

Long-Term Unemployment: Attached and Mismatched?

David Wiczer*

Department of Economics, Federal Reserve Bank of St. Louis

Draft: March 23, 2013

Please see sites.google.com/site/davidwiczer for an updated version

Abstract

The rate of long-term unemployment spiked during the Great Recession. To help explain this, I exploit the systematic and counter-cyclical differences in unemployment duration across occupations. This heterogeneity extends the tail of the unemployment duration distribution, which is necessary to account for the observed level of long-term unemployment and its increase since 2007. This paper introduces a model in which unemployment duration and occupation are linked; it measures the effects of occupation-specific shocks and skills on unemployment duration. Here, a worker will be paid more for human capital in his old occupation but a bad shock may make those jobs scarce. Still, their human capital partly “attaches” them to their prior occupation, even when searching there implies a longer expected duration. Hence, unemployment duration rises and becomes more dispersed across occupations. Redistributive shocks and business cycles, as in the Great Recession, exacerbate this effect.

For quantitative discipline, the model matches data on the wage premium to occupational experience and the co-movement of occupations’ productivity. The distribution of duration is then endogenous. For comparison’s sake, if a standard model with homogeneous job seekers matches the job finding rate, then it also determines expected duration and understates it. That standard model implies just over half of the long-term unemployment in 1976-2007 and almost no rise in the recent recession. But, with heterogeneity by occupation, this paper nearly matches long-term unemployment in the period 1976-2007 and 70% of its rise during the Great Recession.

*Special thanks to Victor Ríos-Rull and Fatih Guvenen for continued advice, support and chalkboard time. Thanks also to Alex Bick, Jonathan Heathcote, Dirk Kreuger, Iourii Manovski, Ellen McGrattan, Amanda Michaud, Josep Pijoan-Mas, Kjetil Storesletten, Laura Sunder-Plassmann, the participants of The Chicago Fed Search and Matching Workshop, The Penn Macro Club, The University of Illinois Summer Workshop, and Minnesota Macro-Labor Workshop. Great thanks to Ludo Visschers for such detailed discussion on Carrillo-Tudela and Visschers (2013) and helping me understand his paper and my own. All errors are my own.

1 Introduction

During recessions, there is a fall in the average rate at which unemployed workers find jobs. This implies an increase in unemployment duration and the share of long-term unemployed. However, the observed rise in unemployment duration during recessions outpaces the fall in the average finding rate. In general, if all unemployed workers find jobs at the same rate, then the duration distribution is exponential and this understates the observed distribution, which has a fatter tail. If this single finding rate matches the average finding rate, it understates the share of long-term unemployed by almost a half. Cyclically, the average finding rate implies only $\frac{3}{4}$ of the time-series standard deviation of unemployment duration. Since 2007, the fall in the average finding rate captures less than $\frac{2}{3}$ of the rise in long-term unemployment, which more than doubled. Instead, one must allow for heterogeneity—some workers will take much longer to find a job than others.

In this paper, I incorporate occupations into an otherwise standard search and matching model, à la Pissarides (2000). In it, jobs require occupation-specific skills and their productivity is affected by occupation-specific shocks. Searchers whose occupation is suffering cyclically lower hiring face a simple trade-off: jobs in their own occupation may be more difficult to get, but they will benefit from their occupation-specific skills and earn a higher wage if they find one. By choosing to remain “attached” to their prior occupation and concentrating their search on these types of jobs, workers extend their unemployment duration. I show that the data is consistent with this mechanism; there are sizeable differences in the job finding rate conditional on one’s prior occupation and, these dispersion is counter-cyclical. Finally, set the model’s state to match counter-parts in the Great Recession and then test whether it can replicate the distribution of unemployment duration in this period.

As in the data, the model generates counter-cyclical dispersion in finding rate and duration between occupations. This is because occupations differ in their cyclical sensitivity. Hence, some occupations fare better than others in a recession and so those attached to the wrong occupation will suffer much longer unemployment spells compared to searchers skilled in other occupations. This makes the tail of the distribution of unemployment rate extend by even more than would be implied by the fall in the average finding rate. Can my mechanism account for the rise in long-term unemployment observed during the Great Recession and recessions generally? To put this differently, if recessions affect searchers differently depending upon their prior occupation, how much of the rise of long-term unemployment can be accounted for by searchers whose occupations suffer greatly during a recession?

To answer this quantitative question, I estimate the wage loss associated with switching occupations after a separation. This quantifies the motive for workers to search within their

own occupation. On the other side, I measure the occupational productivity that drives fluctuations in the hiring demand in each occupation. The model then predicts how much searchers will favor their old occupation and how much this will increase their unemployment.

I use data on workers employment, occupation and wage histories in the Survey of Income and Program Participation (SIPP) from 1996-2007 to estimate the fall in wages workers will face if they begin a job in a new occupation. Because the wages are only observable from occupations optimally chosen by searchers, I take a structural approach to this estimation. The same endogenous selection biases exist in model-generated data, so I use indirect inference to estimate the structural parameters governing wage loss.

To estimate the process of occupation-specific shocks, I again employ a structural approach. Because workers who change occupations lose human capital and are less productive, observed output per capita is endogenous. I use indirect inference with data on observed output per capita to estimate the joint process for exogenous shocks to occupational productivity. The specification of these shocks prescribes that “mismatch” is a *cyclical* phenomenon: shocks are mean-reverting and the increase in dispersion comes largely from heterogeneity in their cyclical sensitivity.

This basic mechanism—some unemployed searchers are “mismatched,” their skills do not resemble those used in the more highly productive occupations—helps explain why unemployment duration rises more in recession. This effect increases unemployment duration beyond that which can be explained by a single matching rate. Since 1976, standard deviation of the log of average unemployment duration was 0.24. A standard search and matching model, as presented in Pissarides (2000), predicts only 5% of this volatility. In contrast, my model predicts two-thirds.

Important for realistic long-term unemployment, the cross-sectional standard deviation of duration is counter-cyclical. In the data its correlation with average unemployment is 0.37. In the model, the correlation is 0.6, while any model with a single rate generates no dispersion. Dispersion is counter-cyclical because occupations’ productivity differs in cyclically sensitivity and adjustments costs are asymmetric. In other words, the dispersion in productivity increases symmetrically in high and low ebbs of the cycle but adding labor is slower than shedding it.

This cyclical asymmetry also relates to a very important intuition about the cyclical-ity of the value of skills. In recessions the value of skills rises, and this makes it harder to switch occupations. The reasoning is two fold: non-linearity in the returns which is induced by the less-volatile outside options and endogenous separation rates that are higher for unskilled workers. When an aggregate productivity shock hits, both experienced and inexperienced loose productivity symmetrically, but the total size of the match surplus was

always smaller for the inexperienced workers; hence, their probability of separation rises by more and proportionately because their outside option in unemployment is still

As a side benefit, the model delivers a “hook” in the Beveridge curve. The efficiency of matching is endogenous and procyclical, a correlation of 0.44 with average productivity. So the probability an average job applicant finds a match is lower in recession than would be predicted by the average number of vacancy postings. When the matching efficiency falls in recession, that shifts out the Beveridge curve, which also appears as an upward hook. All of this occurs because workers can choose where to search, and by searching more where there are relatively few vacancies, as in their own occupation, they lower the average matching efficiency.

Finally, I use this model to measure how much of the drastic increase in duration during The Great Recession can be attributed to the mismatch between unemployed workers’ occupation-specific skills and the set of occupation-specific productivity shocks. To begin this exercise, I take as given the state of the economy: (i) the distribution of unemployed workers split by prior occupation in 2007 and (ii) the history of aggregate and occupation-specific labor productivity shocks from 2008-2010. The model then determines the unemployment dynamics. In this period, the rise in long-term unemployment duration was particularly striking, an 88% increase. The model generates a 70% rise, whereas the average finding rate from the data implies only half of the increase and a simple Mortensen-Pissarides model implies almost no rise.

To consider policy, I am currently computing a version in which unemployment benefits are extended from 6 to 24 months, as happened in the Great Recession recession. The model is also uniquely suited to consider directed interventions, which may be more cost effective by targeting badly occupations.

The rest of the paper proceeds as follows: in Section 2 I review related literature and then discuss some motivating data on unemployment duration in Section 3. In Sections 4 and 5 I describe the model and my quantitative strategy, respectively. A discussion of its business cycle properties follows in Section 6. I test it with data from the Great Recession in Section 7 and then conclude in Section 8.

2 Related Literature

To understand the nature of a recession, I borrow from a long literature on counter-cyclical risk. Lilien (1982) and Abraham and Katz (1986) present competing views for why the dispersion of employment growth across sectors should widen in a recession. Whereas Lilien (1982) and many others since have speculated that the variance of idiosyncratic shocks is

itself stochastic and counter-cyclical, the other side attributes counter-cyclical dispersion to differences in the cyclical sensitivity. In some contexts, endogeneous variables such as unemployment and vacancies could be used to discriminate the two specifications, but Hosios (1994) casts doubt, showing that dispersion shocks may entail vacancy and unemployment co-movements very much like aggregate shocks. His results shows why a asymmetric shocks may bring very large changes in unemployment, as we have observed. This paper takes a specification that most closely follows Abraham and Katz (1986), but also incorporates some stochastic dispersion components via a process of unobservable factors. My heterogeneity in cyclical sensitivity means that productivity dispersion responds symmetrically to expansion and recessions. However, the dispersion across occupations in labor variables—employment growth, unemployment rate, and unemployment duration—is counter cyclical due to asymmetries in the structure of the model.

In the Great Recession, several studies have approached to what extent it was asymmetric and what are the ramifications of this asymmetry. On the one hand, Lazear and Spletzer (2012) are skeptical that unemployment is “structural,” but acknowledges that there is industrial asymmetry in the affect of the recession and that “mismatch” rose during the recession. Hobijn (2012), on the other hand, posits that the peculiar composition of new job postings across occupation and industry has significantly slowed the number of successful new matches in the recession and recovery. In this camp, Mehrotra and Sergeyev (2012) use factor analysis to isolate shocks that affect only certain sectors, a technique used in this paper as well. They find strong evidence that the recession affected different sectors differently. Finally, Sahin et al. (2012) try to connect the unemployment rate to the asymmetry of vacancy postings and unemployed workers across occupations. They find the effect of “mismatch” is at most $\frac{1}{3}$ of the increase in unemployment in the Great Recession.

My model builds on the “islands” structure of Lucas and Prescott (1974), who introduce this basic trade-off faced by agents in my model: an unemployed worker must choose whether to stay or go. I have some common ancestry with Alvarez and Shimer (2011) and Carrillo-Tudela and Visschers (2013) in that all extend Lucas and Prescott (1974) to include within-islands unemployment. In Carrillo-Tudela and Visschers (2013) as here, unemployment occurs because of search and matching frictions as in Pissarides (2000) as well as between island unemployment. Workers may be unemployed because they are mis-allocated across islands or because of search frictions that exist on all islands. In both papers workers develop specific human capital that affects their probability to search elsewhere and both have non-trivial distributions of finding rates that are lower among longer unemployed workers. Compared to Carrillo-Tudela and Visschers (2013), the nature of shocks and market structure is quite different. I include occupation-specific shocks and search islands are segmented by

occupation and prior occupation, whereas their islands are more diverse, labelled by human-capital and match quality, and non-aggregate shocks are to match quality. I also model endogenous separations differently from them, as I follow den Haan et al. (2000). And whereas unemployed workers search directionally in my model, theirs may stay or sample the distribution. Finally, the dynamics of the islands in my model is quite different from theirs. Their islands can shut down leading to “rest unemployment,” in which workers wait for an island to reappear, and contributes significantly to unemployment in their model but not in mine. Given these differences in the theoretical structure, our mappings to data are also quite different, e.g., the reallocative force in my model, occupational shocks, come from productivity, whereas they use workers’ flows and the distribution of finding rates.

Occupation-specific human capital plays a crucial role in my model, as it is the source of heterogeneity amongst the unemployed. Ljungqvist and Sargent (1998) also study how searchers may be unemployed for longer because of their human capital from their prior job. They focus on how the secular increase in “turbulence,” which they map to transitory income shocks, implies a long-term increase in unemployment duration in Europe. In their mechanism, the long-term unemployed are separated workers whose skills are lost (antiquated) and chose to enjoy their relatively high unemployment benefit. Their job-finding rate is low because their unemployment benefits increase their outside option. On the other hand, skilled unemployed in my model are still suitable employees in their own occupation. They have relatively low job-finding rate outside their own occupation because they have a high outside option in another sector and because they rarely search in these other labor markets. However, workers in my model are still quite suitable in their own occupations. These two models will deliver starkly different policy implications regarding targeted spending and changes to unemployment benefits, both of which will be explored in greater detail.

In my framework, as in Ljungqvist and Sargent (1998), unemployment duration and finding rate are negatively-correlated because of composition effects. For both, workers enter unemployment with characteristics that lower their matching probability and, definitionally, are a larger fraction of the long-term unemployed than the rest of the unemployed workers. There is a long tradition of papers analyzing the negative correlation between unemployment duration and job finding rate. Clark and Summers (1979), Machin and Manning (1999) and Elsbey et al. (2008) all chronicle this feature in the US and Europe in both expansions and recessions. According to this literature, the correlation is remarkably robust, but its cause is more difficult to discern. Heckman (1991) describe the econometric task of identifying duration dependence, when an individual’s finding rate falls because of the duration of his unemployment spell, or composition. Heckman and Singer (1984) establishes some separability conditions under which identification can be unravelled, and Heckman (1991) expands

on this further. Starkly, my model is going to abstract from duration dependence and will study only the composition effect generated by the mechanism.

The empirical side of this study borrows from a literature on earnings dynamics and occupational choice. The focus on occupation-specific human capital is, in large part, justified by work such as Kambourov and Manovskii (2009), which emphasizes its importance in wage determination. The occupation choice process built into my equilibrium model will bear strong resemblance to the conditional logit model estimated by Boskin (1974). Finally, Altonji et al. (2009) also use indirect inference to estimate earnings in a richer environment but with some of the same complications. They consider a model of wage determination with the discrete choice to switch jobs. To smooth over this discrete choice, they use a logit-like model first proposed in Smith and Keane (2004), a feature already present in the structure of my model.

3 Descriptive Data on Unemployment Duration

In the Great Recession, unemployment duration¹ rose unprecedentedly, as shown in Figures 1 through 3. Notably, this rise went beyond that expected by a uniform fall in the job finding rate. In Figure 2, I plot the fraction of long-term unemployed. Complementary, Figure 3 plots the true unemployment duration and that which is implied by single finding rate that matches outflows from unemployment. The single finding rate’s duration is always lower than the true duration, because some workers have exceptionally long durations, which pulls up the average.

