambda}$ is already over-identified through the cost of vacancies κ and efficiency of the matching function A. Setting

$\theta_0 \to \infty$ is not a normalization but retains a lot of flexibility in the distribution of s. While assumption As. 1' rules out certain hump-shaped distributions for s, it still nests the main edge cases of interest. In particular, the distribution family nests the full information case when $k_0 \to \infty$. Generally, k_0 governs the speed of learning in the model. To see this, take the level of entropy of the prior distribution, defined below.

$$H(f) := -\int_0^1 f(s)\log(f(s))ds = 1 - \frac{1}{k_0} - \log(k_0)$$

Entropy is a measure of the average level of '*information*' drawn from a realization of a random variable. In our context, information will take the form of job offers (or the lack of) and the more uncertain the prior, as measured by entropy, the more important the information will be. To see this intuitively, take the two extreme values of k_0, that is $k_0 \to 0$ and $k_0 \to \infty$. In both cases, the prior distribution will be degenerate at $s = 0$ and $s = 1$, respectively. The value of Shannon Entropy $H(f) \to -\infty$ and any new information, good or bad, will have no impact on the prior and there will be no learning. The amount of learning is maximized when the uncertainty of the prior is maximized. The function $H(f)$ is maximized when $k_0 = 1$ which corresponds to a uniform distribution on $[0, 1]$.

In total, there are 11 parameters to calibrate, listed in Table 3. We calibrate 5 parameters externally. We set the discount rate to a value of 5% annually. The monthly separation rate is 2% of which we attribute two thirds to reallocative separations, which corresponds roughly to the share of separations that are quits in the JOLTS data[5]. We set the matching function elasticity with respect to the number of unemployed to 0.3, following Borowczyk-Martins et al. (2013). Finally, we normalize $A = 1$, which simply pins down the scale of vacancies but has no other allocative consequences.

There are 6 remaining parameters that are calibrated internally. We follow the convention to calibrate b by targeting the replacement rate, defined as the ratio of b over the mean wage in the economy. We calibrate to a replacement rate of 0.4, which is what is estimated by the OECD for a single person without children. This is in line with values commonly used in the literature, for example in Shimer (2005). As discussed in sections 6 and 7, some of the puzzles we address in this paper can in principle be resolved by choosing a much higher or much lower target for the replacement rate. For example, a higher value for unemployment benefits will imply greater propagation of an aggregate shock in the full information model. Likewise, a much lower and negative value leads to higher frictional wage dispersion. By choosing an empirically plausible intermediate value we thus show how our model can naturally reconcile both puzzles. We can pin down κ by imposing a realistic unemployment rate, which we set to 5%. Since the rate at which the employed lose their job has been calibrated externally, the unemployment rate is thus pinned down by the rate at which the unemployed find work.

The parameter β governs the share of surplus appropriated by workers. We pin it down by using estimates of the pass-through elasticity of wages with respect to productivity shocks. When β is high, so is the pass-through rate, i.e. productivity increases lead to higher wage increases.

[5]This is consistent with our interpretation that η-shocks can be thought of as preference shocks that sufficiently incentivize reallocation. The same argument is made in Alvarez and Shimer (2011) to calibrate a similar parameter in a model in which workers switch across industry.

Vice-versa, a low β leads to a low pass-through rate. We use the result from Lamadon et al. (2022) who estimate this elasticity to be equal to 0.13 on average. In our application, it is very easy to find the harmonic mean of the elasticity of wages with respect to productivity:

$$\mathbb{E}_z\left[\varepsilon(z,n,\tau)^{-1}\right] = \mathbb{E}_z\left[\left(\frac{d\log w(z,n,\tau)}{d\log z}\right)^{-1}\right] = 1 + \frac{1-\beta}{\beta}\frac{\alpha}{1+\alpha}\min\{1, z^{\star}(n,\tau)\}$$

The parameter α is the parameter of the Pareto distribution and governs the dispersion in match productivity. To pin this down we match the monthly rate at which workers encounter job offers. We hit a target encounter rate of 0.67, in accordance with the average monthly number of offers for a typical unemployed worker as reported in Faberman et al. (2022). This is equivalent to calibrating to the acceptance rate of jobs, since we calibrate to the unemployment rate and therefore implicitly impose the job finding rate. To see why productivity dispersion governs acceptance decisions, take the two extreme values of α. The limit as α goes to infinity, the productivity distribution is degenerate. Thus all offers are identical, and if any productivity draw is acceptable all must be. Conversely, as α tends to one, the mean productivity tends to infinity as does the option value of rejecting an offer. Thus, all offers are rejected. It turns out that an intermediate value of α of approximately 4.73 hits the encounter rate.

Table 3: Parameters and calibration targets

Parameter	Description	Value	Source/Target
Externally calibrated			
r	Discount rate	0.05/12	5% annual
δ	Separation rate into the same market	$0.02\times\frac{1}{3}$	2% monthly job loss probability
η	Reallocation rate into the new market	$0.02\times\frac{2}{3}$	$\approx$ 2/3 of separations are worker quits in JOLTS
ω	Matching function elasticity	0.3	Borowczyk-Martins et al. (2013)
A	Matching efficiency	1	Normalization
Internally calibrated			
b	Flow value of unemployment	0.458	40% replacement rate $b/\overline{w}$
α	Pareto parameter of productivity distribution	4.728	0.67 monthly encounter rate.
β	Worker bargaining power	0.109	Wage-productivity pass through of 0.13
κ	Vacancy posting cost	1.881	5% unemployment rate
Regression coefficients			
Data:	$\log(\text{elicited belief}_{it}) = -0.076\log(\text{duration}_{it}) + 0.343\text{Moved State}_{it} + \alpha_i + \epsilon_{it}$		
Model:	$\log(\text{elicited belief}_{it}) = -0.079\log(\text{duration}_{it}) + 0.343\text{Moved State}_{it} + \alpha_i + \epsilon_{it}$		
k_0	Shape parameter of suitability distribution	1.10	
χ	Reallocation cost	19.91	

Note: Moments are targeted such that other than the regression coefficients all moments are matched perfectly. The model's regression coefficients are the mean from simulations of 100000 workers over 6 years in the model and the regression specifications are as reported in the third column of the full sample in Table 2.

There are two remaining parameters to be identified. First, the cost incurred when one switches markets, denoted by χ. Second, the parameter k_0 that governs the shape of the distribution of suitability, over which worker's update their expectation. To pin these down we revisit the supportive evidence presented in section 2. In particular, coefficients from the full sample regression reported in column 3 of Table 2 are targeted. The elasticity of beliefs with respect to unemployment duration is determined by the speed of learning. As discussed, this is governed by the entropy of the prior. Hence learning is maximized when k_0 is equal to one, and minimized as k_0 approaches zero or infinity. In the model, workers voluntarily switch markets when they become sufficiently pessimistic about their original market. Thus after a voluntary market switch, a worker will become more positive about their labor market opportunities. The size of the increase, the coefficient on those who move state (switch market), is determined by how much pessimism a worker can tolerate. If the cost of switching (χ) is small, a worker will readily switch market and there will be only marginal increases in optimism following a change in market. Conversely, if the cost χ is large, a worker will not switch market until they have extremely low expectations about their market, and thus following a change in market will see large increases in perceived opportunities.

The parameter k_0 is calibrated at 1.10, implying that the density of suitability is upward sloping. A worker in a new market expects to be suitable for 52.4% of jobs. The model could accommodate a faster learning process with a slightly higher value of k_0, with a distribution closer to the uniform. The market switching cost χ is estimated at 19.91. This corresponds to a switching cost of approximately 17 months of the average wage in the economy. The structural microeconometric literature has a wide range of estimates for the cost of switching markets. For example, our estimate is small relative to Kennan and Walker (2011) who estimate the cost of switching across geographic space at \$312,000. Or Artuç et al. (2010) who estimate the cost of a worker switching industry at 13 times their annual earnings on average. However, there also exist much lower estimates in the literature as well. The lowest estimate we have come across is Xu (2018) who estimate that the cost of switching occupation as being two thirds of a worker's monthly wage.

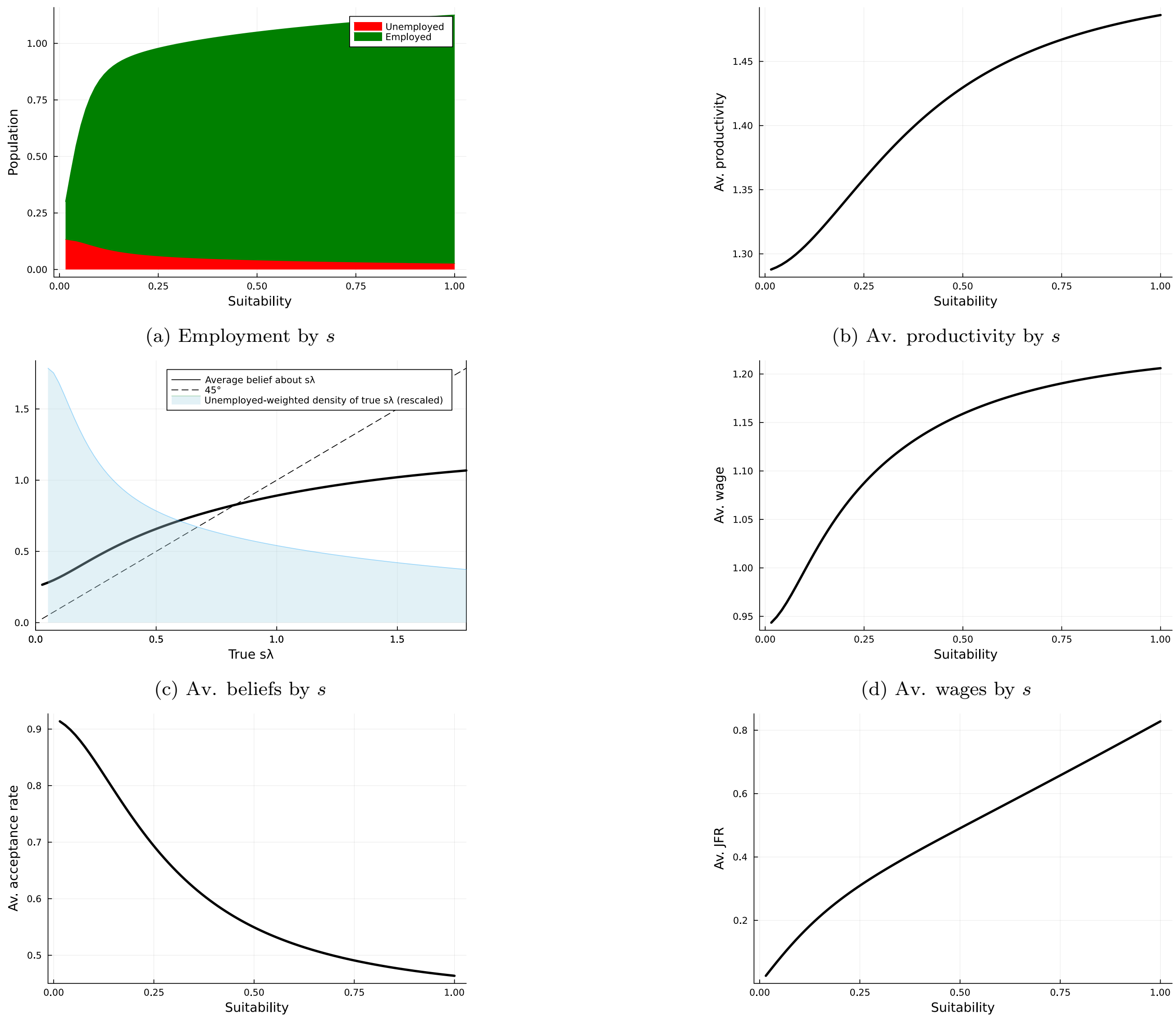
As was discussed in section 2, workers who switch market can do so voluntarily or involuntarily. In the model, involuntary switches come about through the Poisson rate η. If this occurs, it is not clear if after a change in market a worker will become more optimistic or pessimistic about their labor market opportunities. If one conditions on a worker having deteriorating views about their prospects, given they receive no job offers, it is more likely that an observed market switch is voluntary. Similarly, an unemployment spell has an ambiguous impact on a worker's beliefs. If one conditions on a worker receiving no offers, a worker's beliefs will unambiguously decline with the duration of their unemployment spell. Thus, when we run the same regression conditioning on a worker receiving no job offers we obtain the following regression coefficients in a model simulation of 100000 workers over 6 years:

$$\log(\text{elicited belief}_{it}) = -0.114 \log(\text{duration}_{it}) + 0.461 \text{Moved State}_{it} + \alpha_i + \epsilon_{it}$$

For the reasons discussed, both coefficients are significantly larger in absolute magnitude. Although these coefficients are not specifically targeted, when one compares the model's results to the results from the data, the final column of Table 2, they are surprisingly close to one another. We interpret this out of sample fit as adding further credence to the mechanism of our model.

5.2. Summary of model output

Figure 2 summarizes the allocations produced by the calibrated model by displaying the employment distribution as well as averages of variables by worker suitability s. The job finding rate is the product of the contact rate ($s\overline{\lambda}$) and the acceptance rate (given by panel (e)), and is increasing in suitability. Consequently, since separations are fixed across markets, the unemployment rate declines with s. The parameter $k_0 > 1$ and hence the distribution of suitability is upward sloping. Unemployed workers leave bad markets when their estimate of market conditions becomes sufficiently pessimistic. Hence workers, on average, sort into more suitable markets. These two forces result in a unimodal distribution of suitability.



(a) Employment by s

(b) Av. productivity by s

(c) Av. beliefs by s

(d) Av. wages by s

(e) Av. acceptance rates by s

(f) Job finding rates by s

Figure 2: Allocations in the baseline model

Note: Job offer and finding rates in panels (c) and (f) are represented as instantaneous monthly hazards in continuous time.

Productivity and wages are increasing in s, since the more suitable workers have more encounters and can be more selective about which encounters to accept - this can be verified by the fact that the average acceptance rate - the ratio of matches per encounter - in each market is also falling in s.

The figure demonstrates that the model endogenously generates biased beliefs. Workers in bad markets are overly optimistic while workers in good markets are overly pessimistic. This pattern of beliefs in consistent with the findings of Mueller et al. (2021). Irrespective of their true suitability, a worker enters each market with the same prior. If a worker is unsuitable for some specific market and s is close to zero, their posterior mean will converge to the truth from above and hence will on average be larger than the true encounter rate. The reverse argument follows for a worker with high suitability, who is on average overly pessimistic.

6. Model implications

The calibrated model provides us with three consequences of incorporating an information friction in an otherwise standard labor search model, which we believe are worthy of discussion. First, a natural consequence of our model is declining starting wages with the duration of an unemployment spell. This is of consequence to a recent literature evaluating the cost of job loss which typically attribute this relationship to depreciation of human capital while in unemployment. Second, unlike the canonical search model, our calibrated model can generate a level of frictional wage dispersion that is consistent with data. Third, the model highlights new sources of inefficiencies which we document. Overall welfare losses of misinformation are large, and account for more than a third of total output.

6.1. Wage dependence on duration

There is a burgeoning literature that attempts to explain the persistence in the cost of job loss through the lens of a search model of the labor market. A common feature of Ortego-Marti (2016), Burdett et al. (2020), Flemming (2020), and Jarosch (2021) is that when in unemployment a worker loses human capital. The extent of the *scarring effect* of unemployment in all four papers is disciplined by a regression akin to (15).

$$\log\left(w_{it}^{0}\right) = \beta_0 + \beta_1 \text{dur}_{it} + \beta_2 \log\left(w_{it}^{P}\right) + \varepsilon_{it} \tag{15}$$

The specification takes the newly employed and regresses the log starting wage on the duration of unemployment and a measure of previous wages. While the first wage following an unemployment spell can be, on average, lower than a typical wage through the job ladder (Jarosch (2021)) or wage-tenure contracts (Burdett et al. (2020)), there is nothing present in these models other than human capital depreciation that would generate a sensitivity with respect to the duration of the unemployment spell. Consequently, the existence of a mechanism omitted from these studies explaining the phenomenon would lead to inaccuracy in the calibration of human capital depreciation.

In the framework presented here, absent any human capital depreciation, there is a mechanism that generates a negative relationship between unemployment duration and re-employment wages. As a worker spends time in unemployment, assuming they have not turned down wage offers, their perception of their labor market opportunities deteriorates. Over time the worker becomes more pessimistic and is willing to accept lower and lower productivity jobs. Further, their outside option falls, and hence the same productivity job will lead to a lower bargained wage. The equivalent specification to (15) in Flemming (2020) estimates a monthly value of β_1

for the US labor market at -0.03.[6] If one interprets this as being entirely driven by human capital loss in unemployment, then it implies it would take less than two years in unemployment for one's human capital to halve. Simulating worker histories from the baseline model we run the regression specification in (15), using the worker's previous wage as the right hand side variable. Our estimate of β_1 is -0.015. In comparison with the regression in Flemming (2020), the results suggest that the rate of human capital depreciation is over-stated by a factor of two.

6.2. Degree of Frictional Wage Dispersion

Hornstein et al. (2011) document the inability of the canonical search model to generate the degree of frictional wage dispersion seen in the data. To understand this, under full information a worker's reservation wage $\underline{w}$ is the wage to which they are indifferent between employment and unemployment. One can then decompose the degree of wage dispersion potentially generated, measured as the ratio of the reservation wage to the mean wage in the economy as the sum of the replacement rate and a worker's search option. Since the search option increases in the mean wage, for *reasonable* values of the replacement rate, the model is unable to generate a substantial difference between the minimum acceptable wage and the mean wage at large. Implementing this decomposition on the calibrated full information model yields the following.

$$\underbrace{\frac{\underline{w}}{\mathbb{E}(w)}}_{\text{min-mean ratio 96\%}} = \underbrace{\frac{b}{\mathbb{E}(w)}}_{\text{rep. rate 40\%}} + \underbrace{\frac{\beta}{\mathbb{E}(w)}\tilde{\lambda}\int S(z')^{+}d\Gamma(z')}_{\text{search option 56\%}}$$

Under full information, a worker's reservation wage is 96% of the mean wage they can command in the labor market. While it is difficult to accurately estimate the degree of frictional wage dispersion, this number is a lot higher than what is typically found in the data, implying that the degree of frictional wage dispersion is too low in the model. For context, if one defines the minimum as the tenth percentile of the wage distribution, depending on the controls, Hornstein et al. (2007) find a min-mean ratio between 55% and 61%, see Table 2 of their paper.