Essentially, all this discussion boils down to Jensen’s inequality. To illustrate, suppose individuals have finding rate rate f_t^i and f_t^* is the finding rate that matches the aggregate monthly flow from unemployment to employment. This is to say $f_t^* = 1 - \frac{u_{t+1} - u_{t+1}^{<1}}{u_t}$, where $u_{t+1}^{<1}$ are the unemployed for less than one period, i.e. the newly separated workers. As pointed out by Shimer (2012), whether $f_t^i = f_t^* \forall i$ or not

$$\int_i f_t^i di = f_t^* = 1 - \frac{u_{t+1} - u_{t+1}^S}{u_t}$$

¹Everywhere in this paper, unemployment duration refers to the expected time before a math into employment. The BLS defines unemployment duration as the average time unemployed of the current pool of unemployed people. This however, conflates slow finding rate with the inflow rate of unemployed. If there are many newly unemployed people, this will tend to *reduce* duration by their measure.

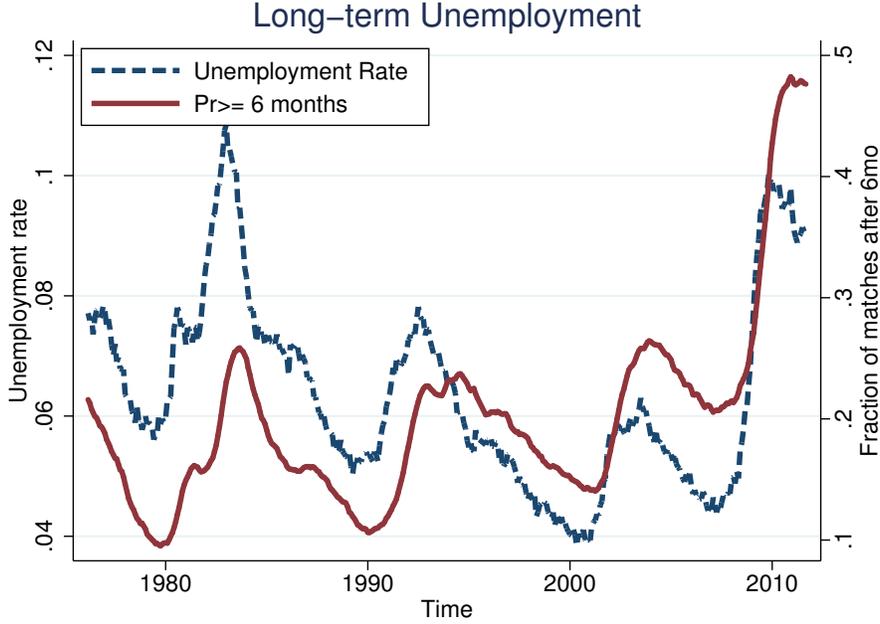


Figure 1: The unemployment rate (blue) and fraction of unemployed whose expected duration is ≥ 6 mo.

However, the expected duration with heterogeneous finding rate will be quite different:

$$\int_i \frac{1}{f_t^i} di \geq \frac{1}{f_t^*}$$

Many others have rejected that $f_t^i = f_t^* \forall i$ in the US and elsewhere, see e.g. Machin and Manning (1999) or Elsby et al. (2008)². I use the unemployment duration question in the CPS to construct a non-parametric Kaplan-Meier estimator for the hazard rate from unemployment to employment. For individuals of duration d , the estimator f_t^d , is the cumulative hazard over a month for finding employment. As in Clark and Summers (1979) I count those who leave the labor force as if they were not in the at-risk population

$$f_t^d = \frac{F_{t+1}^d}{u_t^d - I_{t+1}^d} \quad (1)$$

Where u_t^d are the unemployed in duration bin d , F_{t+1}^d are the subset who find employment in the next period and I_{t+1}^d are those who leave the labor force in period $t + 1$. As can be seen in Figure 4, there is considerable heterogeneity in finding rate. Moreover, in the Great

²It is often the practice to convert monthly rates into continuous time rates. I leave all finding rates at a monthly interval for consistency with this paper's model, which will be discrete time with month-long periods.

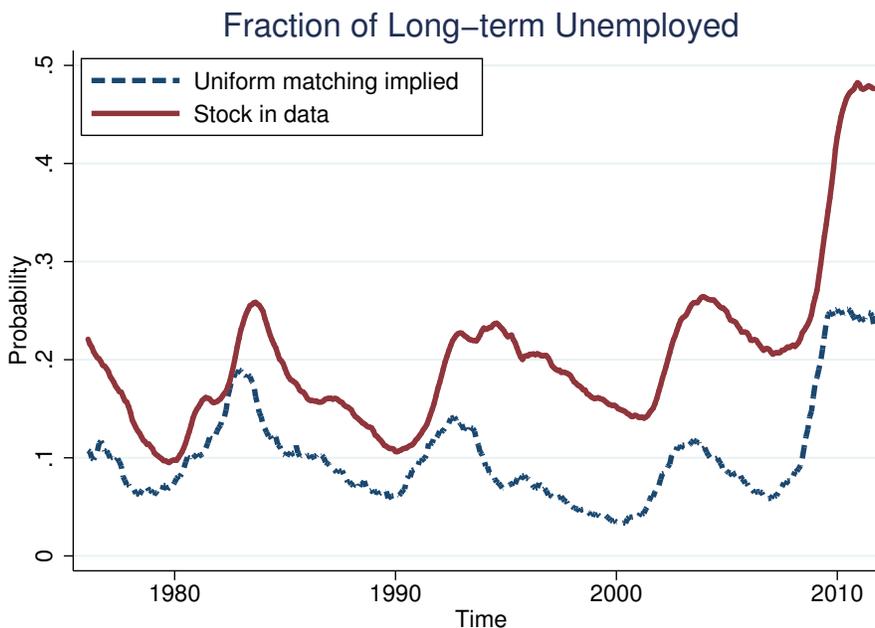


Figure 2: The probability of being unemployed for at least 6 mo.s before finding a job

Recession there was a large fall in the finding rate and especially for longer-durations.

These fluctuations at long durations are borne out in Figures 3 and 1, which were constructed using the durations implied by the estimator in Equation 1. These show that the finding rate at long durations can greatly affect unemployment duration over the cycle.

3.1 Unemployment Duration and Occupation

This paper will link the apparent heterogeneity in finding rates to differences in prior occupation. Further, I will show that this heterogeneity is affected by business cycles. While other studies, e.g. Hornstein (2012), attribute differences in finding rate to inherent and unobservable heterogeneity, the connection to prior occupation has several advantages. Most importantly, my take is more “structural” in attributing causality to agents’ primitive choices in which the incentives are observable. If observed dispersion in finding rate is due to attachment to one’s prior occupation, then this heterogeneity is not policy neutral, as implied by estimating fixed effects to an individual’s finding rate. Moreover, unemployment duration fluctuates greatly and differently in different cycles. With a more structural interpretation, we can predict what conditions affect this. In particular, duration will increase if occupation-specific skills become more important or there is more mismatch between highly productive occupations and the skills of the unemployed.

Why is prior occupation a good dimension on which to separate people? A good deal

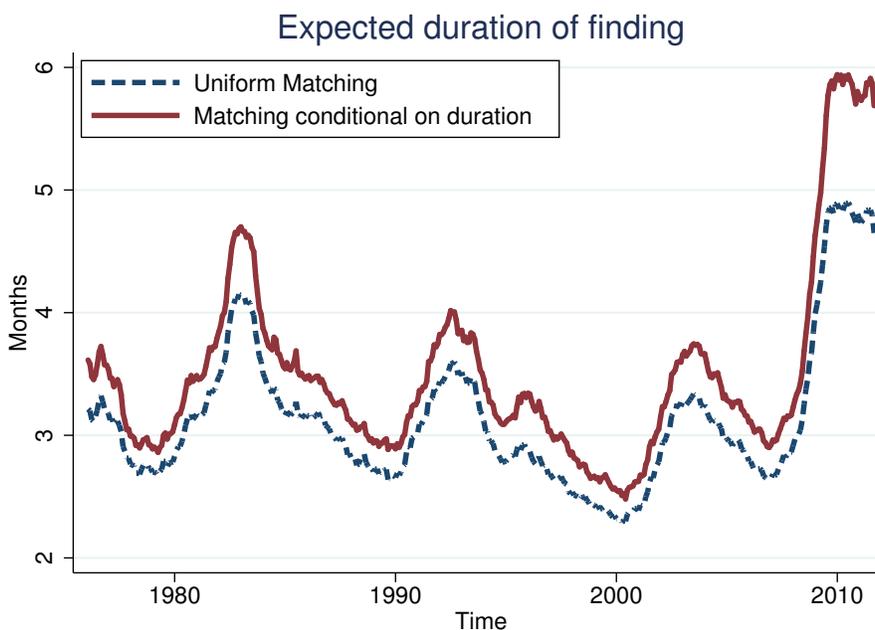


Figure 3: The mean duration before finding a job

of scholarship has been devoted to the importance of occupation-specific experience and skills. Kambourov and Manovskii (2009) were very influential in highlighting the returns to occupational tenure as being larger than other forms of tenure, such as employer or industry tenure. In their baseline, they attribute to occupational tenure a 5-year return between 12-20%. Sullivan (2010) adds some nuance to this result, pointing out that on average occupational tenure has higher returns than industry tenure, but there are some jobs in which industry matters more, e.g. management. He uses NSLY data, rather than PSID in Kambourov and Manovskii (2009), but still find average 5-year returns between 15-25%. Occupational returns are between twice and eight times as high as industry returns when both are included in the regression. All of these studies have to deal with two significant problems that make estimates of the return less certain. First, adding experience in occupation, industry and employer to a simple wage regression is fraught with problems of endogeneity because if the unobservable quality of the match is better it will tend to last longer. Altonji and Shakotko (1987) present one instrumental strategy but this is not without its problems, discussed later in the quantitative section. The weight of the evidence, however, points to occupation-specific skills being especially important.

Unemployment is observably different depending on one's prior occupation. As can be seen in Figures 5(a) and 5(b), unemployment differs across occupations both in its length

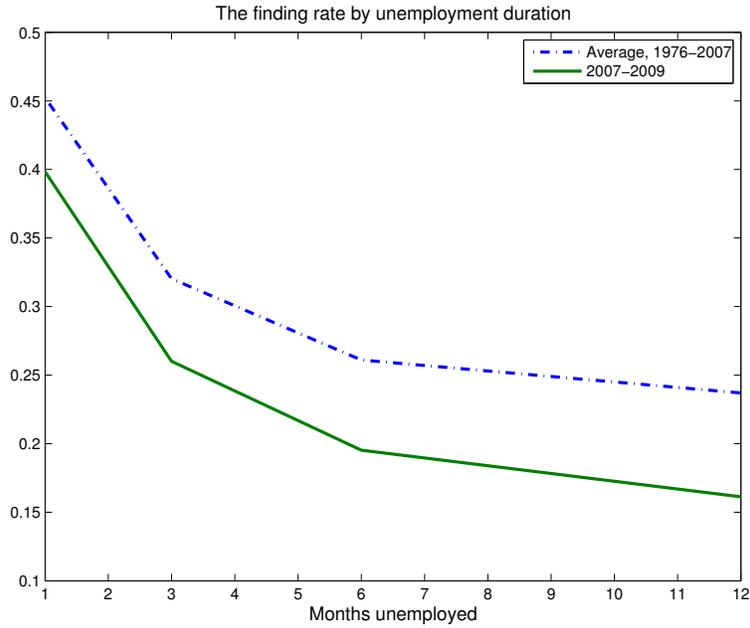


Figure 4: The job finding rate falls as the duration increases

and breadth. The left are the differences in unemployment rate across occupations³. Note that this dispersion is generally counter-cyclical. The correlation between the aggregate unemployment rate and the log of unemployment across occupations is 0.3.

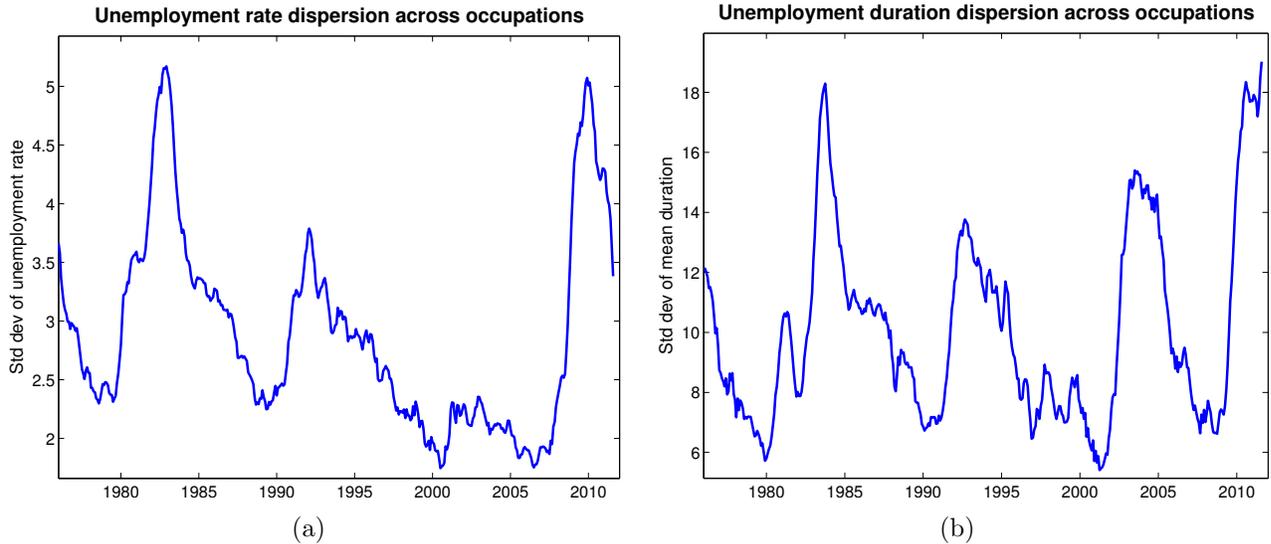


Figure 5: Dispersion across occupations is counter-cyclical

Even more important for this study, Figure 5(b) shows that the expected duration of

³For occupation j , this is the number of unemployed whose prior occupation was j divided by those working and unemployed that last worked in j .

	1978-2007		2008-2010	
	Coef	Std Err	Coef	Std Err
Dummy Occ Switch	8.68	(0.294)	18.98	(0.844)
Dummy Ind Switch	4.53	(0.293)	4.44	(0.839)

Table 1: weeks unemployed = $D_{occ \neq occlly} + D_{ind \neq indly} + \text{demographics} + t$

unemployment varies across occupations. The observed dispersion in duration is consistent with unemployed workers who are attached to their prior occupation. Otherwise, if all workers abandoned their old skills and minimized their duration, every unemployed worker could have the same finding rate. Just like unemployment dispersion, the dispersion in duration is counter-cyclical: if attachment differentiates searchers, recessions exacerbate these effects. It is slightly less cyclical than unemployment dispersion; the correlation between the unemployment rate and log-duration is 0.37.