There have been several recent papers that highlight mechanisms that facilitate a search model generating a greater degree of wage dispersion. Broadly speaking, the strategies for generating more dispersion come in one of two forms: ether reducing the value associated with unemployment; or alternatively, reducing the search option value. An example of the former, Ortego-Marti (2016) includes human capital depreciation in unemployment, hence a worker will accept a lower wage offer to preserve future human capital. Alternatively, Bradley and Gottfries (2021) derive a model in which opportunities arrive akin to stock-flow matching and the employed endogenously draw wage offers from a distribution that stochastically dominates that of the unemployed. This leads to less of an option value for the unemployed. In our model, within a market the search option is fixed and the perceived value of unemployment only changes with changing beliefs. To understand the novel mechanism of our model, consider the following

[6]This number is taken from column 3 of Table 1. A monthly number is obtained my multiplying the weekly estimate of -0.007 by 4.33.

decomposition of the calibrated baseline model.

$$\underbrace{\frac{\underline{w}}{\mathbb{E}(w)}}_{65\%} = \underbrace{\frac{b}{\mathbb{E}(w)}}_{\text{rep. rate } 40\%} + \underbrace{\frac{\beta}{\mathbb{E}(w)}\lambda(0,0)\int S(z',1,0)^{+}d\Gamma(z')}_{\text{search option } 51\%} \underbrace{-(r+\eta)\frac{\chi}{\mathbb{E}(w)}}_{\text{information friction -30\%}}$$
$$+ \underbrace{\frac{1}{\mathbb{E}(w)}\left(\lambda(0,0)\left(V_u(1,0)-V_u(0,0)\right)+\frac{\partial V_u(0,0)}{\partial \tau}\right)}_{\text{learning option } 4\%}$$

In this case, the wage $\underline{w}$ is the lowest possible acceptable wage. This wage corresponds to the reservation wage of a worker indifferent between remaining in their market and switching markets. The replacement rate, as a consequence of the calibration, is identical across both decomposition exercises. The perceived search option varies with a worker beliefs, here the reference group are those newly entering a market, and the value is similar in both the full information and the baseline model. The final two terms are novel to this paper. First, the learning option which plays little role, captures that by simply remaining in unemployment, a worker's value will change through learning. This will either increase discretely with job offers or deteriorate smoothly with the absence of offers. It is the term labeled information friction which allows the model to generate closer to the level of frictional wage dispersion in the data. The size of the cost of mobility determines how long a worker is willing to stay in what they perceive as a market with a low job finding rate. As $\chi \to \infty$ a worker will never leave a market voluntarily and hence even if they believe the job finding rate to be zero, and thus there is no option value, they will remain fixed in their market. It is thus this heterogeneity in beliefs, that governs the heterogeneity in acceptable wages.

Although not directly targeted, our model generates a min-mean ratio of 65%. Introducing the information friction thus allows us to generate substantially more wage dispersion than the full information benchmark. The new friction brings the model much more in line with the estimates reported in Hornstein et al. (2007), which range up to 61%[7]. Thus, while there may be other omitted sources worthy of modeling, frictional wage dispersion in the model becomes much more empirically plausible. It should also be noted that we have used an annualized discount rate of 5%. If workers put more value on future earnings, the option value of unemployment increases and the model thus can generate less frictional wage dispersion. Leaving all else constant, Figure A.4 in the Appendix plots the min-mean ratio varying the calibrated discount rate. Lowering the discount rate does indeed lower the amount of wage dispersion the model can generate. However, even with a rate as low as 1% the model is able to generate a min-mean ratio of 72%, which is significantly more frictional wage dispersion than under the full information benchmark.

6.3. Welfare loss of the informational friction

We now turn to evaluating the welfare loss of missing information. In our model, the cost of missing information comes in two forms. First, workers do not observe their suitability for their own market. This leads to two forms of inefficiency. Some workers switch out of markets that, had their suitability been revealed to them, they would have stayed in. Other workers stay

[7] Hornstein et al. (2007) report the inverse of this number, which they estimate to be 1.64 in their paper.

in markets they would switch out of if they knew their true suitability. The welfare losses are compounded in general equilibrium. Since workers congest markets they are not suitable for, the incentives for firms to post vacancies are reduced relative to a full information benchmark. This leads to a reduction of the encounter rate, which implies higher unemployment and lower learning rates.

Second, workers also have no information about any other markets besides their own. This provides a strong motive for workers to stay in a market even if their own assessment of suitability is mediocre. The reason is that switching workers have to bear the cost of moving χ, but then cannot target markets with high suitability. Instead, their suitability is a random draw. Even after switching, workers then need time to learn about that draw. Therefore, the outside option in any market consists of a costly transition process instead of a one-time transition into the ideal market, as is the case under full information.

We disentangle the relative contribution of both frictions by comparing three models where we sequentially remove the two information frictions. First, the baseline model which contains worker ambiguity over both their suitability in their own market, and the suitability over jobs in alternative markets. Second, we construct a version of the model in which workers learn about their suitability in their own market as soon as they enter it. In this model, workers are always fully informed about their current market and switching decisions take the form of a threshold rule s^*. If suitability $s < s^*$, workers switch immediately upon entry. If $s \geq s^*$, then workers stay in their market until an η-shock displaces them. We call this model the "partial information model". Details on this model can be found in Appendix A.14. Third, the full information model described in section 4 in which workers are fully informed over their suitability in their own market, and all others. All models are computed with the parameter values given in Table 3.

Total welfare is evaluated in each of the three models. In a model with distorted beliefs, a welfare definition is not obvious. Should one maximize expected or realized benefits? Since there are no distortions in the beliefs in the full information model, to create as fair a comparison as possible we assume a social planner would want to maximize realized benefits. Since workers are risk-neutral, maximizing total welfare is equivalent to maximizing flow values. Total welfare is defined by Ω below. The four terms represent the sum of the flow benefit received by the unemployed, the output produced by the employed, net of vacancy and market switching costs.

$$\Omega := b\overline{u} + \int_0^1 \sum_{n=1}^{N} \int_0^{T^\star(n)} e(n,\tau,s)\mathbb{E}\left[z|z \geq z^\star(n,\tau)\right] d\tau ds - \kappa\overline{v} - \chi\left(\eta + \xi\right)$$

Here, ξ denotes the number of voluntary switches which is equal to $\int_0^1 \sum_{n=0}^{N} u(n, T^\star(n), s)ds$ in the baseline model, $\eta\frac{F(s^*)}{1-F(s^*)}$ in the partial information model and 0 in the full information model.

We find that welfare is 34.4% higher in the full information model than in the baseline and 2.4% higher in the partial information model than in the baseline. This means that the pure informational friction of having to learn about ones own market is substantial but there is an order of magnitude larger welfare loss associated with not knowing ones suitability in a market in advance of switching into it. The large welfare gains from learning about alternative markets resonates with the findings of Belot et al. (2022). In a randomized field experiment in the UK, Belot et al. find inexpensive job advice that informs workers about alternative occupations with high labor demand can increase job finding probabilities by between 20% and 40%.

In our baseline model, workers have a fixed prior about their suitability in a prospective market. In reality, one might think that workers are able to retrieve some signals about their true suitability in a market ahead of market entry. If that is the case, our welfare loss estimate can be understood as an upper bound, which is applicable if no additional information is available. Our estimate of a 2.4% welfare loss that arises purely from having to learn about ones own market is sizable in itself and indicates that workers lose significantly from having to make acceptance and switching decisions under uncertainty.

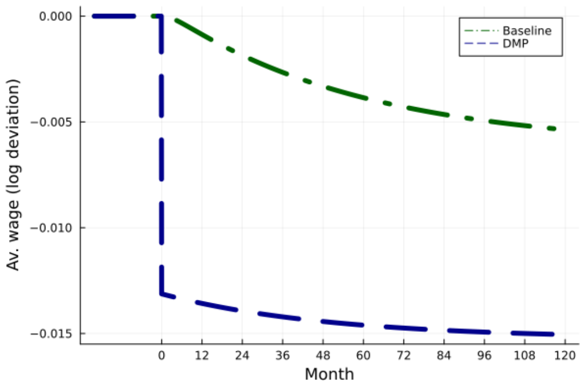
7. Aggregate Shock

The final exercise of the paper is to understand the implications of worker learning on wages and employment over the business cycle. In the calibration of canonical models of macro-labor there exists a dissonance between generating sufficient wage dispersion in the cross-section, and sufficient volatility in employment in the time-series. We have shown in section 6.2 that in the presence of learning we can overcome the problems documented by Hornstein et al. (2011) and generate sufficient frictional wage dispersion without relying on a large and negative replacement rate. When one introduces aggregate shocks into such models a similar obstacle arises. As pointed out by Shimer (2005) these models will ordinarily fail in generating enough volatility in employment over the business cycle. One resolution to the model's lack of propagation is to follow the *small surplus* calibration of Hagedorn and Manovskii (2008) and choose a very high flow value of unemployment, close to the mean productivity of the employed. As we show in this section, in addition to the puzzle of frictional wage dispersion, our model can also reconcile the Shimer puzzle with an intermediate flow value of unemployment calibrated to a replacement rate of 0.4.

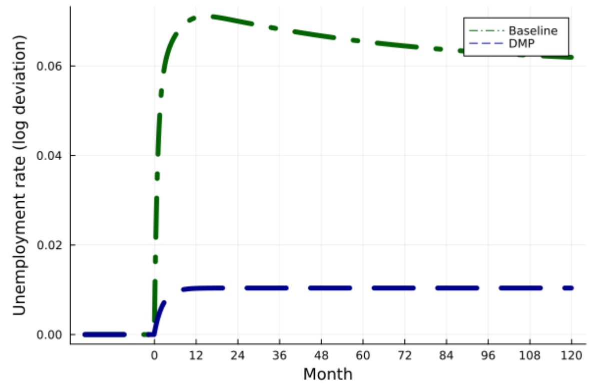
A commonly proposed propagation mechanism is to impose wage rigidity onto the canonical labor search model (see for example Hall (2005), Hall and Milgrom (2008), or Pissarides (2009)). If the wages of new hires are rigid, such that they vary little over the business cycle, then hiring costs remain high in downturns and the wage is unable to smooth firm recruitment over the cycle. Although the baseline model presented is derived in steady-state, one can imagine that in our framework, wage rigidity would emerge naturally out of steady-state. Given an unforeseen shock to the economy, rather than observing it, workers learn through experience about the new state of the world. In the model, wages are the solution to a Nash bargaining protocol. The bargaining outcome depends on beliefs which respond sluggishly, and hence so too will wages. A similar mechanism is present in Menzio (2022). In Menzio's framework a share of workers are fully informed of the aggregate shock and operate under rational expectations. A complementary share have stubborn beliefs, believing aggregate productivity is fixed at its unconditional mean. The existence of agents with stubborn beliefs prevents wages from fully adjusting and propagates employment fluctuations in a way akin to our setting.

We consider a permanent and unanticipated shock to the share of jobs a worker is suitable for. Rather than being able to operate in a share s of all jobs in a market, after the shock, the worker is instead suitable for $(1-\varepsilon)s$ of all jobs. It is assumed that the shock has no impact on your current job. In the baseline model, we assume the shock is not observed, but agents update their beliefs over their new suitability gradually. The scaling property of the gamma distribution

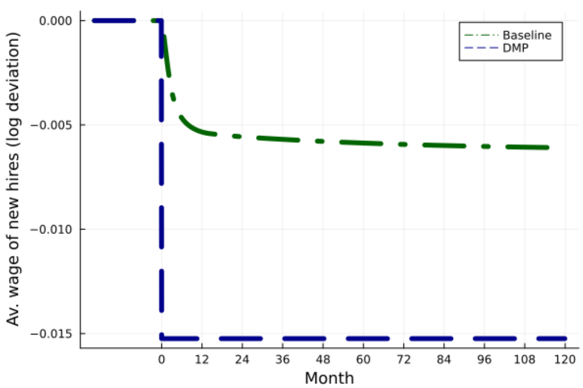
Figure 3: Transition dynamics in response to a one-time permanent shock to suitability



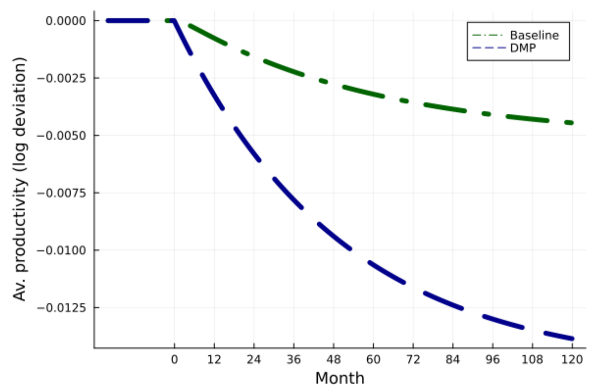
(a) Wages



(b) Unemployment



(c) Wages (new hires)



(d) Productivity

Note: All impulse responses are presented as log deviations from the initial steady-state equilibrium. The suitability shock implies that irrespective of their initial market, following the permanent shock a worker is suitable for 5% less jobs in a market. The baseline model is calibrated according to Table 3 and the full information (DMP) model is calibrated according to Table A.5.

means that when sampling from a new market the distribution of suitability is given by

$$f(s) = \left(\frac{k_0}{\overline{s}^{k_0}}\right) s^{k_0 - 1} \quad \text{for} \quad s \in [0, \overline{s}] \quad \text{where, } \overline{s} = 1 - \varepsilon.$$

In the full information model, since suitability is observable it seems natural that the shock too would also be observable. After a worker is hit by the reallocation shock η, with full information, they will move to the market where they are suitable for most jobs $\overline{s} = 1 - \varepsilon$. Hence, in this setting, a share ε of meetings are of no use to the worker and firm, and the shock is isomorphic to a shock to the matching efficiency parameter of the matching function. Figure 3 considers the impulse response of the benchmark and full information model following a shock of $\varepsilon = 0.05$, at time zero.

The response of the full information model is well understood. Market tightness, vacancies divided by unemployment, jumps immediately to its new lower steady-state level. Consequently, the wages of new hires also move discontinuously to a new lower level. The impact on aggregate wages, unemployment, and productivity is more protracted. Existing matches continuously renegotiate wages, and hence we see an immediate drop on impact in aggregate wages. However, with a larger vacancy cost there are fewer job offers and the unemployed are thus less discerning in which jobs to accept, thus lowering their reservation productivity threshold $z^\star$. This leads to a further fall in aggregate wages and a decline in labor productivity.

By contrast, in the baseline model the new state of the world is slowly revealed to workers through a learning process. Following a hike in the cost of posting a vacancy there are fewer vacancies, because of the free entry condition. However, the reduction in job offers is not immediately apparent to the unemployed who learn slowly from experience in unemployment. As workers learn more about the market, they are willing to accept a lower wage and thus after a point unemployment begins to decline. Since workers are slow to learn, the impact on employment is much larger than in the baseline model. The cost of a firm hiring increases by the same amount in both specifications, but under full information this is somewhat compensated by lower wages, and thus dampening the cycle. In the baseline model, at the point of impact workers are uninformed and hence their perceived employment opportunities are unchanged despite the fall in labor demand. This results in larger changes in employment. The same impact can be seen in productivity. In the full information model workers immediately drop their reservation wage and thus productivity drops substantially. In the baseline model it takes time for reservation wages to adjust, and thus productivity moves far less.

8. Conclusion

This paper presents a tractable model to understand the importance of learning in a search theoretic model of the labor market. A worker's labor market history becomes consequential for decisions in the present. Tractability is retained as one's history can be full characterized by two sufficient statistics. Strikingly, the model reconciles features of the data that are typically missing in macro labor models. In particular, the model consistent with data produces: the pattern of bias implied by worker's labor market expectations; falling starting wages with the duration of an unemployment spell; and the level of cross-sectional frictional wage dispersion.

The cost of misinformation is large. In a hypothetical world in which all agents are fully informed, the economy would grow by approximately one third. Most economic losses originate from workers making decisions under uncertainty that are shaped by limited information about other markets. Thus, large welfare improvements can be made by informing unemployed workers about their outside options. We believe a model of this type can help tailor such an intervention in the labor market, and this could prove to be a fruitful future research agenda. Finally, we show following an unanticipated shock to the economy our baseline model generates an order of magnitude more propagation to employment than a full information economy. Since learning is protracted, wage rigidity arises endogenously and hiring costs do not adjust as much over the cycle, leading to larger employment fluctuations.

References

Fernando Alvarez and Robert Shimer. Search and rest unemployment. *Econometrica*, 79(1): 75–122, 2011.

Erhan Artuç, Shubham Chaudhuri, and John McLaren. Trade shocks and labor adjustment: A structural empirical approach. *American Economic Review*, 100(3):1008–1045, 2010.

Michéle Belot, Philipp Kircher, and Paul Muller. Do the long-term unemployed benefit from automated occupational advice during online job search? *mimeo*, 2022.

Daniel Borowczyk-Martins, Grégory Jolivet, and Fabien Postel-Vinay. Accounting for endogeneity in matching function estimation. *Review of Economic Dynamics*, 16(3):440–451, 2013.

Jake Bradley and Axel Gottfries. A job ladder model with stochastic employment opportunities. *Quantitative Economics*, 12(4):1399–1430, 2021.

Kenneth Burdett and Tara Vishwanath. Declining reservation wages and learning. *The Review of Economic Studies*, 55(4):655–665, 1988.

Kenneth Burdett, Carlos Carrillo-Tudela, and Melvyn Coles. The cost of job loss. *The Review of Economic Studies*, 87(4):1757–1798, 2020.

John Conlon, Laura Pilossoph, Matthew Wiswall, and Basit Zafar. Labor market search with imperfect information and learning. *mimeo*, 2021.

Richard W Cottle and George B Dantzig. Complementary pivot theory of mathematical programming. *Mathematics of the Decision Sciences*, 1:115–136, 1968.

Jason Faberman, Andreas I Mueller, Ayşegül Şahin, and Giorgio Topa. Job search behavior among the employed and non-employed. *Econometrica*, 90(4):1743–1779, 2022.

Armin Falk, David Huffman, and Uwe Sunde. Self-confidence and search. *IZA Working Paper No. 2525*, 2006.

Jean Flemming. Skill accumulation in the market and at home. *Journal of Economic Theory*, 189:105099, 2020.

Francisco Gonzalez and Shouyong Shi. An equilibrium theory of learning, search, and wages. *Econometrica*, 78(2):509–537, 2010.

Marcus Hagedorn and Iourii Manovskii. The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review*, 98(4):1692–1706, 2008.

Robert E Hall. Employment fluctuations with equilibrium wage stickiness. *American Economic Review*, 95(1):50–65, 2005.

Robert E Hall and Paul R Milgrom. The limited influence of unemployment on the wage bargain. *American economic review*, 98(4):1653–1674, 2008.

Andreas Hornstein, Per Krusell, and Gianluca Violante. Frictional wage dispersion in search models: A quantitative assessment. *The Federal Reserve Bank of Richmond, Working Paper Series*, (06-07), 2007.

Andreas Hornstein, Per Krusell, and Gianluca Violante. Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review*, 101(7):2873–2898, 2011.

Gregor Jarosch. Searching for job security and the consequences of job loss. *mimeo*, 2021.

John Kennan and James Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011.

David Kreps. Anticipated utility and dynamic choice. *Econometric Society Monographs, no. 29, Cambridge: Cambridge University*, pages 242–274, 1998.

Thibaut Lamadon, Magne Mogstad, and Bradley Setzler. Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review*, 112(1): 169–212, 2022.

Guido Menzio. Stubborn beliefs in search equilibrium. *NBER Working Paper*, (29937), 2022.

Dale T Mortensen and Christopher A Pissarides. Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415, 1994.

Andreas Mueller and Johannes Spinnewijn. Expectations data, labor market and job search. *Handbook of Economic Expectations, forthcoming*, 2022.

Andreas Mueller, Johannes Spinnewijn, and Giorgio Topa. Job seekers' perceptions and employment prospects: Henterogeneity, duration dependence and bias. *American Economic Review*, 111(1):324–363, 2021.

Victor Ortego-Marti. Unemployment history and frictional wage dispersion. *Journal of Monetary Economics*, 78:5 – 22, 2016.