Looking more closely at the micro-level data behind counter-cyclical dispersion, note that unemployment duration is longer for those that switch occupations and that this effect increased markedly in the Great Recession. In Table 1, I present the results of a simple regression of unemployment duration on occupation and industry switching dummies.

So, how much of the downward-sloping employment finding rate can be explained purely by occupation-level heterogeneity? Figure 6 takes the mean finding rate within a given occupation since 1976 and then constructs the hazard rate at durations 1, 3, 6 and 12 months by taking the population weighted average⁴ finding rate over occupations. It is important to emphasize that this is not a complete answer to the central question of this paper. Figure 6 considers the finding rate for individuals within occupations without considering the endogenous choice they are making whether or not to switch occupations and change their finding rate. Rather, it takes as given what they choose to do. If workers from an occupation are more less likely to abandon their occupation than merely justified by their desire to preserve their skills, e.g. one's occupation has utility value, then the data will have even more heterogeneity. Conversely, if workers are more willing to switch for whatever reason, then occupation-level hazard rates will understate the heterogeneity in finding rates due to occupation-specific human capital.

Taking the mean finding rate within an occupation misses much of the heterogeneity, especially of the very fast job finders. But, it still incorporates the downward slope because some occupations have, on average slower finding rates. And hence, this level of heterogeneity introduces more long-term unemployment than with a constant hazard. The Great Recession

⁴Weights here are by the fraction of the unemployed population at that duration. Hence, if x_ℓ is the fraction of workers initially from a given occupation, those in the at-risk population by duration d are $x_\ell e^{-f_\ell d}$

seemed to cause a near parallel shift down for the hazard function, rather than rotating it as we see with full heterogeneity. This shows that pure occupation-level heterogeneity cannot explain all of the rise in long-term unemployment, though it picks up some.

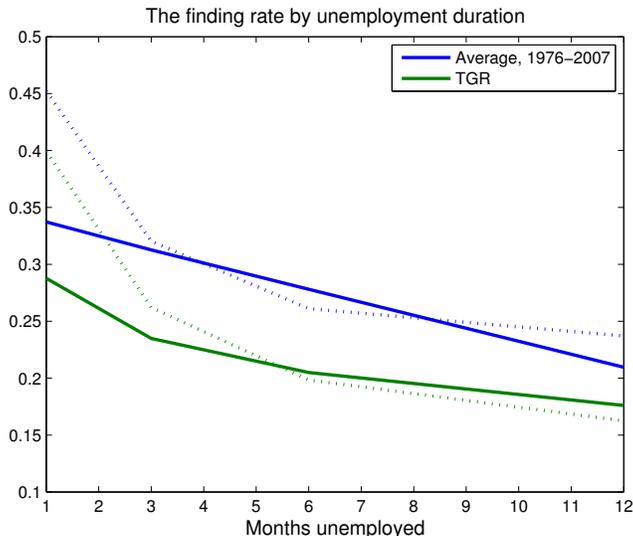


Figure 6: Finding rate using only occupation-level heterogeneity. Dotted lines are the actual finding rates

While Figures 5 and 6 are illustrative that unemployment is different across occupations, they are merely illustrative. As already alluded, the finding rate by occupation includes multiple confounding phenomena aside from the mechanism in question, the effect of human capital. I have already mentioned that search behavior systematically differs across occupations for reasons other than only those implied by human capital. Also, occupations may experience unemployment differently because of unobservable characteristics associated with occupation choice. For example, some occupations may attract older or younger workers who generally have different finding rates. These other factors will affect the occupation-level heterogeneity we observe but are not really attributable to the occupation-specific skills.

More directly commenting on Figure 6, this figure does not incorporate all of the heterogeneity in finding rate that is attributable to human capital and will be in the model. Very basically, even given the potential wage loss from switching, some will switch and some will not. This means that the finding rate among those skilled in an occupation will vary and hence the slope of the finding rate curve may be steeper than Figure 6 that considers only between-occupation differences.

To isolate and quantify the effect on duration of one's attachment over the business cycle, I will estimate the skill loss associated with moving between occupations and how productivity evolves jointly across occupations. I now present my model, which joins these

two forces, one pulling workers away from their occupation and the other keeping them in place.

4 The Model

I will present a model of directed search, where the search markets are occupations. Occupations experience productivity shocks and moving across occupations incurs a cost. The model is designed such that parameters can be calibrated and so that the costs of moving across occupations and the process for occupation-specific shocks can both be structurally estimated. Then, the model will address:

1. Are business cycle fluctuations in unemployment and duration due to mismatch?
2. Can the productivity changes and skill distribution in the Great Recession explain long-term unemployment?

To address the first question, I analyze simulated histories from the model feeding in productivity processes and occupation switching costs that match the data. For the second, I begin with a pool of unemployed whose prior occupation matches that seen at the beginning of the Great Recession. I feed in shocks that match those observed in this recession and use the model to generate fluctuations in finding rates.

The model highlights the basic trade-off faced by agents: whether to stay and search narrowly within their own occupation or leave and search more broadly. In the former, there will be generally a lower finding rate, but their return is higher. By switching, they get a lower wage upon matching but it may increase their chance of getting a job. This is described graphically in Figure 7.

When the economy is hit by shocks this behavior affects unemployment duration through two channels. First, because occupations differ in their cyclical sensitivity, during a downturn workers from harder hit occupations have particularly low finding rates. Second, because separations are endogenous and occupation-specific, the workers from low-ebb occupations are both disproportionately represented in the unemployment pool and also have a low finding rate. Again, this effect is worse during recessions.

4.1 Technology and preferences

Time is discrete. Production is split into J “occupations,” and workers have skills suited for various occupations. New workers coming from occupation ℓ provide $\omega_{\ell d}$ in destination occupation d . Here $\omega_{\ell\ell} = 1$ and $\omega_{\ell d} < 1$. Labor in an occupation is aggregated linearly in

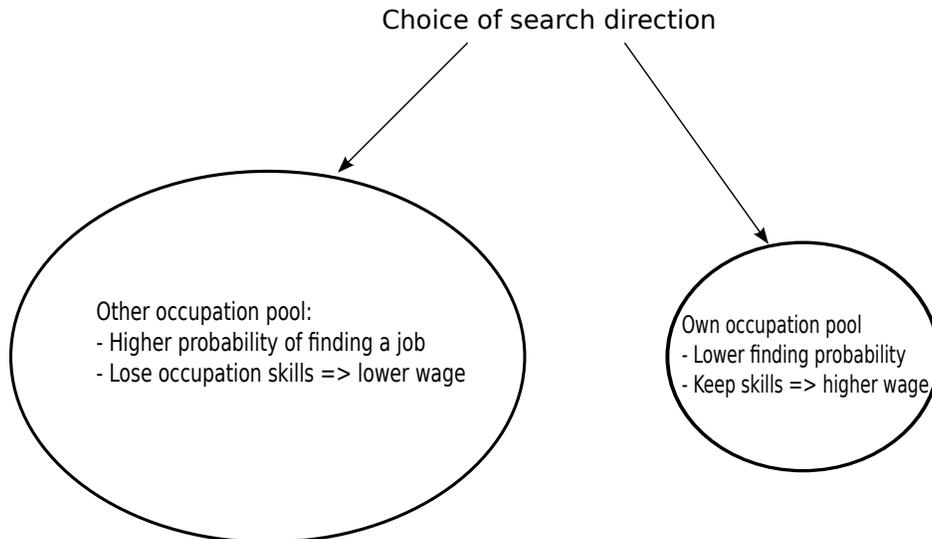


Figure 7: A searcher decides whether to look narrowly or outside of his field

each type. So if $x_{\ell d}$ is the measure of workers working in d who last worked in ℓ then the productive labor force of d is $L'_d = \sum_{\ell=0}^J \omega_{\ell d} x'_{\ell d}$. Note that production is done by the labor force at the end of the period after all transitions have completed, hence it uses L'_d rather than L_d that came into the period.

Buffeting these occupations are idiosyncratic shocks z_d , which are affected by the average level of productivity, Z and a set of factors \mathbf{f}_t that are unobservable but help determine co-movements. Aggregate productivity follows a simple AR(1) process and the vector of factors follows a VAR. Hence, productivity shocks are described the by system

$$Z_t = \rho_Z Z_{t-1} + \epsilon_t \quad (2)$$

$$z_{d,t} = \lambda_{f,d} \mathbf{f}_t + \lambda_{Z,d} Z_t + \rho_z z_{d,t-1} + (1 - \rho_z) \zeta_{d,t} \quad (3)$$

$$\mathbf{f}_t = \Gamma \mathbf{f}_{t-1} + \eta_t \quad (4)$$

For future notation, it will be helpful to define \mathcal{Z} as the state of the joint $Z, \mathbf{f}, \{z_d\}$ process. The dispersion in coefficients $\lambda_{Z,d}$ model differences in cyclical sensitivity while the factors \mathbf{f}_t and loadings $\lambda_{f,d}$ allow for co-movements amongst some occupations additional to the aggregate cycle. The form is meant to be parsimonious because there are 22 occupations and so an estimate for $E[\zeta \zeta']$ is infeasible.

Several additional shocks govern worker transitions. They become experienced at rate τ meaning they provide productivity $\omega_{dd} = 1$. If an experienced worker is separated, he can keep that experience if he matches with the same occupation. Those who are separated while inexperienced become unattached to any occupation, and upon matching with and

occupation d will provide ω_{0d} .

To model endogenous separations, workers who enter the period with a job draw disutility from work $\xi_i \sim H$ which, if it is large enough, will provoke a separation. The separation policy will have a cutoff property, so for each (ℓ, d) type there will exist a $\bar{\xi}$ such that for $\xi < \bar{\xi}$ the match is no longer profitable the separation probability is $s = H(\bar{\xi})$. These shocks are i.i.d., following den Haan et al. (2000) who show that persistence of these shocks is not necessary to match dynamics of separations. By modeling separation-inducing shocks as preferences rather than productivity shocks, I deviate from den Haan et al. (2000) and most of the earlier literature. The problem is that productivity shocks complicate the task of matching observed productivity fluctuations. With productivity shocks, there is “cleansing” over the cycles meaning that the primitive shocks are more volatile than those observed.⁵

When they are searching for a job, there is also another shock $\psi \sim F$ that captures their love of the new job. At the beginning of a period of searching, the agent i sees shocks from each occupation he may choose, $\{\psi_{i,j}\}_{j=1}^J$, but only experiences $\psi_{i,d}$ in his destination occupation if he successfully finds a job.

Agents utility is linear and they enjoy consumption on top of the shocks. To summarize, worker i who stays matched the entire period will experience flow utility $c_i + \xi_i$. If that worker was newly matched in occupation d in that period he experiences $c_i + \psi_{i,d}$.

4.2 Unemployment benefits

Workers who are unemployed receive a flow utility b_i . If they are still eligible for benefits, $e = 1$, this is a 50% replacement of their former salary.⁶ With probability δ benefits expire, $e = 0$, and then they receive food stamp support, as in Nakajima (2012), at about 17% of average earnings. This random expiration modelling technique dates to Fredriksson and Holmlund (2001) and significantly reduces computational load and when I discuss the quantitative results, I will delve further into the implications of this simplification.

Note that unemployed workers do not enjoy leisure utility, which Hagedorn and Manovskii (2008) and others have shown plays an important role in matching the volatility of vacancy postings. This is because those working experience disutility from work and that plays a similar role in reducing the expected surplus from a match. In the baseline version of the model, I will set $\delta = 0$ but discuss as an extension when it takes more realistic values.

⁵den Haan et al. (2000) elide this nuance.

⁶I give 50% replacement of the average earnings of experienced or inexperienced workers. This is only a tiny difference in benefit but saves me from adding an entire state.

4.3 Search and market structure

Each (ℓ, d, e) has its own labor search market $m \in M$, where $M = \{(\ell, d, e) : \ell \in \{0, \dots, J\}, d \in \{1, \dots, J\}, e \in \{0, 1\}\}$. This implies that there is a tightness θ for workers with productivity level $\omega_{\ell, d}$ and with benefits e . The matching function has constant returns to scale and the standard conditions on derivatives. The finding rate for workers is $p(\theta^m)$ and for firms is $q(\theta^m) = p(\theta^m)/\theta^m$, where $q' < 0 < p' < \infty$. κ is the posting cost.

An individual unemployed worker with prior experience ℓ^* and eligibility e^* chooses a vector $\{g^m\}_{m \in M_{\ell^*, e^*}}$ as the probability of applying to destination markets m . M_{ℓ^*, e^*} is the set of markets for which this worker can apply, $M_{\ell^*, e^*} = \{(\ell, d, e) | \ell = \ell^*, d \in \{1, \dots, J\}, e = e^*\}$. Given the state of the worker, denote the tightness he faces in market m as θ^m . Hence, his realized match probability is going to be $\sum_{m \in M_{\ell^*, e^*}} g^m p(\theta^m)$.

Wages are renegotiated each period by generalized Nash bargaining. Hence, they depend upon the output of the match, $z_d \omega_{\ell d}$ and the preference shocks (ξ, ψ) . The firm's bargaining weight is μ . This renegotiation ensures that separations are mutual and that no pareto-improving transfer could occur.

Timing

The timing is such that first all uncertainty is revealed, then matches are made and finally they produce. This means that there may be unemployment stints lasting less than a single period which helps match flows in the data to the discrete time model.

The period is divided into stages as follows:

1. Shocks to productivity $Z, \{z_d\}$
2. Employed workers become experienced
3. Random utilities are realized
4. Separations occur
5. Workers choose their search direction and matches happen
6. Production and consumption occurs
7. Benefits expire

4.4 Households

I now describe the recursive problem of workers in this economy. For all households, the state is x, \mathcal{Z} and their type. The individual's type is defined by ℓ, d, e , the experience, current occupation and benefit eligibility. The value for unemployed workers enjoying benefits is $U(\ell, 0, 1, x, \mathcal{Z})$, those with expired benefits is $U(\ell, 0, 0, x, \mathcal{Z})$ and $U(\ell, d, 1, x, \mathcal{Z})$ for employed workers. For expositional clarity, I will subdivide the problem into three component value functions U_w, U_b for working and unemployed workers, respectively.

The value function, written from the perspective of stage 1, is then

$$U(\ell, d, e, x, \mathcal{Z}) = \mathbb{I}_{d>0} U_w(\ell, d, 1, x, \mathcal{Z}) + \mathbb{I}_{d=0} U_b(\ell, 0, e, x, \mathcal{Z}) \quad (5)$$

Note that, $w^m, s^m, \bar{\xi}^m, g^m, \theta^m$ are all functions of the aggregate state (x, \mathcal{Z}) , but notational convenience I suppress this dependence. The component value-functions are:

$$U_w(\ell, d, 1, x, \mathcal{Z}) = \max_{\bar{\xi}^{\ell d 1}} (1 - \tau \mathbb{I}_{\ell \neq d}) \left[(1 - s^{\ell d 1}) \left(\int_{\bar{\xi}^{\ell d 1}}^0 w^{\ell d 1}(\xi, 0) + \xi dh(\xi) + \beta E U(\ell, d, 1, x', \mathcal{Z}') \right) + s^{\ell d 1} U_b(0, 0, 1, x, \mathcal{Z}) \right] + \tau \mathbb{I}_{\ell \neq d} U_w(d, d, e, x, \mathcal{Z}) \quad (6)$$

$$U_b(\ell, 0, e, x, \mathcal{Z}) = \int_{\psi} \max_{\{g^m\}_{m \in M_{\ell, e}}} \sum_{m \in M_{\ell, e}}^J p(\theta^m) g^m (w(0, \psi_m) + \psi_m + \beta E [U(\ell, d \in m, 1, x', \mathcal{Z}')]) + \left(1 - \sum_m p(\theta^m) g^m \right) (b(\ell, 0, e) + \beta E [U_b(\ell, 0, e', x', \mathcal{Z}')]) df(\{\psi_m\}) \quad (7)$$

Note that if $\ell = d$, then $\tau \mathbb{I}_{\ell \neq d} = 0$ and

$$U_w(d, d, 1, x, \mathcal{Z}) = \max_{\bar{\xi}^{d d 1}} (1 - s^{d d 1}) \left(\int_{\bar{\xi}^{d d 1}}^0 w^{d d 1}(\xi, 0) + \xi dh(\xi) + \beta E U(d, d, 1, x', \mathcal{Z}') \right) + s^{d d 1} U(d, 0, 1, x, \mathcal{Z}) \quad (8)$$

All workers take as given equilibrium conditions on θ^m, w^m , which are described below. Expectations for the law of motion of $\mathcal{X}(x) = x'$ is consistent with the actual law of motion. Note that s^m already includes the integration over ξ , i.e. $s^m = H(\bar{\xi}^m)$

Firms

The representative multi-worker firm produces uses many occupations and posts vacancies in any labor market, $\{v^m\}_{m \in M}$. With aggregation, these vacancies will determine tightness in each market, but the individual firm takes tightness, $\{\theta^m\}$, and wages, $\{w^m\}$ as given.

In addition, the firm takes as given the distribution of ψ that will arrive, induced by the household's maximizing direction choice. Call \tilde{F} this extreme value distribution. Because production is linear in workers there is no externality problem à la Stole and Zwiebel (1996). However, using a single firm employing multiple types of workers simplifies the problem of what happens when a worker becomes experienced.

$$\begin{aligned} \Pi(\mathbf{L}, x, \mathcal{Z}) = & \max_{\{v^m\}_{m \in M}} \sum_{\ell=0, d=1}^J \left[\omega_{\ell d} z_d L'_{\ell d} - \left(L'_{\ell d} - \sum_{e \in \{0,1\}} v^{\ell d e} q(\theta^{\ell d e}) \right) \int_{\bar{\xi}^{\ell d}}^0 w^{\ell d 1}(\xi, 0) dh(\xi) \right] \\ & - \sum_m \int_{\psi} v^m q(\theta^m) w^m(0, \psi) d\tilde{f}(\psi) + \kappa v^m + \beta E[\Pi(\mathbf{L}', x', \mathcal{Z}')] \end{aligned} \quad (9)$$

Where the law of motion for the labor force \mathbf{L}' is given by:

$$L'_{\ell d} = (1 - \tau)(1 - \hat{s}^{\ell d 1})L_{\ell d} + q(\theta^{\ell d 0})v^{\ell d 0} + q(\theta^{\ell d 1})v^{\ell d 1} \quad (10)$$

$$L'_{dd} = (1 - \hat{s}^{dd 1})L_{dd} + q(\theta^{dd 1})v^{dd 1} + q(\theta^{dd 0})v^{dd 0} \quad (11)$$

Here, the \hat{s} denotes the firms' expectations the aggregate separation policies. Of course, in equilibrium all these expectations will be consistent with aggregate behavior.

The wage bill depends on cut-offs, $\{\bar{\xi}^m\}$ and the distributions of $\{\psi_m\}$ induced by the searchers' problems. However, because these shocks are not persistent it is enough for the firm to know the aggregate state when for their forecast value in the next period.

4.5 Equilibrium

To define equilibrium, I give conditions for tightness and wages, which play the role of prices, and the laws of motion for workers of various types. A recursive competitive equilibrium is

- A set of functionals

$$- U : \{0, J\}^2 \times \{0, 1\} \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow \mathbb{R}_+$$

$$- \Pi : [0, 1]^{J(J+1)} \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow \mathbb{R}_+,$$

- Policy functions:

$$- \{g^m\}_{m=0}^{2J(J+2)} \text{ where } g^m : \mathbb{R}_+^J \times \{0, J\}^2 \times \{0, 1\} \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow [0, 1]$$

$$- \{\bar{\xi}^m\}_{m=0}^{J(J+2)} \text{ where } \bar{\xi}^m : \{0, J\}^2 \times \{0, 1\} \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow \mathbb{R}_-$$

$$- \{v^m\}_{m=0}^{2J(J+2)} \text{ where } v^m : [0, 1]^{J(J+1)} \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow \mathbb{R}_+$$

- Wages, $\{w^m\}_{m=0}^{2J(J+2)}$, where $w^m : \mathbb{R}_- \times \mathbb{R}_+ \times [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow [0, 1]$

- Market tightness, $\{\theta^m\}_{m=0}^{2J(J+2)}$, where $\theta^m : [0, 1]^{2(J+1)^2} \times \mathbb{R}^4 \rightarrow [0, 1]$

Though not primary equilibrium objects, it will ease notation to first define the aggregate number of applicants, a^m , in each market $m = \ell, d, e$

$$a^m = \begin{cases} \bar{g}^m \left(x_{\ell 01} + \bar{s}^{\ell \ell 1} (x_{\ell \ell 1} + \tau \sum_{j \neq \ell} x_{j \ell 1}) \right) & \ell > 0, e = 1 \\ \bar{g}^m \left(x_{001} + (1 - \tau) \sum_{j=0}^J \sum_{k \neq j} \bar{s}^{jk1} x_{jk1} \right) & \ell = 0, e = 1 \\ \bar{g}^m (x_{\ell d 0}) & e = 0 \end{cases}$$

Here \bar{g}^m , and \bar{s}^m are both evaluated at the average (defined below). Distinguishing these from the agent-level counterparts matters a great deal for \bar{g}^m because while the linearity implies that in equilibrium $g^m \in \{0, 1\}$ but averaging over, $\bar{g}^m \in [0, 1]$

- Tightness in market $m = (\ell, d, e)$ satisfies $\theta^m = \frac{v^m}{a^m}$
- Expectations $\mathcal{X}(x)$ are consistent with aggregate laws of motion, where policies $\bar{g}^{\ell d}$ are evaluated at the aggregate, $\bar{g}^m(\cdot) = \int_{\psi_i} g^m(\psi_i; x, \mathcal{Z}) dF(\psi)$

$$x'_{dd1} = (1 - \bar{s}^{dd1}) \left(x_{dd1} + \tau \sum_{\ell=0, \ell \neq d}^J x_{\ell d1} \right) + p(\theta^{dd1}) a^{add1} + p(\theta^{dd0}) a^{add0} \quad (12)$$

$$x'_{\ell d1} = (1 - \bar{s}^{dd1}) (1 - \tau) x_{\ell d1} + p(\theta^{\ell d1}) a^{\ell d1} + p(\theta^{\ell d0}) a^{\ell d0} \quad (13)$$

$$x'_{\ell 01} = \left(1 - \sum_d \bar{g}^{\ell d1} p(\theta^{\ell d1}) \right) \left((1 - \delta) x_{\ell 01} + \bar{s}^{\ell \ell 1} \left(x_{\ell \ell 1} + \tau \sum_{j \neq \ell} x_{j \ell 1} \right) \right) \quad (14)$$

$$x'_{001} = \left(1 - \sum_d \bar{g}^{0d1} p(\theta^{0d1}) \right) \left((1 - \delta) x_{001} + (1 - \tau) \sum_{d=1}^J \sum_{l=0, l \neq d}^J \bar{s}^{\ell d1} x_{\ell d1} \right) \quad (15)$$

$$x'_{\ell 00} = \left(1 - \sum_d \bar{g}^{\ell d0} p(\theta^{\ell d0}) \right) x_{\ell 00} + \delta \left(1 - \sum_d \bar{g}^{\ell d1} p(\theta^{\ell d1}) \right) x_{\ell 01} \quad (16)$$

- The firm is representative, so $L_{\ell d} = x_{\ell d1}$

A few endogenous variables are pinned down in equilibrium:

- There is free entry, so tightness satisfies:

$$\kappa = q(\theta^{\ell d e}) \left(\omega_{\ell d} z_d - \int_{\psi} w^{\ell d e}(0, \psi) d\tilde{f}(\psi) + \beta E \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}') \right) \quad (17)$$

Where $\Pi_{\ell d}$ is the derivative with respect to $L_{\ell d}$

- Separations are mutual and the cutoff satisfies

$$\omega_{\ell d} z_d + \bar{\xi}^{\ell d 1} + \beta E[U_w(\ell, d, 1, x', \mathcal{Z}') + \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}')] = \begin{cases} U_w(\ell, 0, 1, x, \mathcal{Z}) & \ell = d \\ U_b(0, 0, 1, x, \mathcal{Z}) & \ell \neq d \end{cases} \quad (18)$$

- Wages $w^{\ell de}(\xi, \psi)$ are set by bargaining, where with firms' weight is μ .

$$w^{\ell de}(\xi, \psi) = (1 - \mu) (\omega_{\ell d} z_d + \beta E \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}')) - \mu (\xi + \psi - b + \beta E U_w(\ell, d, 1, x', \mathcal{Z}') - \beta E U_b(\ell, 0, e, x', \mathcal{Z}')) \quad (19)$$

Note that wages depend on the entire distribution of shocks, because the workers' outside option allows him to move around. This has the effect of compressing wages across occupation types for inexperienced workers who will readily switch and also raising them compared to their more experienced counterparts except in the most productive occupation in the economy.

4.6 Discussion of the model assumptions and equilibrium

4.6.1 Work history truncation

I truncate the work history after two periods, so that those who lose a job when inexperienced become “unattached” rather than storing two periods of their work history. This assumption imbues the model with several, mostly innocuous peculiarities, especially with regard to wage dynamics. Wages in this model fluctuate non-monotonically over one's tenure. They are higher in the first period of employment than the second, and then rise again upon gaining tenure. This is because the outside option in the worker's first period is to stay unemployed but with skills from to his own occupation, i.e. $b(\ell, 0, 1, \cdot) + \beta E[U_b(\ell, 0, 1, x', \mathcal{Z}')]$. In subsequent periods, if he matched in a new occupation, the outside option becomes $b(\ell, d, 1, \cdot) + \beta E[U(0, 0, 1, x', \mathcal{Z}')]$. He is penalized both by a lower unemployment benefit and a lower continuation value that reflects lost human capital. To avert this problem, I would have define the problem of workers who are displaced after having labor history ℓ, d , and extend the state space an additional element. Instead, I assume that human capital from occupation ℓ is lost upon taking a job type d . This lumps together work histories $(\ell, d, 0) \forall \ell, d : \ell \neq d$ into a category $(0, 0)$.

Clearly, histories need to be truncated somewhere to keep the problem tractable. I choose this setup because it gets the problem of unemployed workers correct even if it misses the problem of currently employed workers. It simplifies the process by which old, unused skills

depreciate by just assuming that it is immediate.

4.6.2 Efficiency

Like all models with a search friction, to establish efficiency in this model, I must confront the congestion externality. It will turn out that fairly standard conditions are enough, but this is not a foregone conclusions. Efficiency in my model is complicated by the ability for agents to choose their search market. In addition to the vacancy posting condition, I must establish that another policy, search direction, is inline with a planner's choice.

Proposition 1 *The competitive equilibrium is an efficient allocation if the bargaining weight of the worker μ is equal to the elasticity of the worker's finding rate, $\frac{\partial p(\theta^m)}{\partial \theta^m} \frac{\theta^m}{p(\theta^m)} \forall m \in M$ and the match surplus is split proportionally between worker and firm.*

This proposition goes one step beyond the standard Hosios condition because even if job posting can be set at the efficient level, the search direction choice, $\{g^m\}$ requires another condition. Note that the elasticity of the matching function has to be constant, as with a Cobb-Douglas matching function, because θ will generally vary across $m \in M$. As we know, in a model with search frictions and generalized Nash bargaining, the congestion externality affects a competitive equilibrium. The Hosios condition sets the bargaining weight such that the job posting decision is socially optimal and in a simple Mortensen-Pissarides model, this is enough. In this model, however, workers also choose a search direction. The socially optimal search direction considers the entire surplus that would be gained whereas search direction responds only the workers' share in the competitive equilibrium. My result is reminiscent of Marimon and Zilibotti (1999), who also show efficiency in a model with heterogeneous searchers choosing where to allocate themselves. My logic is somewhat different from theirs, however. Their searchers had to choose whether to stay and bargain or continue searching, and the Hosios affected them just like firms' very similar posting decision. On the other hand, my searchers weigh different directions within the period, so their bargaining weight should be neutral with respect to the relative returns across occupation destinations. With any distributional assumptions, the variance of the preference shock $var(\psi)$ is scaled by the workers' share and so is the return to any direction. So long as all of the returns are scaled evenly, the competitive equilibrium corresponds to the social planner's solution. This last condition is provided by Nash bargaining, in which agents in the labor market equally split the total match surplus.

Tightness is the same that would come out of a planner problem so long as the correct bargaining weight is applied and wages are negotiated individually, as in most similar models. I show this in Appendices B and C. The basic intuition is the same as in most search and

matching models following Hosios (1990). As is explicated in greater detail, the solution to the job creation decision in the planner's problem for each submarket is determined by:

$$\kappa = p'(\theta^{\ell de}) \left(\omega_{\ell d} z_d - b(\ell, d, e) + \int \psi \tilde{f}(\psi) d\psi + \beta E[W'_{\ell d1} - W'_{\ell 0e}] \right) \quad (20)$$

Where $W_{\ell d1}, W_{\ell 0e}$ are the partial derivatives of the planner's objective with respect to $x_{\ell d1}, x_{\ell 0e}$. Using our assumptions on the matching function, that $\theta q(\theta) = p(\theta)$ we can substitute $\frac{\partial p}{\partial \theta} = \tilde{\mu} \frac{p(\theta)}{\theta} = \tilde{\mu} q(\theta)$ where $\tilde{\mu}$ is the elasticity of $p(\theta)$ with respect to θ . It is clear then, that setting $\tilde{\mu} = \mu$, we get the competitive markets condition in Equation 17 so long as the quantity inside the parentheses is equal to the match surplus and that is the same in both competitive and centralized versions.