Christopher A Pissarides. The unemployment volatility puzzle: Is wage stickiness the answer? *Econometrica*, 77(5):1339–1369, 2009.

Tristan Potter. Learning and job search dynamics during the great recession. *Journal of Monetary Economics*, 117:706–722, 2021.

Michael Rothschild. Searching for the lowest price when the distribution of prices is unknown. *Journal of political Economy*, 82(4):689–711, 1974.

Robert Shimer. The cyclical behavior of equilibrium unemployment and vacancies. *American economic review*, 95(1):25–49, 2005.

Johannes Spinnewijn. Unemployed but optimistic: Optimal insurance design with biased belief. *Journal of the European Economic Association*, 13(1):130–167, 2015.

Melvin Stephens. Job loss expectations, realizations, and household consumption behavior. *The Review of Economics and Statistic*, 86(1):253–269, 2004.

Ming Xu. Understanding the decline in occupational mobility. *mimeo*, 2018.

Appendix A. Appendix

Appendix A.1. Sample Restrictions to the SCE

Our baseline sample is taken from the SCE. The sample is a rotating panel in which individuals are surveyed every month for up to 12 months. Attention is restricted to unemployed workers between June 2013 and February 2020. This leaves a sample size of 4,140 which covers 1,392 different individuals. Two further restrictions are imposed on the data and the effect on the sample size is reported in Table A.1. First we restrict individuals to be between the ages of 20 and 65, inclusive. In the data, job seekers report their expected probability that they will be employed within the next three and twelve months. We follow the approach of Mueller et al. (2021) and omit those who either do not respond to either question or report a higher probability of finding a job in three months than in twelve.

Table A.1: Sample Restrictions

	Sample Size	Individual Respondents
Unrestricted Sample	4,140	1,392
Sample size after restriction:		
Age restriction	3,469	1,159
Consistent elicited belief response	2,221	859

Labor Market Supplement

To garner information on the job offers received when unemployed, the labor market survey, a supplement of the SCE is used. The survey is conducted for a subset of respondents every four months. The first in our sample is March 2014 and the final survey is conducted in November of 2019. To have information on job offers the first requirement is that respondents are included in the labor market survey. As is shown in Table A.2 this reduced the sample size by almost one half. The supplement contains a question: "*how many job offers did you receive in the last four months?*" From the unemployed respondents who answer zero, we can identify unemployment spells, or subsets of spells in which workers received no job offers. Inspection of Table A.2 shows that this reduces the sample by almost a further two thirds.

Table A.2: Subsample Restrictions

	Sample Size	Individual Respondents
Restricted Sample	2,221	859
Included in the labor market supplement	1,267	296
Identified as no offer spells	692	211

To review, the full sample in section 2 is the final row of Table A.1, and the no offer sample is the final row of Table A.2.

Appendix A.2. Interacting with Search Intensity

Table A.3 includes a measure of search intensity. For the measure of search intensity we once again rely upon the Labor Market survey, a supplement to the main survey. We define an unemployed worker searching with low intensity if they report fewer than two hours spent on job search activities in the last seven days. Approximately 10% of the unemployed are deemed low intensity under this metric.

Table A.3: Interacting Table 1 with Search Intensity

	Spells without job offers			Interacted with search intensity		
	(1)	(2)	(3)	(1)	(2)	(3)
Low intensity	0.161	0.110		0.013	-0.152	
	(0.106)	(0.127)		(0.189)	(0.236)	
Log unemployment duration months	-0.159***	-0.127***	-0.149**			
	(0.032)	(0.031)	(0.068)			
Log unemployment duration months × low intensity				-0.105	-0.032	0.343**
				(0.066)	(0.079)	(0.172)
Log unemployment duration months × (1-low intensity)				-0.165***	-0.138***	-0.174**
				(0.035)	(0.033)	(0.071)
Demographic controls		x			x	
Spell fixed effects			x			x
Observations	595	592	673	595	592	595
R^2	0.064	0.211	0.764	0.065	0.214	0.732

The specification is as in Table 1 interacted with a measure of search intensity. Attention is restricted to unemployed workers without any job offers in their unemployment spell.

Appendix A.3. Nested Distributions of (As. 1)

Table A.4 lists distributions nested by the truncated gamma distribution used in the exposition of the model.

Table A.4: Nested Distributions

Distribution	Parameter Restriction	Probability Density
Truncated gamma	$\theta_0 > 0, k_0 > 0, \bar{s} \in (0,1]$	$f(s) = \frac{1}{\gamma\left(k_0, \frac{\bar{s}}{\theta_0}\right)\theta_0^{k_0}} s^{k_0-1} e^{-\frac{s}{\theta_0}}$
Uniform $[0, \bar{s}]$	$\theta_0 \to \infty, k_0 = 1$	$f(s) = \frac{1}{\bar{s}}$
Degenerate at $\bar{s}$	$\theta_0 \to \infty, k_0 \to \infty$	$F(s) = \begin{cases} 0, & \text{for } s < \bar{s} \\ 1, & \text{for } s \geq \bar{s} \end{cases}$
Truncated exponential (parameter $1/\theta_0$)	$k_0 = 1$	$f(s) = \frac{1}{\left(1-e^{-\frac{\bar{s}}{\theta_0}}\right)\theta_0} e^{-\frac{s}{\theta_0}}$
Truncated erlang (scale parameter $1/\theta_0$ and shape k_0)	$k_0 \in \mathbb{Z}^+$	$f(s) = \left(1 - \sum_{n=0}^{k_0-1} \frac{1}{n!} e^{-\frac{\bar{s}}{\theta_0}} \left(\frac{\bar{s}}{\theta_0}\right)^n\right)^{-1} \frac{s^{k_0-1} e^{-\frac{s}{\theta_0}}}{\theta_0^{k_0}(k_0-1)!}$
Beta distribution $(k_0, 1)$	$\theta_0 \to \infty, \bar{s} = 1$	$f(s) = k_0 s^{k_0-1}$
Truncated chi squared with $\upsilon \in \mathbb{Z}^+$ degrees of freedom	$\theta_0 = 2, k_0 = \frac{\upsilon}{2}$	$f(s) = \frac{1}{2^{k_0}\gamma(k_0, \frac{\bar{s}}{2})} s^{k_0-1} e^{-\frac{s}{2}}$
Truncated Schulz-Zimm	$\theta_0 = \frac{1}{k_0}$	$f(s) = \frac{{k_0}^{k_0}}{\gamma(k_0, k_0\bar{s})} s^{k_0-1} e^{-k_0 s}$

Note: If the distribution is specified as truncated, the supremum is in each case $\bar{s}$. Each specified parameter restriction is in addition to those specified for the truncated gamma distribution.

Appendix A.4. Proof of Result 1

Given rational expectations, a worker's prior distribution over suitability is given by

$$f(s|k_0, \theta_0) = \frac{1}{\gamma\left(k_0, \frac{\bar{s}}{\theta_0}\right)\theta_0^{k_0}} s^{k_0-1} e^{-\frac{s}{\theta_0}} \quad \text{for } s \in [0, \bar{s}]. \qquad \text{(As. 1)}$$

If a worker is suitable for a proportion S of jobs, the offer arrival rate is $\Lambda = S\bar{\lambda}$. Defining function $\Lambda = g(S) = S\bar{\lambda}$ then by the transformation technique, the density of λ is

$$\begin{aligned} f_\Lambda(\lambda) &= f_S\left(g^{-1}(\lambda)\right)\left|\frac{ds}{d\lambda}\right| \\ &= \frac{1}{\gamma\left(k_0, \frac{\bar{s}}{\theta_0}\right)\theta_0^{k_0}} \left(\frac{\lambda}{\bar{\lambda}}\right)^{k_0-1} e^{-\frac{\lambda}{\bar{\lambda}\theta_0}} \left|\frac{1}{\bar{\lambda}}\right| \\ &= \frac{1}{\gamma\left(k_0, \frac{\bar{s}}{\theta_0}\right)\left(\theta_0\bar{\lambda}\right)^{k_0}} \lambda^{k_0-1} e^{-\frac{\lambda}{\bar{\lambda}\theta_0}} \\ &= \frac{1}{\gamma\left(k_0, \frac{\bar{s}\bar{\lambda}}{\theta_0\bar{\lambda}}\right)\left(\theta_0\bar{\lambda}\right)^{k_0}} \lambda^{k_0-1} e^{-\frac{\lambda}{\bar{\lambda}\theta_0}} \quad \text{for } \lambda \in [0, \bar{s}\bar{\lambda}]. \end{aligned}$$

Appendix A.5. Proof of Result 2

Let X be the duration of time a worker waits for a job offer. Given job offers arrive according to a Poisson process with rate λ. The waiting time for an offer is exponential, and given

Assumption 1, the prior over lambda is uniform on $[0, \overline{\lambda}]$. Thus if a worker has received n offers in the market, and is currently unemployed, they have $n+1$ data points to infer the value of λ with $(x_1, ...x_n, x_{n+1})$. Where for $i = \{1, 2, ...n\}$, x_i is an uncensored job offer waiting time, and x_{n+1} is the time lapsed without a job offer in their current unemployment spell. Implementing Bayes' rule and defining $\boldsymbol{x} := (x_1, ...x_n)$ yields

$$f(\lambda|\boldsymbol{x}, x_{n+1}) = \left(\frac{f(\boldsymbol{x}|\lambda)}{f_{\boldsymbol{x}}(\boldsymbol{x})}\right)\left(\frac{1 - F(x_{n+1}|\lambda)}{1 - F_x(x_{n+1})}\right)\pi(\lambda) \propto \left(\prod_{i=1}^{n} f(x_i|\lambda)\right)(1 - F(x_{n+1}|\lambda))\,\pi(\lambda)$$

where,

$$\begin{aligned}
\prod_{i=1}^{n} f(x_i|\lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \\
(1 - F(x_{n+1}|\lambda)) &= e^{-\lambda x_{n+1}} \\
\pi(\lambda) &= \frac{1}{\gamma\left(k_0, \frac{\overline{s}\overline{\lambda}}{\theta_0\overline{\lambda}}\right)\left(\theta_0\overline{\lambda}\right)^{k_0}} \lambda^{k_0-1} e^{-\frac{\lambda}{\overline{\lambda}\theta_0}} \quad \text{for } \lambda \in [0, \overline{s}\overline{\lambda}] \\
&= 0 \quad \text{otherwise.}
\end{aligned}$$

So,

$$\begin{aligned}
f(\lambda|\boldsymbol{x}, x_{n+1}) &\propto \left(\prod_{i=1}^{n} f(x_i|\lambda)\right)(1 - F(x_{n+1}|\lambda))\,\pi(\lambda) \\
&\propto \lambda^{n+k_0-1} e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0} + \sum_{i=1}^{n+1} x_i\right)} \quad \text{for } \lambda \in [0, \overline{s}\overline{\lambda}]
\end{aligned}$$

Hence the posterior depends on the total duration of unemployment $\tau := \sum_{i=1}^{n+1} x_i$, the number of job offers accumulated n and the parameters governing the prior of workers (θ_0, k_0). Thus rather than having $n+1$ state variables to infer the posterior of λ, the worker need only keep track of the number of offers they have received n and the total length of time spent in unemployment $\tau := \sum_{i=1}^{n+1} x_i$. Rewriting the posterior density in terms of these two sufficient statistics yields

$$f(\lambda|n, \tau) = \mathcal{C}\lambda^{n+k_0-1} e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0} + \tau\right)} \quad \text{for } \lambda \in [0, \overline{s}\overline{\lambda}].$$

Where $\mathcal{C}$ is the constant of integration, to pin down the constant $\mathcal{C}$, we ensure the posterior density integrates to one. Hence,

$$\begin{aligned}
\frac{1}{\mathcal{C}} &= \int_0^{\overline{s}\overline{\lambda}} \lambda^{n+k_0-1} e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0} + \tau\right)} d\lambda \\
\frac{1}{\mathcal{C}} &= \left(\frac{1}{\overline{\lambda}\theta_0} + \tau\right)^{-n-k_0+1} \int_0^{\overline{s}\overline{\lambda}} \left(\lambda\left(\frac{1}{\overline{\lambda}\theta_0} + \tau\right)\right)^{n+k_0-1} e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0} + \tau\right)} d\lambda
\end{aligned}$$

Define $u = \lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)$, $du = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)d\lambda$, and hence

$$\frac{1}{\mathcal{C}} = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-n-k_0}\int_0^{\bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)} u^{n+k_0-1}e^{-u}du$$

The lower incomplete gamma function is defined as,

$$\gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt$$

and hence,

$$\frac{1}{\mathcal{C}} = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-n-k_0}\gamma\left(n+k_0, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right).$$

Finally, substituting the constant of integration back into the posterior density yields

$$f(\lambda|n,\tau) = \frac{1}{\gamma\left(n+k_0, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{n+k_0}\lambda^{n+k_0-1}e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)} \quad \text{for } \lambda\in[0,\bar{s}\overline{\lambda}]$$

This is a truncated gamma distribution with shape parameter $n+k_0$ and scale parameter $1\Big/\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)$.

Appendix A.6. Proof of Result 3

The mean of the posterior for a worker with labor history summarized by τ and n is given by

$$\begin{aligned}\lambda(n,\tau) &= \int_0^{\bar{s}\overline{\lambda}}\lambda f(\lambda|n,\tau)d\lambda \\ &= \frac{1}{\gamma\left(n+k_0, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{n+k_0}\int_0^{\bar{s}\overline{\lambda}}\lambda^{n+k_0}e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)}d\lambda\end{aligned}$$

Define $u = \lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)$, hence $du = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)d\lambda$

$$\begin{aligned}\int_0^{\bar{s}\overline{\lambda}}\lambda^{n+k_0}e^{-\lambda\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)}d\lambda &= \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-n-k_0}\int_0^{\bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)}u^{n+k_0}e^{-u}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-1}du \\ &= \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-n-k_0-1}\gamma\left(n+k_0+1, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)\end{aligned}$$

Substituting the solution of the integral into the mean of the posterior gives,

$$\lambda(n,\tau) = \left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)^{-1}\frac{\gamma\left(n+k_0+1, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)}{\gamma\left(n+k_0, \bar{s}\overline{\lambda}\left(\frac{1}{\overline{\lambda}\theta_0}+\tau\right)\right)} \quad \text{(Re. 3)}$$

Appendix A.7. Derivation of the HJB equation for the unemployed

Assuming that the time interval Δ is small enough such that,

$$\Delta\eta - \int_{\tau}^{\tau+\Delta}\int_{0}^{\infty}\lambda f(\lambda|n,\tau')d\lambda d\tau' < 1$$

Then a discrete time representation of the value function is,

$$\begin{aligned} V_u(n,\tau) = \Delta b + \frac{1}{1+r\Delta}\Bigg[\Bigg(\int_{\tau}^{\tau+\Delta}\int_{0}^{\infty}\lambda f(\lambda|n,\tau')\Big(V_u(n+1,\tau') + \mathbb{E}_z\left[\max\left(\beta S(z,n+1,\tau'),0\right)\right]\Big)d\lambda d\tau'\Bigg) \\ + \Delta\eta V_u(0,0) + \left(1-\Delta\eta - \int_{\tau}^{\tau+\Delta}\int_{0}^{\infty}\lambda f(\lambda|n,\tau')d\lambda d\tau'\right)V_u(n,\tau+\Delta)\Bigg] + o(\Delta) \end{aligned}$$

Where $o(\Delta)$ encompasses terms that go to zero faster than Δ, for example workers getting more than one job offer in a period Δ. Define $\lambda(n,\tau)$ as the mean of the posterior density such that, $\lambda(n,\tau) = \int_0^\infty \lambda f(\lambda|n,\tau)d\lambda$. Hence the HJB equation simplifies to,

$$\begin{aligned} V_u(n,\tau) = \Delta b + \frac{1}{1+r\Delta}\Bigg[\Bigg(\int_{\tau}^{\tau+\Delta}\lambda(n,\tau')\Big(V_u(n+1,\tau') + \mathbb{E}_z\left[\max\left(\beta S(z,n+1,\tau'),0\right)\right]\Big)d\tau'\Bigg) \\ + \Delta\eta V_u(0,0) + \left(1-\Delta\eta - \int_{\tau}^{\tau+\Delta}\lambda(n,\tau')d\tau'\right)V_u(n,\tau+\Delta)\Bigg] \end{aligned}$$

Multiply both sides by $(1+r\Delta)$ and rearranging yields,

$$\begin{aligned} (1+r\Delta)V_u(n,\tau) &= \Delta b(1+r\Delta) + \Bigg(\int_{\tau}^{\tau+\Delta}\lambda(n,\tau')\Big(V_u(n+1,\tau') + \mathbb{E}_z\left[\max\left(\beta S(z,n+1,\tau'),0\right)\right]\Big)d\tau'\Bigg) \\ &\quad + \Delta\eta V_u(0,0) + \left(1-\Delta\eta - \int_{\tau}^{\tau+\Delta}\lambda(n,\tau')d\tau'\right)V_u(n,\tau+\Delta) \\ rV_u(n,\tau) &= b(1+r\Delta) + \frac{1}{\Delta}\left(\int_{\tau}^{\tau+\Delta}\lambda(n,\tau')\Big(V_u(n+1,\tau') - V_u(n,\tau)\Big)d\tau'\right) \\ &\quad + \frac{1}{\Delta}\left(\int_{\tau}^{\tau+\Delta}\lambda(n,\tau')\mathbb{E}_z\left[\max\left(\beta S(z,n+1,\tau'),0\right)\right]d\tau'\right) + \eta\left(V_u(0,0) - V_u(n,\tau)\right) \\ &\quad + \frac{1}{\Delta}\left(1-\Delta\eta - \int_{\tau}^{\tau+\Delta}\lambda(n,\tau')d\tau'\right)\left(V_u(n,\tau+\Delta) - V_u(n,\tau)\right) \end{aligned}$$

Finally, taking the limit as $\Delta \to 0$ gives the HJB equation for the value of unemployment as,

$$\begin{aligned} rV_u(n,\tau) = b + \lambda(n,\tau)\left(V_u(n+1,\tau) - V_u(n,\tau)\right) + \lambda(n,\tau)\mathbb{E}_z\left[\max\left(\beta S(z,n+1,\tau),0\right)\right] \\ + \eta\left(V_u(0,0) - V_u(n,\tau)\right) + \frac{\partial V_u(n,\tau)}{\partial\tau} \end{aligned}$$

Appendix A.8. Rewriting the value function

The Nash bargaining assumption allows us to write

$$\mathbb{E}_z\left[\max\left(V_e(z,n+1,\tau)-V_u(n+1,\tau),0\right)\right]=\mathbb{E}_z\left[\max\left(\beta S(z,n,\tau),0\right)\right]$$