As in any search and matching model, we also need separations to be efficient. This is fairly straightforward given our assumption that wages adjust with the shocks ξ . The separations occur such that at the cutoff $\bar{\xi}^{\ell d1}$ the surplus value of the match is 0 after all the shocks of the period have realized. Just as in the competitive case, at the cutoff, the value of producing is the same as the outside value of a searcher:

$$\omega_{\ell d} z_d + \bar{\xi}^{\ell d1} + \beta E W'_{\ell d1} = W_{\ell 01}$$

Next, we have to check that efficiency does not fail in the g^m policies. In the competitive equilibrium, the searcher only realizes a fraction of the surplus upon choosing a direction, whereas the social planner realizes its entirety. The scale is different in the social planner and competitive equilibrium, but not the relative returns to different directions. So long as the shocks are also equally scaled, this is not a problem.

This is easiest to see if we take a simplification of the model. Suppose shocks were Gumbel-distributed with standard deviation σ_ψ and the matching probability was the same everywhere, p . Call R_d the total return for a searcher if he matches in occupation d . In this case, the reservation value drops from the likelihood. The social planner, considering the entire return chooses the probability of applying to occupation j as

$$\bar{g}^j = \frac{\exp \{p^j R_j / \sigma_\psi\}}{\sum_d \exp \{p^d R_d / \sigma_\psi\}}$$

And the probability a searcher goes to occupation j in competitive equilibrium is given by

$$\bar{g}^j = \frac{\exp \left\{ \frac{p^j (1-\mu) R_j}{(1-\mu) \sigma_\psi} \right\}}{\sum_d \exp \left\{ \frac{p^d (1-\mu) R_d}{(1-\mu) \sigma_\psi} \right\}}$$

And here $1 - \mu$ scaling both return and shock cancel. This need not be the case if the outside option, which is not scaled by $1 - \mu$, did not cancel.

4.6.3 The role of preference shocks for efficiency

The preference shocks for new arrivals serve to add convexify the equilibrium choice set. In addition, they give a degree of freedom to the equilibrium choices and without them, we could not generally have efficiency. To understand this, first note there are two jump variables determining θ instead of the usual one. In every search and matching model, vacancies fluctuate to satisfy the free-entry condition and thus set the matching probability. But here, \bar{g}^m can also adjust immediately. Its optimality condition equates expected returns across occupation choices. With preference shocks, this is tantamount to choosing a cutoff above which the searcher chooses this direction. Without these shocks however, returns must all equate and also satisfy the free entry condition.

Proposition 2 *In an equivalent economy but in which $\sigma_\psi = 0$, firm entry in the competitive equilibrium is generically not efficient.*

Consider a world in which there is no random utility ψ . In this case, for a given ℓ, e the planner chooses a degenerate distribution for $\{g^m, \theta^m\}_{m \in M_{\ell e}}$ while generically, searchers in a competitive market will choose $g^m \in (0, 1)$ and $\theta^m > 0$ in more than one market.

Detail for this proof is given in a yet to be completed appendix. To understand the logic, consider that without the shock we would have a system of conditions to determine search direction

$$\begin{aligned} g^m &: p(\theta^m) (w^m - b + \beta E[U(\ell \in m, d \in m, \cdot) - U(\ell, 0, \cdot)]) \\ &= p(\theta^n) (w^n - b + \beta E[U(\ell, j \in n, \cdot) - U(\ell, 0, \cdot)]) \quad \forall m, n \in M_{\ell, e} \text{ and } \forall \ell, e \end{aligned} \quad (21)$$

and a single equation to determine entry

$$\kappa = q(\theta^{\ell de}) (\omega_{\ell d} - \bar{w}^{\ell de} + \beta E\Pi_{\ell d}) \quad (22)$$

In a competitive economy, this implies values for both θ^m, w^m , much like competitive search in the tradition of Moen (1997). In general g^m is not degenerate and workers apply across a variety of locations. But, for the planner consider any $m = (\ell, d, e), n = (\ell, j, e)$. The conditions for any search direction that obtains in equilibrium are

$$g^m : p(\theta^{\ell d}) (\omega_{\ell d} - b + \beta E[W'_m - W'_{\ell 0e}]) = p(\theta^n) (\omega_{\ell j} - b + \beta E[W'_n - W'_{\ell j 0}]) \quad (23)$$

Here W_m is the derivative of the planner's objective with respect to a worker of type $m = (\ell, d, e)$. But, free entry must also satisfy

$$\kappa = q(\theta^m) (\omega_{\ell d} - b + \beta E[W_m - W_{\ell 0 e}]) \quad (24)$$

To satisfy Equation 24 for every ℓd pair, the model cannot generally also satisfy Equation 23 except in knife-edge parameterizations. To see this, combining these two conditions to yield

$$\kappa \frac{p(\theta^m)}{q(\theta^m)} = \theta^m = \kappa \frac{p(\theta^n)}{q(\theta^n)} = \theta^n \quad (25)$$

But if tightness everywhere is the same, the cumulative value of every match must also be the same. Given the linear structure of production, i.e. no decreasing returns, this cannot generally be satisfied. Substituting back into 24 and simplifying, this implies the gross return in both occupations must be equal:

$$\omega_{\ell d} + \beta E[W_m] \stackrel{?}{=} \omega_{\ell j} + \beta E[W_n] \quad (26)$$

For the planner's problem, search direction is almost always degenerate for all workers because Equation 26 cannot be satisfied, whereas in a competitive economy with bargained or *posted* wages this is not generally the case. That wage posting and directed search fails to deliver efficiency is somewhat interesting.

4.7 Cyclical dynamics of finding rate heterogeneity

A crucial result of the model is that the dispersion of unemployment duration is counter-cyclical. To deliver this, there must be a mechanism such that occupation switching takes longer in recession. Remember Table 1, showing that switchers are unemployed for longer and the effect is larger in the Great Recession. This empirical effect is part of how the model delivers counter-cyclical dispersion in unemployment duration across occupations. In a recession, the occupations that are most cyclically sensitive have the worst productivity and the workers attached to this occupation are unemployed for a very long time whether they find a job in their own or other occupations.

There are two channels that slow the finding rate for recession-affected occupations and slow reallocation away from them. The first is the same as a standard search and matching model, the surplus size falls and vacancy postings decline. The second, is that occupation switching becomes more difficult. The reasoning for this is subtle: the total size of the surplus for experienced workers is larger and hence less elastic with respect to a productivity

shock. This logic is the same as that of Hagedorn and Manovskii (2008), which shows that vacancies are much more volatile in a market in which the total match surplus is small. The same size shock to the flow value of a match has a large effect on the relative size of the surplus and hence a large effect on vacancy posting.

In the context of my model, this effect means a higher volatility of vacancies in markets for switchers, in which the value of the match $\omega_{\ell d} z_d - \xi + \beta(E[U_w(\cdot) + \Pi_{dd}(\cdot)])$ is smaller and hence closer to the flow value of unemployment $b + \beta E[U_b(\cdot)]$. In recessions, few postings imply that it is more difficult to switch, burnishing the unemployment rate of occupations that are hit hardest by the recession.

Related, the experience premium within an occupation is pro-cyclical and this means inexperienced workers are relatively expensive during recessions. To put this differently, $\frac{\omega_{dd}}{\omega_{\ell d}} > \frac{w_{dd}^g}{w_{\ell d}^g} > \frac{w_{dd}^b}{w_{\ell d}^b}$, where g stands for *good times* and b for *bad*. The crucial insight is that the workers' outside option is relatively constant, whereas the match value depends more strongly on the shock level.

5 Quantitative Strategy

Crucially, workers in the model weigh their attachment to their prior occupation against the forces pushing them away. I will try to match these to the data as precisely and with as much detail as possible. The former will come from the wage gap between inexperienced and experienced workers in each occupation and the latter will come from the occupation-specific shock processes.

5.1 Occupation-specific shocks

The shock process described by Equations 2-4 can be directly estimated if we have information on value-added by occupation, workers per occupation *and their working history*. This final aspect will prove to be the trickiest. The trouble arises from the way we model $\{\omega_{\ell d}\}$ as a vector of relative productivities. Hence, when workers shift into a new occupation, they affect the observed productivity in that occupation but not the actual shock that hit it. Hence, the primitive shock process is not directly observable, but only a process that combines these exogenous driving forces and the endogenous movement of workers.

The Bureau of Economic Analysis publishes annual industry accounts that provide the industry-specific value-added and number of workers in terms of bodies and full-time equivalents. To map from industries to occupations, I pool CPS surveys over the year to compute the portion of workers attached to each occupation working in each industry. For example,

in 2000 88% of labor in SOC 31, Health Care Support Occupations, are in the industry code Health Care and Social Assistance and the rest are spread over other industries. The same year, 22% of those working in Business and Financial Operations Occupations are in the Finance and Insurance industry and there are also high concentrations elsewhere.

I then attribute value added to the occupation corresponding to each industry in which they work, weighted by the fraction of the workers there. To get value added per worker, I then divide by the number of workers in that occupation. Because the CPS is only a snapshot and contains sampling error, I use a 4 year moving average to smooth it. Also, because my sample of the CPS begins in 1962 but the BEA data begins in 1948, I use the average during the 1960s for all of the years prior. This is going to bias downwards my estimate of the variance of occupation-specific productivity.

5.1.1 Adjusting for endogenous worker experience

The problem now is that I treated every worker as equal, whereas in the theory workers of different experience levels will have different average productivity levels in that occupation. The model counterpart to observed output per capita is endogenous and so I can not match productivity shocks in the model to observed output per capita in the data. I cannot adjust the data observation for observed experience either because the data is insufficient. In the CPS, I only observe their job experience, but not occupational experience. Even in the SIPP, I observe occupational experience but only their prior occupation if they transitioned during the survey period, which is relatively infrequent given the shortness of the SIPP panel. Further, even if I knew their prior occupation, to be consistent with the model, I would also need to know the tenure in that occupation, to find out if they are providing $\omega_0 d$.

The model, however, has predictions for the number of workers that will be inexperienced and from whence they came. So, I can make the model analog of observed productivity, given this endogenous composition consistent with the actual observations. I will choose a stochastic process for \mathcal{Z} such that the output per worker in an occupation is the same in the model and data.

To be specific, there are two major differences between the occupation-level productivity shocks in the model and that which is observed in data:⁷ (1) I only observe industry-level productivity and (2) observed productivity in my model is endogenous because the composition of workers determines the output per head.

Notice, that I could relate observed industry shocks, $z_{i,t}$ to occupation shocks $z_{d,t}$ if I

⁷A minor difference is that I can only make observations every year, though the model is monthly. I simply estimate the process as if I only observe once every 12 observations.

assume that distributions of types are the same across industries.

$$z_{i,t} = \sum_d w_{i,d} \sum_\ell \omega^{\ell d} x'_{\ell d} z_{d,t} \quad : \quad w_{i,d} = \frac{\Pr[occ \cap ind]}{\Pr[ind]} \quad (27)$$

I can get $w_{i,d}$ from the CPS. However, the object $\omega^{\ell d} x'_{\ell d}$ is endogenous. $x'_{\ell d}$ is the density of workers with productivity $\omega^{\ell d}$ —endogenous and unobservable.

If, however, I pretend $\omega^{\ell d} = 1 \forall \ell, d$, then the endogenous experience distribution, \mathbf{x} does not matter. I go to the data with this false assumption and estimate a process for $\{\{z_d\}, Z, \mathbf{f}\}, \tilde{\mathfrak{Z}}^{data}$. I parameterize this by the same factor process as described in Equations 2-4. Hence, the auxiliary model is the dynamic process for productivity, given that $\omega^{\ell d} = 1 \forall \ell, d$. I estimate the auxiliary model in the data taking model-generated data with this same assumption, that $\omega^{\ell d} = 1 \forall \ell, d$, to find $\tilde{\mathfrak{Z}}^{model}$. Then, I choose model parameters for the process in Equations 2-4 to solve

$$\min LR(\tilde{\mathfrak{Z}}^{data}, \tilde{\mathfrak{Z}}^{model})$$

An alternative way to view this strategy is that I measure output per capita by occupation in both the model and data. To map from industry to occupation, I assign industry output to occupation on the basis of the relative population. Then in the data I have a measure of output per person conditional on occupation and I can compute this same object in the data. The auxiliary model can then be described as the stochastic process for average output per capita within an occupation.

The model does not have a notion of industry, so for equation 27 to work, I need to assume that workers from occupations are equally productive in each industry. Otherwise, rather than attributing the output proportional to the workers there, there would need to be some weighting. If I were explicit on industries with specific shocks, one would not expect the productivity distribution to be the same. Just as workers flow in response to occupation shocks, they would also move for industry shocks.

The distance metric I choose is the likelihood ratio rather than the Wald metric for two reasons. First, there are many parameters in the set of $\lambda_{f,d}, \lambda_{Z,d}$ and weighting them is somewhat fraught, whereas the LR metric gives a natural weighting for observations. Also, the magnitude of the unobservable factors and coefficients is indeterminate; one could multiply each $\lambda_{f,d}$ by 2 and divide the factors by 2. One can, of course, use a normalization to constrain this degree of freedom, but then that normalization would implicitly affect the weights. By construction, my shock sequence reproduces the data exactly if I were to recompute industry-level shocks using Equation 27.

5.1.2 Specification of the stochastic process

The factor process is explicitly designed to introduce time-varying volatility between occupations. As mentioned, this is a feature explored in Lilien (1982) and Abraham and Katz (1986). More contemporary papers have introduced this into search and matching, notably Schaal (2012). Fernandez-Villaverde and Rubio-Ramrez (2010) discuss several ways to model time-varying volatility. They point out that a finite number of regimes complicates using perturbation techniques to solve the model, as I do here. Instead, they focus on stochastic volatility or ARCH models. There are a number of shortcomings to modelling occupation specific shocks as they suggest:

$$\begin{aligned}z_{d,t} &= \rho_z z_{d,t-1} + \sigma_t \zeta_{d,t} \\ \log \sigma_t &= \rho_\sigma \log \sigma_{t-1} + \varphi_t\end{aligned}$$

As one might imagine, this process is difficult to estimate precisely on the limited data sample I have. Several methods, indirect inference, GMM and quasi-ML all run into the same basic problem, that estimating a dynamic process for a second moment is essentially using 4th moment information. With only 60 observations on variance from 1947-2007, we end up with quite a wide bound at two standard deviations $\rho_\sigma \in (0.118, 0.613)$. Furthermore, the process requires some correlation of the shock structure to work in some desirable quantitative aspects. For example, it is probably important to have some correlation between φ_t and ϵ from Equation 2. Without an estimate for $E[\zeta_{d,t}, \zeta'_{d,t}]$, we cannot capture the co-movements among occupations, which is an important feature because occupations that tend to experience similar shocks are also similar in skill space. This plays a large role in how “far” workers switch and how much wage they lose.

5.2 Cost of switching occupations

The relative productivity of inexperienced workers, $\{\omega_{\ell d}\}_{\ell=0, d=1}^J$ needs to be estimated to match the costs of switching occupations. I use indirect inference to estimate ω as a function of the O*NET skills of the occupation pair. To summarize the method, I quantify occupations’ relation to each other by their skill requirements, as published by the US DOL’s O*NET. Then, I assume $\omega_{\ell d}$ is function of the difference between ℓ and d . This motivates a regression on the relative wage between (ℓ, d) workers and (d, d) workers and the difference in O*NET skills in the two occupations. I then match the coefficients of this regression between model-generated data and observations in the SIPP from 1996-2007.

Were it not for the random utility to choosing occupations changes to $\{\omega_{\ell d}\}$ would cause

discontinuous changes in policies and observable endogenous variables. If in equilibrium workers are making a transition from ℓ to d , a change to $\omega_{\ell d}$ by even just an infinitesimal amount will cause the direction of search to change discretely. If these new parameters cause workers to make an ℓ to j switch, now the observed return to skill will change discontinuously because workers are gaining experience in a different occupation. This is precisely the situation that Smith and Keane (2004) address. Their solution is to modify the discrete choice into a probability, which takes on a similar form as the policy $g^{\ell d}$ that convexifies choices in my model.

It should also be noted, that I have assumed occupation stayers experience no wage loss. This restriction can be justified from a modelling perspective, as it focuses the analysis on the additional wage loss for switchers. More importantly, several studies have shown that workers separated but who stay in their occupation have almost no wage loss. Kambourov and Manovskii (2009) make this point using data from the CPS Displaced Worker Survey and Fujita (2011) corroborates the finding using SIPP data. While Fujita (2011) finds some moderate wage loss for occupation stayers, it is statistically insignificant.

5.2.1 O*NET data

The US Department of Labor’s O*NET database collects data on occupations that can be used to quantify the differences between them. It is the successor to the Dictionary of Occupational Titles, which classified the types of tasks necessary to work in a particular occupation. The O*NET expands upon this, providing quantitative information on skills, knowledge, and abilities required to work in a particular occupation. For every occupation each descriptive element gets a score for its “importance” in that occupation and the “level” at which it is performed. It collects survey information from both workers and their supervisors and but publishes only the rescaled means.

To characterize occupations, I use the “skills” measure, because in my model, that which attaches workers to their occupation is learned on the job. O*NET’s skills are most close to this sort of human capital learned on the job. To process this data, I first combine importance and level scores by a Cobb Douglas with elasticity 0.5, as in Blinder (2009). Next, I reduce the 35 skills down to three principal components. This is because I will eventually have to solve for coefficients on each skill dimension and also because they are largely redundant. The first three components explain about 80% of the variation between all of the O*NET occupations. These occupations, however, are finer than the two-digit SOC code-defined occupations I will eventually work with so I take a simple average over O*NET occupations within an SOC occupation. Finally, I rescale the measures by replacing the component, with its quantile rank value amongst the occupations. This leaves each occupation with a skill

value between zero and one in three categories.

5.2.2 The auxiliary model relating occupational skills to productivity

To map O*NET data into the model, I assume that the productivity gap between experienced and inexperienced workers is proportional in logs to the difference in skill intensity. More precisely, suppose

$$\omega_{\ell d} = e^{\sum_{i=1}^3 \beta_i (k_{i,d} - k_{i,\ell})}$$

One would expect $\{\beta_i < 0\}$, meaning that inexperienced workers will see a larger wage gap if their old job is less intensive in a certain skill than the current one.

Now consider a linear approximation of log wages around the average wage of the experienced worker, \bar{w}^{dd} . Using the bargained wages in Equation 19, this yields a convenient linear equation relating $\{\beta\}$ to the wage gap between experienced and inexperienced workers:

$$\log \left(\frac{w^{\ell d}}{\bar{w}^{dd}} \right) \propto \sum_i \beta_i (k_{i,d} - k_{i,\ell}) + const - \mu \psi_d \quad (28)$$

Again, expect $\beta_i < 0$ as a larger deficiency in skills from the prior occupations means a larger value for $(k_{i,d} - k_{i,\ell})$ but also a lower $\log w^{\ell d}$ relative to the average wage in the occupation.

To use this as my auxiliary model, I can just regress the experience premium on the skill difference in both SIPP and model-generated data. Before using wages in the data, I first regress them on sex, age, age squared and college education. I use residual wages because all of these complicating factors in wage determination do not exist in my model, but are quite significant in the data. See 12 at the end for results of the first stage regression.

Our theory says that in the model, the constant includes information on market tightness and the outside option of workers with experience from occupation ℓ . This implies that the auxiliary model is doubly misspecified. By assumption, the errors are not normal and also there should be ℓ fixed effects.

For the average wage for experienced workers I need to choose who to count as “experienced.” Consistent with the model, I would take workers whose occupational tenure is greater than $\frac{1}{\tau}$. Unfortunately, this data operation now depends on calibrated model values and the calibration depends on the results I find this regression. I iterate, first estimating the auxiliary model assuming everyone with any tenure is experienced, then calibrating for τ and then using this for the expected duration before becoming experienced in the data. As it turns out, the regressions are almost unaffected by the selection criteria for the experienced workers in constructing \bar{w}^{dd} .

As seen in Table 2, the auxiliary model is quite precisely estimated and skill differences are

strong predictors of differences in the wage premium for experience. To interpret negative the coefficients in all of the composite skills, a beginner suffers a larger wage gap in an occupation using a skill much more intensively than his old occupation. Put another way, of those starting in an occupation, the ones paid least are coming from occupations that used skills less than the new occupation. As is often the case O*NET elements are quite good predictors for wages. We see that in the auxiliary model' estimation.

	Coefficient	T-statistic	Model	Structural
Skill 1	-0.331	-6.88	-0.35	-0.95
Skill 2	-0.294	-6.20	-0.32	-0.87
Skill 3	-0.186	-4.33	-0.19	-0.52
const	-0.289	-23.98		

Table 2: Auxiliary model for the value of occupational skills on productivity

The indirect inference procedure is fairly successful at minimizing the residual between data and model coefficients. The binding function maps larger structural coefficients for skill loss into these regression coefficients, which is again in line with our expectations. Occupational selection means that the wage residuals are truncated from below and so the coefficients are attenuated towards zero.

5.3 Calibration of the economy

I will calibrate the rest of the parameters in the model. I will use pre-2008 data on both matching and occupational switching behavior. Before listing all of the variables to calibrate, I need to describe functional forms for the matching function and the distribution of disutility shocks.

5.3.1 Heteroskedasticity in the random utility term

In this paper, I model random deviations in search behavior through exponentially-distributed preference shocks. Because these shocks are scaled by the probability of finding a job in that search direction, which differs across markets, the effective variance is heteroskedastic. My distributional assumption accommodates this while preserving a closed form for the probability of taking a given search direction. Alternatively, if I were to model perturbations from the optimal search direction by individual shocks to the finding rate, I would avoid this issue. These options have slight differences in math and interpretation but, subject to distributional assumptions can be mostly isomorphic.

In my choice I follow a long tradition of partially random match quality back to Boskin (1974) and Miller (1984). In my implementation, the basic interpretation is that landing a

job in some occupations is more pleasant than others. For this framework, we can write the return to searching in direction d

$$p(\theta^{\ell d}) (w^{\ell d}(\psi_d) - b + \beta E[U'_{\ell d} - U_{\ell 0}] + \psi_d) \quad (29)$$

We can separate this into a deterministic and stochastic component because the Nash bargaining assumption makes $w(\psi)$ linear in ψ , as in Equation 19. I rewrite Equation 29 as

$$R_{\ell d} + p(\theta^{\ell d})(1 - \mu)\psi_d$$

This is *almost* like the standard additive random utility model except that the random component is scaled by $p(\theta^{\ell d})$. This means that if the variance of ψ is σ_ψ^2 across all choices d , the effective variance of the shock to return is actually $(p^{\ell d}(1 - \mu))^2 \sigma_\psi^2$. Whereas with additive random utility, Boskin (1974) and others use the tractability of Gumbel-distributed shocks, we are now breaking the i.i.d. assumption. If we were to keep Gumbel-distributed errors with heteroskedastic shocks, the direction policy would be chosen according to a heteroskedastic logit, as described in Bhat (1995). While the distribution has a nice interpretation as the first order statistic from a sample of exponentially-distributed utility, it implies policies for $\{g^{\ell d}\}$ that have no closed form description. In particular, the probability of an individual choosing occupation d would be

$$\int_{-\infty}^{\infty} \prod_j G \left(\frac{R_d - R_j + ((1 - \mu)p^{\ell d} \sigma_\psi) \psi_d}{(1 - \mu)p^{\ell j} \sigma_\psi} \right) g(\psi) d\psi_d$$

where G and g are the CDF and PDF of the Gumbel distribution, $G(x) = e^{-e^{-x}}$. Though this is only one integral, rather than 22 without the independence assumption, it is still computationally intensive to evaluate and would considerably increase the computational burden to solve and simulate the model.

With an exponential distribution for ψ , the probability $g^{\ell d}$ has a closed form despite the heteroskedasticity. Another likely candidate, the Fréchet distribution is still convenient with heteroskedasticity (as seen in Eaton and Kortum (2002), but only when the average return is zero. To my knowledge, exponentially-distributed shocks have not been used in economics models of discrete choice since Daganzo (1979). In Appendix A I show that the probability a worker chooses occupation d takes a closed form:

$$\prod_{j=1}^J \frac{t^{\ell j}}{t^{\ell d} + t^{\ell j}} e^{t^{\ell d}(R_{\ell d} - R_{\ell j})} \quad (30)$$

where $t^{\ell j}$ is the precision parameter, the inverse of the standard deviation, of shock $\psi^{\ell j}$.

5.3.2 The matching function

The choice of matching function is actually non-trivial in this model. The most common form is Cobb-Douglas, $m(u, v) = \phi_0 u^{\phi_1} v^{1-\phi_1}$, however this has the problem that in many cases, v rises high enough that matching probability may be greater than one. As is often the case in directed search models with many submarkets in which agents search, this top constraint actually will bind frequently for some submarkets.

Instead of Cobb-Douglas, I follow den Haan et al. (2000) and Hagedorn and Manovskii (2008) using the form $m(u, v) = \frac{vu}{(v^{\phi_1} + u^{\phi_1})^{1/\phi_1}}$. This is still homogenous of degree one but has the property that $p(\theta) < 1 \forall \theta > 0$. The downside, however, is that this form assumes an elasticity with respect to θ . Generally, ϕ is chosen to match the average finding rate and then the level of θ implies an elasticity of $p(\cdot)$ with respect to θ . This is problematic because for values of θ which are “realistic”⁸ this elasticity is too high. By modifying the matching function to

$$m(u, v) = \phi_0 \frac{uv}{(v^{\phi_1} + u^{\phi_1})^{1/\phi_1}}, \quad p(\theta) = \phi_0 \frac{\theta}{(1 + \theta^{\phi_1})^{1/\phi_1}} \quad (31)$$

I have both ϕ_0, ϕ_1 to match both level and elasticity of the job finding rate.

5.3.3 Separations

For the distribution of disutility to govern shocks, I use a negative-exponential as the tail with a mass point at a disutility of zero. Most workers experience no disutility from work, but some fraction can, if they stay in the match, experience a great deal.

$$Pr[\xi < x] = \begin{cases} \lambda_0 e^{\lambda_1 x} & \text{if } x < 0 \\ 1 & \text{if } x = 0 \end{cases} \quad (32)$$

The form ensures that there are always some separations, as it has infinite negative support, but also there is a cap for a given type. No more than λ_0 of any type (ℓ, d) will separate. This contrasts with many other models of this sort which have a lower bound on the number of separations but no upper limit. The reason for this modelling assumption is two-fold. First, it allows me to parsimoniously parameterize a distribution that has both level of separations and elasticity with respect to productivity. Second, the inexperienced workers will have a much higher separation rate, and if it is unbounded on the top, then negative

⁸As described in Hagedorn and Manovskii (2008), one can use help wanted listings as a proxy for v and the entire unemployment pool for u and generally the economy-wide θ is approximately 0.63.

productivity shocks affect inexperienced workers by too much and few workers are able to become experienced. Figure 8 plots the distribution in Equation 32 to help visualize the distribution and illustrate the magnitudes involved.

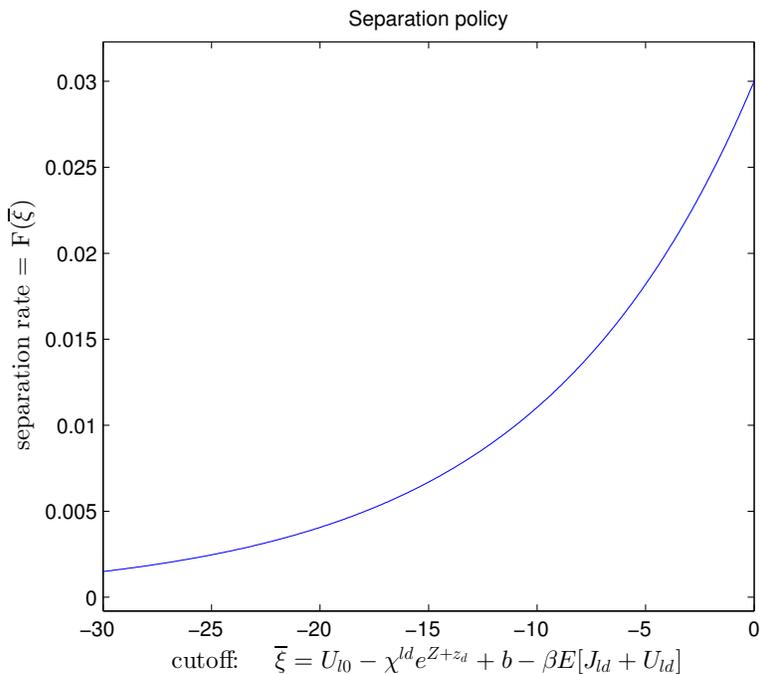


Figure 8: The distribution tail for disutility

5.3.4 Choosing targets

The parameters then that we have to match are $\phi_0, \phi_1, \lambda_0, \lambda_1, \tau, \sigma_\psi$. Though all the parameters affect all of the targets, there are some heuristic relationships. All of these variables are chosen to match the pre-2008 economy. κ and μ also must be set, which I take from the literature. $\kappa = 0.26$ corresponds to that used in Hagedorn and Manovskii (2008), while the firms' share, μ is not easily observable and the literature uses anything in the range of 0.5 to 0.97. I choose a relatively conservative, middle level, 0.66. The results are not particularly sensitive to values anywhere between 0.5 and 0.8, though at very high levels the model becomes difficult to calibrate. The unemployment replacement rate is 40% of the prior working wage, which is approximately average across US states. Rather than targeting the level of separations, the measurement of which suffers from time-aggregation bias as pointed out in Shimer (2012), I choose separations so that the average unemployment rate is 5.5%.