The expected surplus expression can be written as

$$\begin{aligned}
\mathbb{E}_z\left[\max\left(\beta S(z,n,\tau),0\right)\right] =& \frac{\beta}{r+\delta+\eta}\left[\frac{\alpha}{1-\alpha}z^{-\alpha+1}\right]_{\max(z^\star(n,\tau),1)}^{\infty} - \beta\frac{r+\eta}{r+\delta+\eta}V_u(n,\tau)\left(\frac{1}{\max(z^\star(n,\tau),1)}\right)^{\alpha} \\
&+\beta\frac{\eta}{r+\delta+\eta}(V_u(0,0)-\chi)\left(\frac{1}{\max(z^\star(n,\tau),1)}\right)^{\alpha} \\
=& \frac{\beta}{r+\delta+\eta}\Bigg(\frac{\alpha}{\alpha-1}\max\left(z^\star(n,\tau),1\right)^{-\alpha+1}-(r+\eta)V_u(n,\tau)\max(z^\star(n,\tau),1)^{-\alpha} \\
&+\eta(V_u(0,0)-\chi)\max(z^\star(n,\tau),1)^{-\alpha}\Bigg)
\end{aligned}$$

When $z^\star(n,\tau)<1$ and all worker-firm meetings form matches,

$$\mathbb{E}_z\left[\max\left(\beta S(z,n,\tau),0\right)\right]=\frac{\beta}{r+\delta+\eta}\left(\frac{\alpha}{\alpha-1}-(r+\eta)V_u(n,\tau)+\eta V_u(0,0)\right)$$

When the productivity threshold binds and $z^\star(n,\tau)\geq 1$ then

$$\begin{aligned}
\mathbb{E}_z\left[\max\left(\beta S(z,n,\tau),0\right)\right] &= z^\star(n,\tau)^{-\alpha}\frac{\beta}{r+\delta+\eta}\left(\frac{\alpha}{\alpha-1}z^\star(n,\tau)-(r+\eta)V_u(n,\tau)+\eta(V_u(0,0)-\chi)\right) \\
&= z^\star(n,\tau)^{-\alpha}\frac{\beta}{r+\delta+\eta}\left(\frac{\alpha}{\alpha-1}z^\star(n,\tau)-z^\star(n,\tau)\right) \\
&= z^\star(n,\tau)^{1-\alpha}\frac{\beta}{r+\delta+\eta}\left(\frac{1}{\alpha-1}\right) \\
&= \frac{\beta}{r+\delta+\eta}\left(\frac{1}{\alpha-1}\right)\left((r+\eta)V_u(n,\tau)-\eta(V_u(0,0)-\chi)\right)^{1-\alpha}
\end{aligned}$$

Substituting back into the unemployed value function equation (2) yields an expression for $V_u(n,\tau)$ as an expression of model primitives and V_u, n and τ.

$$(r+\eta)V_u(n,\tau) = b+\lambda(n,\tau)\frac{\beta}{r+\delta+\eta}\Bigg[\mathbb{1}(z^\star(n+1,\tau)<1)$$
$$\left(\frac{\alpha}{\alpha-1}-(r+\eta)V_u(n+1,\tau)+\eta(V_u(0,0)-\chi)\right)$$
$$+\,\mathbb{1}(z^\star(n+1,\tau)\geq 1)\left(\frac{1}{\alpha-1}\right)\left((r+\eta)V_u(n+1,\tau)-\eta(V_u(0,0)-\chi)\right)^{1-\alpha}\Bigg]$$
$$+\,\lambda(n,\tau)\left(V_u(n+1,\tau)-V_u(n,\tau)\right)+\eta(V_u(0,0)-\chi)+\frac{\partial V_u(n,\tau)}{\partial\tau}$$

Appendix A.9. Computational appendix

The following outlines the exact computation of the value function for an unemployed worker. The unemployed worker's value function is the one part of the model for which the solution requires some numerical finesse. In contrast, the rest of the model is fairly standard and the numerical details for those parts of the model are not reported here.

Appendix A.9.1. Solving the value function of an unemployed worker

The value function of an unemployed worker is given by

$$rV_u(n,\tau) = \max\Bigg\{\Bigg(b+\lambda(n,\tau)\mathbb{E}_z\left[\max\left(V_e(z,n+1,\tau)-V_u(n+1,\tau),0\right)\right]$$
$$+\,\eta\left(V_u(0,0)-\chi-V_u(n,\tau)\right)$$
$$+\,\lambda(n,\tau)\left(V_u(n+1,\tau)-V_u(n,\tau)\right)+\frac{\partial V_u(n,\tau)}{\partial\tau}\Bigg),$$
$$r(V_u(0,0)-\chi)\Bigg\}$$

We can rewrite this equation as

$$\min\Bigg\{(r+\eta)V_u(n,\tau)-\Bigg(b+\lambda(n,\tau)\mathbb{E}_z\left[\max\left(V_e(z,n+1,\tau)-V_u(n+1,\tau),0\right)\right]+\eta S$$
$$+\,\lambda(n,\tau)\left(V_u(n+1,\tau)-V_u(n,\tau)\right)+\frac{\partial V_u(n,\tau)}{\partial\tau}\Bigg),V_u(n,\tau)-S\Bigg\}=0 \tag{A.1}$$

where $S=[V_u(0,0)-\chi]$. To transform the equation into a fixed point equation solely dependent on V_u and parameters, we further need the following identity:

$$\mathbb{E}_z\left[\max\left(V_e(z,n+1,\tau)-V_u(n+1,\tau),0\right)\right]$$
$$=\begin{cases}\frac{\beta}{r+\delta+\eta}\left(\frac{\alpha}{\alpha-1}-(r+\eta)V_u(n+1,\tau)+\eta S\right) & \text{if } z^\star(n+1,\tau)<1\\ \frac{\beta}{r+\delta+\eta}\left(\frac{1}{\alpha-1}\right)\left((r+\eta)V_u(n+1,\tau)-\eta S\right)^{1-\alpha} & \text{if } z^\star(n+1,\tau)\geq 1\end{cases}$$

where $z^\star(n,\tau) = (r+\eta)V_u(n,\tau) - \eta S$.

Given χ, the solution strategy consists of looping over different outside values S to find the value of S that yields the desired equality $S = [V_u(0,0) - \chi]$. In what follows, we can thus treat S as given.[8]

Within each loop, for a given S, the solution strategy is recursive, starting from a sufficiently high value of $n = \bar{N}$. We assume that, once the a worker reaches $\bar{N}$, both n and τ are fixed forever. By standard arguments, this assumption does not matter as long as $\bar{N}$ is sufficiently large. However, it provides a natural solution for V_u at $n = \bar{N}$. In our application, we choose $\bar{N} = 70$ and verify that a numerically negligible number of agents ever reach this part of the state space. For any τ, $V_u(\bar{N},\tau)$ solves

$$\min\Bigg\{(r+\eta)V_u(\bar{N},\tau) - \Bigg(b + \lambda(\bar{N},\tau)\mathbb{E}_z\left[\max\left(V_e(z,\bar{N},\tau) - V_u(\bar{N},\tau), 0\right)\right] + \eta S\Bigg), \\ V_u(\bar{N},\tau) - S\Bigg\} = 0$$

This yields a solution for $n = \bar{N}$. As shown below, for any n, we can then write the HJBVI for $V_u(n,\cdot)$ as a linear complementarity problem (LCP) given the solution for $V_u(n+1,\cdot)$.

To do this, we start by discretizing and rewriting equation A.1 to fit the following format:

$$0 = \min\{Lv - z, v - S\}$$

Discretizing the differential term in the standard manner (i.e. using finite differences), it is easy to rewrite the problem in this way, where v is a vector capturing the discretized version of $V_u(n,\cdot)$ and any terms involving $n+1$ are treated as known. Concretely, we choose a fine, uniformly spaced grid for τ, which we denote g_τ.[9]

Now, the discretized version of equation A.1 can be written in the above format by setting

$$L = (\rho+\eta)\cdot I - \begin{pmatrix} -\frac{1}{d\tau} - \lambda(n, g_{\tau,1}) & \frac{1}{d\tau} & 0 & \dots & 0 \\ 0 & -\frac{1}{d\tau} - \lambda(n, g_{\tau,2}) & \frac{1}{d\tau} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

and letting z collect all remaining terms.

In a second step, one can rewrite the problem once more as

$$0 = \min\{Lx + q, x\}$$

where $x = v - S$ and $q = -z + LS$. This is the well-known linear complementarity problem (see Cottle and Dantzig (1968)) to which there exist fast solvers.

[8]In practice, when calibrating the model, it can be easier to treat S as a parameter instead of χ, and then set χ accordingly once S has been calibrated. This avoids a loop.

[9]In our application, $g_{\tau,2} - g_{\tau,1} = d\tau = 0.1$, so for example v_1, v_2 and v_3 denote $V_u(n,0), V_u(n,0.1)$ and $V_u(n,0.2)$ respectively. $g_{\tau,3001} = 300$ is our largest grid point. We again verify that only a negligible number of agents ever encounter this boundary.

Casting the problem in this way allows for rapid computation of $V_u(n,\cdot)$ given a solution for $V_u(n+1,\cdot)$ and therefore enables computation of $V_u(\cdot,\cdot)$ on its entire domain, given S. Finding the correct S solving $S = [V_u(0,0) - \chi]$ is then a matter of a simple one-dimensional fixed-point search.