τ , which governs the speed at which workers become experienced and the wage helps the model match the average returns to occupational tenure. Kambourov and Manovskii (2009)

estimate the five year return using PSID data⁹ and correcting for the endogeneity of tenure by instrumenting with average tenure of that occupational stint, a common though imperfect strategy introduced in Altonji and Shakotko (1987). If anything, the returns to tenure are still biased downwards because it may be that unobservable match quality is correlated with labor market experience and tenure is positively correlated with experience. In the context of this model, downwards bias somewhat mitigates the effect of occupational attachment because lower τ implies that fewer become experienced and attached.

σ_ψ , the standard deviation of occupation-specific preference shocks, determines the rate of switching occupations when unemployed. The mapping between this parameter's role in the model and data are not so straightforward, though. For a very low σ_ψ , searchers flock to the single occupation with the largest return and raising it increases the spread of $g^{\ell d}$ regardless of $\{z_d\}$. As described in Section 5.3.1, the standard deviation of shocks, as perceived by workers, actually depends on the finding rate in that occupation. Because the unemployed from some occupations have higher average finding rates than others, they will actually experience larger standard deviation of finding rate shocks. σ_ψ just acts as a baseline, over which endogenous variables act.

For λ_1 , I target the standard deviation of separations across occupations rather than a statistic related to the time-series of the aggregate separation rate. This is to emphasize the role of separations in determining the distribution of unemployed people. To characterize the effect of occupational attachment, we need to have the proper distribution of occupations among unemployed workers. With endogenous separations, when occupation d suffers a low z_d , this increases the number of type d in unemployment through two channels: (1) they have a lower finding rate and (2) they enter unemployment more frequently. Because separations have been made endogenous to get the distribution of unemployed workers correct, I look to the cross-sectional distribution.

5.3.5 Calibration results

Table 3 displays the calibration results. In most dimensions, the model is quite capable of matching the targets. Especially important, parameters of the matching function are matched quite exactly. Of course, these are just average figures and it is up to the model to match facts about the distribution of finding and duration.

⁹They use data from 1968-1993, but the results are not much different when extended to 2006 although changes in the survey make this later data somewhat more suspect.

Target	Model	Data	Source
Change Probability	0.48	0.46	SIPP 1996-2007 (3 Panels)
Job Finding Rate	0.34	0.33	CPS, 1976-2007
Returns to tenure	0.005	0.006	PSID (monthly rate)
Match Elasticity	0.48	0.48	Barichon (2011)
Separation Rate	0.029	0.028	5.5% unemployment
sd(Separation)	0.01	0.01	CPS, 1976-2007

Table 3: Summary of calibration targets

6 Business Cycle Properties

The model provides insight into how unemployment and duration responds to cycles generally. Given the estimated shock process, it generates counter-cyclical dispersion of duration and unemployment rates across occupations. Aggregates, the unemployment rate and duration, also fluctuate more over the cycle than a similarly calibrated Mortensen-Pissarides model with homogenous searchers. Summarily, this is because with heterogeneous searchers in a downturn, some are affected more greatly than others and the fall in their finding rate pulls out the whole duration distribution. This is akin to the logic in Figures 4 and 3, where the negative-sloped finding rate implies duration dynamics that cannot be replicated with homogeneous finding rates.

So, the logic of my results is thus: the finding rate for some occupations falls precipitously and these workers' long duration pulls out the distribution of duration enough to generate realistic levels of long-term unemployment. As discussed in the Section 4.7, in a recession, those who stay searching in their own occupation have slow finding and those who try switching also have especially slow finding because wage compression in recessions make unskilled workers particularly unprofitable. This exacerbates counter-cyclical dispersion. With this heterogeneity in recession, the workers from affected occupations constitute a greater part of the tail of the duration distribution.

To turn to heterogeneity in recession, the prior literature has pointed out two mechanisms by which the economy can generate counter-cyclical risk across segments of the economy, as I have discussed. I generalized Abraham and Katz (1986), making occupations differ in their cyclical sensitivity. In this framework, the variance in productivity increases symmetrically for positive and negative aggregate shocks to Z . Unemployment and duration can be counter-cyclical if the finding rate responds asymmetrically. This asymmetry is built right into the search and matching model through a number of mechanisms. The matching friction implies that hiring in an expansion is slower, while separations are instantaneous in response to a negative shock. In my model, this is exacerbated by occupation-specific skills and

Stat	Model	Data
$\text{corr}(\text{sd}(\log p_{\ell,t}), \log u)$	0.60	0.37
$\text{corr}(\text{sd}(\log u_{\ell,t}), \log u)$	0.63	0.28

Table 4: Business cycle properties of the model, $p_{\ell,t}$ is the finding rate of those coming from occupation ℓ and $u_{\ell,t}$ is the unemployment rate for occupation ℓ

occupational attachment. Here, the composition of unemployed workers affects the finding rate because if there are many attached to low productivity occupations this decreases the effective finding rate, even if there are other high productivity occupations. A negative shock to a particular occupation will increase separations and the pool of searchers in a low finding rate occupation. On the other hand, a positive shock, though some occupations will experience a larger increase in their finding rate than others, is dampened because wages also adjust and the separation rate will not vary so much across occupations.

Table 4 shows the counter-cyclical dispersion in finding rate and unemployment rate. A reader might find it odd that the finding rate’s dispersion has an even higher correlation with unemployment than in the data. There is no reason that the adjustment asymmetries in the model correspond precisely to those in the data. For instance, the elasticity implied by the form of this particular matching function is higher in slack markets. This increases dispersion’s correlation with average productivity. Alternatively, some complementarity between occupations’ output would also reduce the dispersion across occupations. The important aspect is that the model is very close to the data in the relationship between the dispersion across occupations’ finding rate and the cycle—essential for the recessionary increases in long-term unemployment.

6.1 Duration and finding rate heterogeneity

The central target of this paper is the heterogeneity in finding rates engendered by occupation specific human capital. This is neatly summarized by the model’s analogue to Figure 4, which I plot in Figure 9. The finding rate falls from about 36% in the first month to 24% after a year of unemployment. The model misses most in the initial fall in finding rate, where some workers quickly flow out of unemployment but captures the low finding rates of those long in unemployment. This is not to say that none of the searchers in my model find jobs quickly: on average the finding rate for the unattached unemployed, type $(0, 0)$ searching in the most efficient market is 0.47. However, the shocks $\{\psi_{d,i}\}$ make some of them search inefficiently. These shocks allow the model to match switching behavior for the experienced, attached searchers, but the randomness also affects the unattached searchers who would otherwise focus their attention on easier to find jobs.

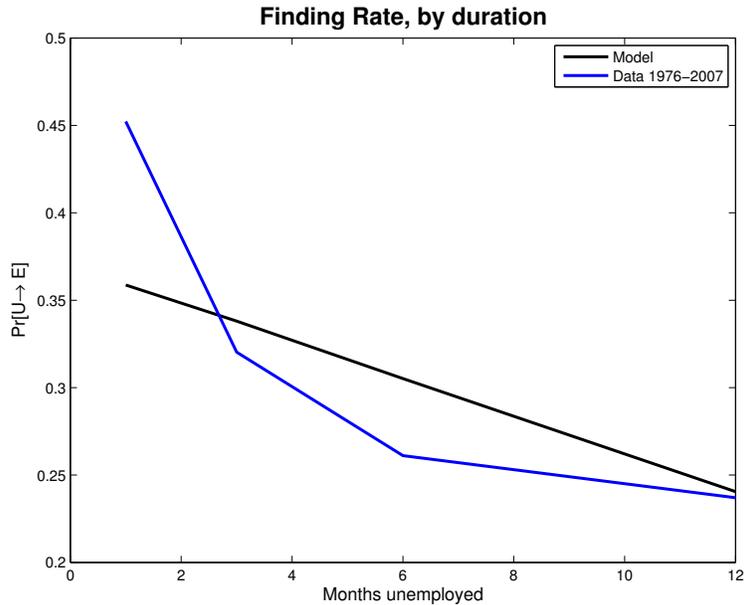


Figure 9: Finding rate by duration with occupation-level heterogeneity implied by the model

The model is fairly successful at matching the expected duration before a match and the cross-sectional standard deviation of duration. For both, the fact that some workers have quite low finding rates at long durations pulls out the moment. At the 10th percentile, the finding rate is only 20%. Essentially, Table 5 shows how the model is relatively successful at matching the tail of the duration distribution.

In Table 5, I present two other models for contrast, a standard Mortensen-Pissarides model and a “No human capital” version. The Mortensen-Pissarides model is calibrated to match the same average finding rate and is presented in the fourth column. The single finding rate means that the expected duration is just $\frac{1}{p_t}$. Because the duration is constrained to be exponentially distributed, its standard deviation is also $\frac{1}{p_t}$. The intermediate model is one in which there are multiple occupations but new hires begin at full productivity even if they came from another occupation. That is, the “No HC” column corresponds to $\omega_{\ell d} = 1 \forall d$. In this case, because $var(\psi_d) > 0$ workers still may apply to occupations which are not the highest productivity.

Notice in Figure 9 that the hazard falls almost linearly. This is a repercussion of the pure composition effect. The slope is not entirely time independent. To see this, consider the continuous time case, if for each i groups, the finding rate is p_i and the initial population when they enter unemployment is $u_{i,0}$ then the hazard is $h(t) = \frac{\sum_i u_{i,0} e^{-p_i t} p_i}{\sum_i u_{i,0} e^{-p_i t}}$ and so the derivative of the hazard is $\frac{-\sum_i p_i^2 u_{i,0} e^{-p_i t}}{\sum_i u_{i,0} e^{-p_i t}} + (h(t))^2$. But, to make the curve more convex, as in the data, it would take either more spread in the finding rates or for p_i to be time-dependent,

as in the case of true duration dependence.

Stat	Model	Data	No HC	MP
E(duration)	3.54	3.73	3.30	3.03
sd(duration)	3.91	4.24	3.66	3.03
Fraction ≥ 6 mo	0.16	0.16	0.12	0.09

Table 5: Duration when matched, implied from CPS 1976-2007

The time series of duration and its cross-sectional variation also matches the data rather well. The correlation between average productivity and average unemployment duration is -0.51, quite counter-cyclical.

6.2 Cyclical volatility and match efficiency

In my model, the variance of unemployment is quite a bit higher than in many other search and matching models, where low volatility is a common problem that was described in Shimer (2005) and addressed by countless papers. Carrillo-Tudela and Visschers (2013) is quite successful in addressing this critique with a model similar to my own. It generates much of fluctuations from “rest unemployment,” and shows that total unemployment volatility increases when reallocation and matching frictions complement each other. Section 4.7 describes why cycles affect some occupations so much worse than others.

The average finding rate fluctuates by more than would be suggested by fluctuations in the average tightness alone. This is because some of the variation comes from changes in the search direction, rather than just congestion. This point is important because it corresponds to findings of Beauchemin and Tasci (2008) and Cheremukhin and Restrepo-Echavarria (2010), who estimate a process for the efficiency of the average matching function and show that it can account for much of the cyclical variation in unemployment rates. My model has endogenous changes in the efficiency of matching because workers may choose their search direction in such a way that reduces their finding chances over the cycle. I define this effective efficiency as a residual

$$\tilde{\phi}_0 = \frac{p \left(\sum_{l=0, d=1}^J \theta^{\ell d} x_{\ell 0} g^{\ell d} \right)}{\sum_{l=0, d=1}^J p(\theta^{\ell d}) x_{\ell 0} g^{\ell d}}$$

This measure is a very clean way to capture the impact of heterogeneity. $\tilde{\phi}_0$ is the “match efficiency,” and it incorporates workers choice to search in markets that are more or less likely to yield a match. Its fluctuations capture the allocative efficiency that is endogenous in this model, as with the same aggregates for searchers and vacancies, we may have very different

finding rates. As can be seen in Table 6, the unemployment and finding rates vary quite a lot over the cycle and much of this is explained by the “efficiency” of the matching function. Clearly, $\phi_1 \neq 1$ implies that $\tilde{\phi}_0 \neq 1$ by Jensen’s inequality, but its procyclicality shows that it plays an important role in unemployment fluctuations in the cycle. Beauchemin and Tasci (2008) calculate a strongly procyclical process that is necessary to match unemployment, shown in Table 6. This mechanism explains some of it.

	Model	Data	MP	No HC
$\text{sd}(\log u_t)$	0.18	0.24	0.021	0.11
$\text{sd}(\log p_t)$	0.11	0.21	0.021	0.06
$\text{corr}(\tilde{\phi}_{0,t}, E_t[z_{d,t}])$	0.439	0.60	0.0	0.11
$\text{sd}(\log \tilde{\phi}_{0,t})/\text{sd}(\log p_t)$	0.205		0.0	0.08

Table 6: Business cycle fluctuations. “Data” for $\tilde{\phi}_0$ corresponds to the accounting procedure in Beauchemin 2012

Of course, my paper is not the first to show that there is significant variation in the hiring rates across occupations, nor to tie this to shifts in the Beveridge curve. Especially convincing, Hobijn (2012), uses novel vacancy data and shows significant heterogeneity in the hires per vacancy across occupations and that the increase in this heterogeneity during the Great Recession and its aftermath explains a great deal of the shift in the Beveridge curve and slow down in average hiring. Barnichon and Figura (2011) also evaluate

My result is also similar in spirit to the exercise in Sahin et al. (2012), though my work does not yet directly calculate their measure of “mismatch.” They discuss how allocative efficiency across occupations, locations or industries maps into match efficiency. The difference being that their cyclical volatility is affected by either a random process for posting cost or process for vacancies themselves. In their setup, these components ensure that vacancies will be sufficiently volatile, a dimension of the data that I do not target.

6.3 Different wage setting regimes

In this section, I explore different modes for setting wages within my model. In the baseline version, I set wages by Nash bargaining with a fairly low bargaining weight to the firm (0.67). We know that this makes wages fluctuate more than their data counterpart and vacancies less so, but this parameterization is also the majority in such models.

Wage setting is an important margin on which to manipulate the model for two reasons. First, my experience with cross-sectional dispersion in different wage setting regimes can give insight into others’ attempts to explain time-series variation. Second, the mapping from the model to occupational switchers’ wage loss in the data depends upon the wage

setting assumption. Because this is a crucial source of identification, it is important to test the results' robustness in this dimension.

After Shimer (2005) fixated the search and matching literature on the MP model's inadequacy in terms of cyclical fluctuations, assumptions on wage setting have garnered considerable attention. This critique holds that the time series of average tightness only fluctuates by a twentieth as much as in the data, given realistic aggregate productivity.

[To be written]

6.4 Expiring Unemployment Benefits

Now consider an economy in which $\delta = \frac{1}{5}$, which is approximately the expiration rate of unemployment benefits during normal times in the US. Workers who are employed expect their benefits to last for 6 months of unemployment.¹⁰ Qualitatively, the effect is to diminish the value of unemployment, lower workers' outside option and increase the matching rate on average¹¹. As was the intuition with switchers' unemployment duration earlier, a larger surplus with the same size of shocks will imply less variation in the vacancy posting and unemployment rates.