Appendix A.10. Kolmogorov Forward Equations

Since the entry rate in all markets is the same, we can denote it as μ and find μ ex post by using the fact that the population has to integrate to one. Recall that

- Job offers are taken whenever $z \geq z^*(n,\tau)$
- The true job finding rate in a market is given by $\tilde{\lambda}_s$
- Workers exogenously reallocate to new markets at rate η and reallocate endogenously whenever they reach $\tau = T^*(n)$

Denote by $e(n,\tau,s)$ and $u(n,\tau,s)$ the mass of employed and unemployed workers in information state (n,τ) and market m. For given n, $\tau < T^*(n)$, the stationary distribution satisfies:

$$0 = \mathbb{I}(\tau < T^*(n-1))P(z \geq z^*(n,\tau))\tilde{\lambda}_s u(n-1,\tau,s) - (\delta+\eta)e(n,\tau,s)$$
$$0 = \mathbb{I}(\tau < T^*(n-1))P(z < z^*(n,\tau))\tilde{\lambda}_s u(n-1,\tau,s) + \delta e(n,\tau,s) - (\tilde{\lambda}_s+\eta)u(n,\tau,s) - \partial_\tau u(n,\tau,s)$$

with absorbing boundaries at $\tau = T^*(n)$. For $n = 0$ and $\tau < T^*(0)$,

$$\partial_\tau u(0,\tau,s) = -(\tilde{\lambda}_s+\eta)u(0,\tau,s)$$
$$\implies u(0,\tau,s) = u(0,0,s)\cdot\exp(-(\tilde{\lambda}_s+\eta)\tau) = \mu f(s)\cdot\exp(-(\tilde{\lambda}_s+\eta)\tau)$$

For $n \geq 1$, the employment distribution can be characterized as a solution to a set of linear non-homogeneous ODEs given the solution for $n-1$:

$$\begin{aligned} e(n,\tau,s) =& \mathbb{I}(\tau < T^*(n-1))P(z \geq z^*(n,\tau))\frac{\tilde{\lambda}_s}{\delta+\eta}u(n-1,\tau,s) \\ \partial_\tau u(n,\tau,s) =& \mathbb{I}(\tau < T^*(n-1))\left(P(z < z^*(n,\tau)) + P(z \geq z^*(n,\tau))\frac{\delta}{\delta+\eta}\right)\tilde{\lambda}_s u(n-1,\tau,s) \\ &- (\tilde{\lambda}_s+\eta)u(n,\tau,s) \end{aligned}$$

with boundary conditions $u(n,0,s) = 0 \,\forall n \geq 1$. It is easy to solve this by integrating factor method and write this in closed form. For $\tau < T^*(n-1)$:

$$
\begin{aligned}
&(\tilde{\lambda}_s+\eta)e^{(\tilde{\lambda}_s+\eta)\tau}u(n,\tau,s) + e^{(\tilde{\lambda}_s+\eta)\tau}\partial_\tau u(n,\tau,s) \\
&= e^{(\tilde{\lambda}_s+\eta)\tau}\mathbb{I}(\tau < T^*(n-1))\left(P(z < z^*(n,\tau)) + P(z \geq z^*(n,\tau))\frac{\delta}{\delta+\eta}\right)\tilde{\lambda}_s u(n-1,\tau,s)) \\
\implies u(n,\tau,s) &= \int_0^\tau e^{(\tilde{\lambda}_s+\eta)(\tilde{\tau}-\tau)}\left(P(z < z^*(n,\tilde{\tau})) + P(z \geq z^*(n,\tilde{\tau}))\frac{\delta}{\delta+\eta}\right)\tilde{\lambda}_s u(n-1,\tilde{\tau},s)d\tilde{\tau} \\
&\forall n \geq 1, \tau \leq T^*(n-1) \\
u(n,\tau,s) &= e^{-(\tilde{\lambda}_s+\eta)(\tau - T^*(n-1))}u(n,T^*(n-1),s) \quad \forall n \geq 1, T^*(n-1) < \tau \leq T^*(n)
\end{aligned}
$$

This establishes that $u(n,\tau,s)$ (and therefore $e(n,\tau,s)$) can be computed recursively.

Appendix A.11. Wages in the baseline model

$$
\begin{aligned}
(r+\delta+\eta)V_e(z,n,\tau) &= w(z,n,\tau) + \delta V_u(n,\tau) + \eta(V_u(0,0)-\chi) \\
(r+\delta+\eta)\left(V_e(z,n,\tau) - V_u(n,\tau)\right) &= w(z,n,\tau) - (r+\eta)V_u(n,\tau) + \eta(V_u(0,0)-\chi) \\
(r+\delta+\eta)S(z,n,\tau) &= z - (r+\eta)V_u(n,\tau) + \eta(V_u(0,0)-\chi)
\end{aligned}
$$

So, by Nash bargaining,

$$
\begin{aligned}
w(z,n,\tau) - (r+\eta)V_u(n,\tau) + \eta V_u(0,0) &= \beta z - (r+\eta)\beta V_u(n,\tau) + \eta\beta(V_u(0,0)-\chi) \\
w(z,n,\tau) &= \beta z + (1-\beta)\left((r+\eta)V_u(n,\tau) - \eta(V_u(0,0)-\chi)\right)
\end{aligned}
$$

Appendix A.12. Calibration table of the full information model

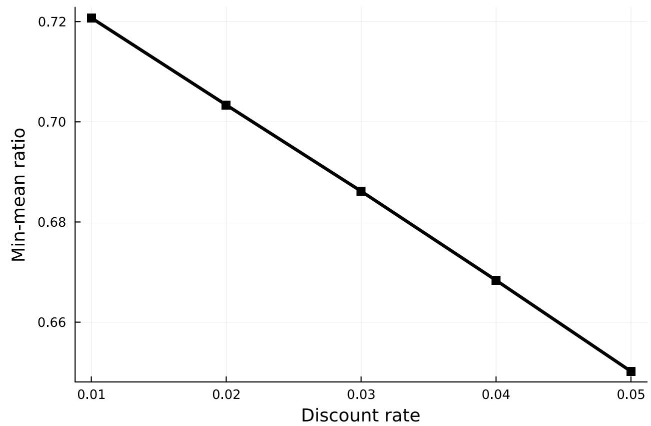
Table A.5: Parameters and calibration targets of the full information model

Parameter	Description	Value	Source/Target
Externally calibrated			
r	Discount rate	0.05/12	5% annual
δ	Separation rate into the same market	$0.02 \times \frac{1}{3}$	2% monthly job loss probability
η	Reallocation rate into the new market	$0.02 \times \frac{2}{3}$	$\approx 2/3$ of separations are worker quits in JOLTS
ω	Matching function elasticity	0.3	Borowczyk-Martins et al. (2013)
A	Matching efficiency	1	Normalization
Internally calibrated			
b	Flow value of unemployment	0.481	40% replacement rate $b/\overline{w}$
α	Pareto parameter of productivity distribution	3.85	0.67 monthly encounter rate.
β	Worker bargaining power	0.106	Wage-productivity pass through of 0.13
κ	Vacancy posting cost	10.13	5% unemployment rate

Note: In the full information model workers are always in their most suitable market, and hence the shape parameter of the suitability distribution k_0 and the reallocation cost χ are immaterial to the equilibrium. Further, the model is not able to generate the duration pattern in beliefs, hence regression moments are omitted.

Appendix A.13. Frictional wage dispersion and the discount rate

Figure A.4: Wage Dispersion and the Discount Rate



Note: The figure plots the min-mean ratio, measures as the mean wage in the economy divided by the minimum wage a worker would conceivably accept. Parameters are calibrated as is reported in Table 3 with the exception of the discount rate which is varied between 1% and 5% per annum.

Appendix A.14. Partial information model

In this section we set out the "partial information model" used in the welfare decomposition of the paper. In this model, a worker is fully informed about their suitability for their own market, but not about the suitability of others, and only learns about that after entering the market. Since the agent is now perfectly informed about the market they inhabit we drop the n and τ state variable and replace with suitability s. Instead of the expected arrival rate $\lambda(n, \tau)$, we can write the true encounter rate $\tilde{\lambda}_s$. The worker and firm value functions are given by:

$$\begin{aligned}(r+\eta)V_u(s) =& b + \tilde{\lambda}_s \max\left(\beta S(z,s), 0\right) + \eta\mathcal{X}\\ (r+\eta)V_e(z,s) =& w(z,s) + \delta\left(V_u(s) - V_e(z,s)\right) + \eta\mathcal{X}\\ (r+\eta)J(z,s) =& z - w(z,s) + \delta\left(0 - J(z,s)\right)\end{aligned}$$

The parameter $\mathcal{X}$ represents the value of having to switch markets. This will be endogenized shortly, but since it enters the value function for unemployment and employment in an isomorphic manner, the surplus function can be written independently of it and is given by.

$$(r+\eta+\delta)S(z,s) = z - b - \beta\tilde{\lambda}_s \int_{\max(z^\star(s),1)} S(z', s)d\Gamma(z')$$

Recall that $\Gamma(z)$ is the cdf of a Pareto with minimum one and parameter α. The surplus generated from a match solves the integral equation above. Notice, there are two potential cases: for some s the reservation productivity may not bind ($z^\star(s) < 1$) and for some s it may bind ($z^\star(s) \geq 1$).

The threshold productivity solves $S(z^\star(s), s) = 0$ and solves the fixed point

$$z^\star(s) = b + \beta\tilde{\lambda}_s \left(\frac{1}{r+\eta+\delta}\right)\left(\frac{1}{\alpha-1}\right) z^\star(s)^{1-\alpha}.$$

The free entry condition is standard, and $\tilde{\lambda}_s$ is pinned down given the Cobb-Douglas constant returns to scale matching function, such that the value in posting a vacancy is zero. What is non-standard, is after an η-shock an unemployed worker must decide whether or not to stay in their new market. The decision can be written as.

$$\tilde{V}_u(s) = \max\{V_u(s), \int_0^{\overline{s}} \tilde{V}_u(s')dF(s') - \chi\}$$

After joining the new market, the suitability of jobs is observable. Thus, is a worker chooses to stay where they are they receive value $V_u(s)$, the first element of the max operator. Leaving the market, and resampling a new market incurs a cost χ and drawing a new suitability, where $F(\cdot)$ is the cdf of suitability across market, given by (As. 1) in its most general form. It is the second term of the max operator that replaces $\mathcal{X}$ in the value functions earlier. The solution to the optimization problem is again a reservation strategy. A worker will stay in a market if their suitability exceeds $s^\star$, and immediately leave if their suitability is less than $s^\star$. The threshold $s^\star$ is the solution to

$$\chi = \int_{s^\star}^{\overline{s}} (1 - F(s))dV_u(s).$$