Table 7 summarizes the changes to important results in the paper. Workers with yet unexpired benefits change their behavior and also workers whose benefits have expired are more likely to switch occupations and generally appear different. A bit less than one in six unemployed workers has expired benefits and their finding rate is generally higher (holding constant the composition) by 7 percentage points.

Stat	Baseline	6 Mo Expiration	Data
E(duration)	3.54	3.49	3.73
Fraction ≥ 6 mo	0.16	0.15	0.16
sd(log u_t)	0.18	0.16	0.24

Table 7: The effect of benefits' expiration

Quantitatively, the change matters. However, it is not obvious that this does not overstate the case of the effect of unemployment benefits expiration. The issue is two-fold: (1) random expiration does not necessarily mimic the experience of workers who know the expiration of

¹⁰Note the timing issue here: workers separate at the beginning of the period but do not face expiration risk. They do not enter "expired" status until the next period after unsuccessfully searching and being unemployed for a month. Thus, an expected period of benefits that is 6 months implies an expiration rate of $\frac{1}{5}$ for workers who enter the period in unemployment.

¹¹Though, because I re-calibrate the model, the average matching rate stays the same

their benefits and (2) the fall in consumption due to benefits expiration is significant but not as great as the fall in income would suggest.

Random expiration of unemployment benefits is computationally convenient but for the unemployed their decision always incorporate the risk that next month they may lose benefits. This means that the value of unemployment for short unemployed workers is lower than it would be given a more realistic, deterministic expiration. This point applies to all such models with random expiration back to Fredriksson and Holmlund (2001).

Though the expiration of benefits may entail a significant fall in income, the concomitant fall in consumption may be quite small. Several papers, most canonically Gruber (1997) estimate a fall in consumption expenditures between $\frac{1}{5}$ to 13 if there were no unemployment benefit. This is significant, but not the same scale as the reduction of income from unemployment benefits to food stamps. Furthermore, Browning and Crossley (2009) show that the composition of expenditures change, i.e. away from durables, and that further reduces the fall in welfare terms.

7 The Great Recession

The Great Recession is a perfect test of the model out of sample. As has been hopefully made clear, the model's parameters are based entirely on data to 2007. It was successful at generating longer-term unemployed in regular cycles. But how does it fare in the extreme conditions recently observed?

Figure 10 (left) plots the dispersion in log-industry productivity. Each industry has its mean and linear trend removed, as in the data fed into the model. Notice that the dispersion in industry-specific productivity is at its highest since the post World War II recessions. In my model, the productivity difference between the best and least productive, the ratio of 90th to 10th percentile, will contribute to the elongation of the tail of the duration distribution.

The fall in average productivity was quite sever, though rather short. However, the rise in dispersion preceded the recession and continued even afterwards. This observation, that dispersion increases even in 2006 echoes Barnichon and Figura (2011), who find an increase in labor market dispersion that presages the Great Recession.

7.1 Setting up the experiment

My experiment takes from the data the distribution of workers, $\{x\}$ and the distribution of shocks \mathcal{Z} . Some manipulations are required to map these observables into model objects because the data on worker experience is not detailed enough and the productivity shocks

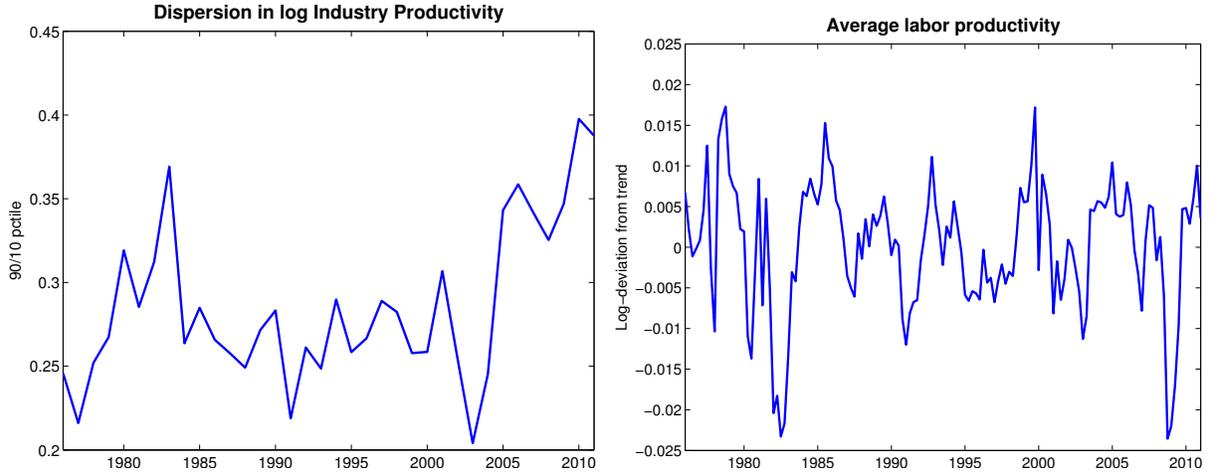


Figure 10: The dispersion of industry productivity is peaks in this recession and average productivity hits a nadir

are only observed through per capita output.

To set up the experiment with x from the data, distributions of both employed and unemployed are necessary because the distribution of employed will determine the distribution of unemployed as the economy evolves over the Great Recession. The problem is that in the SIPP, I only observe a workers occupation and occupation tenure. I would also need to know the prior occupation to map one for one into the model, but this is unavailable. First, I take the number of experienced workers directly from the data, x_{dd} . These are workers who have tenure greater than or equal to the average number of months it takes to become experienced in the model, $\frac{1}{\tau}$. Then, I have some group of workers observed in occupation d but who are inexperienced. I assign them to prior occupations ℓ to match the steady state distribution. I set $\frac{x_{ss,jd}}{x_{ss,\ell d}} = \frac{x_{0,jd}}{x_{0,\ell d}}$, where ss means the steady state and 0 is the beginning of the experiment. To set x_{00} , I take the unemployed workers whose prior occupation tenure was less than $\frac{1}{\tau}$.¹²

Next, I need to ensure that the distribution x from the data is not going to drift on its own because the model's steady-state population in each occupation is different from that in the data. For example, there are some occupations that are very small in the data, but are closer to average size in my model. If I were to just start a simulation with the occupation population from the data, far fewer workers than the model's steady state, there will be a natural drift within the model to attract workers there. I can solve this problem in two ways: (1) I can change the average productivity in each occupation so that the steady state populations correspond between model and data or (2) I can change the size of the population

¹²Not every worker reports his prior tenure, so I take the fraction with less than τ of those who report a tenure. This assumes that there is no systematic trend in not reporting occupational experience.

I draw from the data before feeding it into the model. I take the second course. I adjust every occupation’s population so that the number I feed into the model is the deviation from it’s steady state, multiplying by $\frac{x_{LR,d}^{data}}{\sum_{\ell} x_{ss,\ell d}}$, the data’s long-run average for that occupation’s fraction of the labor market over the model’s steady state for the same statistic.

To map productivity into the model, I again have to take a structural approach. I can observe output per capita annually and by industry. I take the observations for 2008, 2009, 2010 and interpolate between them. Rather than simple, deterministic interpolations, I sample from the estimated monthly process and trace it out to connect the observed points. I take 100 draws of these 24 month periods. For each history, I solve for the underlying productivity in much the same way as in the other simulations. I require that the model-generated composition of workers and the primitive shock imply observed productivity.

7.2 Results

Feeding into the model the set of productivity shocks and distribution of unemployed workers I can, indeed recover a large rise in the rate of long-term unemployment and unemployment duration generally. Notice also that the baseline Mortensen-Pissarides model has almost no response. This is because average productivity moved little, as noted by McGrattan and Prescott (2012). But, even though average productivity moved little, this masks considerable heterogeneity between occupations. In fact, the 90th percentile grew while the median and 10th percentile of occupations’ productivity shrank. In particular, the productivity recovery in 2010 occurred largely at the top of the distribution. This increase in skewness, Kelley’s measure went from XX to XX adds some detail to the increase in dispersion shown in Figure 10.

In the context of the model, an increase in the spread of productivity creates pressure for workers to reallocate. However, the friction associated with occupation-specific skills interferes. As shown in Table 8, average unemployment duration rises in the model, but long-run unemployment is especially prominent as in the data.

	Model	Data	MP	no HC
Expected Duration	3.85	4.36	3.07	3.42
Standard Deviation	4.31	4.60	3.07	3.42
Fraction \geq 6months	0.26	0.30	0.094	0.12

Table 8: Unemployment duration during the Great Recession

Notice especially that the model’s increase in long-term unemployment nearly keeps pace with the rise seen in the data, but average duration is not as great. Long-term unemployment in the model rose by so much because the spreading of productivity also entailed a spreading

of unemployment duration across occupations and the occupations that were affected were affected badly. However, average duration rose by less because at the short end, the finding rate did not fall by much.

Figure 11 tells some of this story. As mentioned, the model’s finding rate falls by less and is less “kinked” due, in part, to my abstraction from duration dependence. This seems to be the explanation for my model’s lower mean duration but relatively high incidence of long-term unemployment. Consider the case in which all of the downward slope were due to duration dependence. Before 6 months, the finding rate was relatively high and after 6 months it went very near zero. This would mean that whomever did not first find a job is now the population of long-term unemployed, and that may be a small number, but the mean unemployment rate is very high. In my model, however, all finding rate heterogeneity comes from differences in the composition of unemployed, and so even a long durations there are still many who have fairly high finding rates and have merely been repeatedly unlucky.

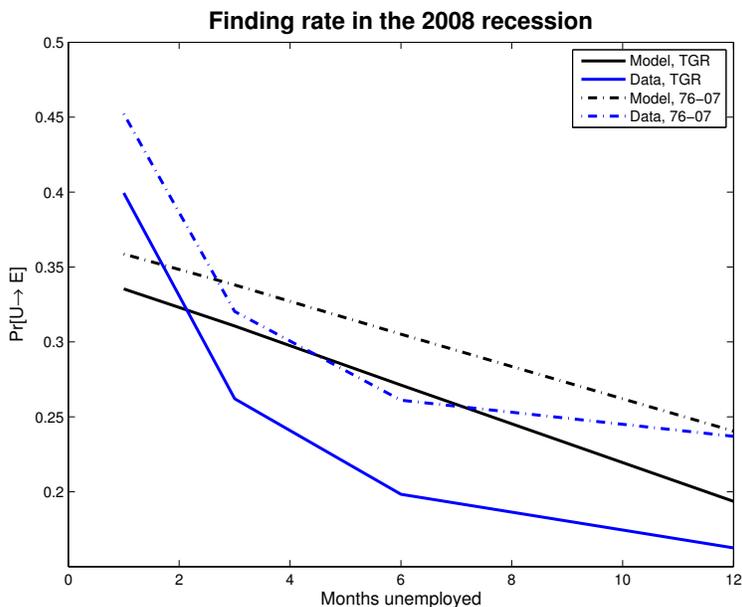


Figure 11: The finding rate rotates downward. but captures only some of the heterogeneity

7.3 Time to switch occupations

Crucially, it takes longer to switch occupations in a recession. This holds in the data and in the model. One way to measure and quantify this effect is a simple regression on the data:

$$\text{weeks unemployed} = D_{occ_t \neq occ_{t-1}} + \text{demographics} + \text{const} + \epsilon$$

I include the industry switching dummy for illustrative purposes, to again remind readers that occupations are a useful unit of analysis. Notice that the industry switchers do not have significantly longer durations during the most recession. The crucial coefficients are shown in Table 9. Conditional on many demographic factors ¹³ workers who switch occupations are unemployed for longer than those who do not. For the full regression, see Table 13.

	1978-2007		2008-2010	
	Coef	Std Err	Coef	Std Err
Data Dummy Occ Switch	7.94	(0.302)	16.16	(0.889)
Model Dummy Occ Switch	6.13		13.74	

Table 9: Marginal effect in weeks of an occupation switch

7.4 The Unemployment Rate

The unemployment rate rises less drastically in the model than the data, but by more than would be predicted by a simple search and matching model. This would be suggested by the prior result that unemployment is more volatile in my model than other search and matching models and it is also consistent with the result suggested by Andolfatto and MacDonald (2004). As Figure 10 (Right) suggests, there was a large fall in average productivity. In Figure 10 (Left) we see this was accompanied by an increase in dispersion that lasted even longer. As seen in Table 10 this generates an increase in unemployment rate, though not as large as was actually seen. Consistent with the design of the experiment, unemployment starts at 5.2%, actually below the calibrated steady state, and then we see that it rises considerably. These figures are just about in-line with the upper bound found in Sahin et al.

Stat	Model	Data
Average u_t	6.72	8.23
Peak u_t	7.02	10.0

Table 10: The effect of benefits' expiration

(2012) when they partly endogenize the posting of vacancies.

7.5 Unemployment benefits extensions

As discussed in section 6.4, expiring unemployment benefits has a sizable impact on the behavior of the model because it affects searchers' outside options. Hence, when the duration

¹³The controls are: race marital status, sex, age, high school and college dummies and an annual trend. The trend is estimated on 1978-2007 and applied to the 2008-2010 data.

of unemployment benefits increases, it non-trivially increases the workers' outside option and encourages them to search more narrowly and for longer. Or, to look from the other side, it makes it more difficult to pay a worker to switch occupations when his outside option rises.

In the Great Recession, several studies have found that the effect from benefits extension on unemployment is non-trivial. Nakajima (2012) chalks 1.5 percentage points to the extension of unemployment benefits.

One of the key features of a structural model is that I can use it for policy experiments. Here, I look at the effect of the extension of unemployment benefits and perform a counterfactual experiment wherein I do not extend the benefits from 6 to 12 months. The results of this experiment are shown in Table 11. Note that, in the new calibration with $\delta = \frac{1}{5}$, the unemployment rate goes up considerably when the unemployment benefits are extended to $\delta = \frac{1}{11}$.

Stat	Baseline	12 Mo	6 Mo	Data
E(duration)	3.85	3.87	3.76	4.36
Fraction \geq 6 mo	0.26	0.27	0.22	0.30
Average u_t	6.72	6.91	6.27	8.23

Table 11: The effect of benefits' expiration

8 Conclusion

The unprecedented rise in unemployment duration during the Great Recession has focused attention onto its causes. One often cited explanation blames a mismatch between skills in demand and those in supply from the unemployed. In this line of thinking, the long-term unemployed are composed of workers with skills in particularly low demand. But, standard theory is ill-suited to quantitatively evaluate this explanation. One needs heterogeneity amongst the unemployed and, more importantly, this heterogeneity must be along workers' skills.

This paper introduces a model with workers who differ by their prior occupation, which summarizes their skills. With discipline from data on the wage premium to occupational experience and the process for occupation-specific shocks, it can quantify the link between unemployment duration and skills. In the model, workers' crucial choice is whether to search in their own occupation where they command a premium for their experience. But, for workers from hard-hit occupations, searching for jobs in their old occupation will extend their duration in unemployment.

I began with a data exercise, I showed differences across the unemployed in their job finding rates are required for understanding the level and cyclical variation of unemployment duration. In any model in which the rate at which workers find a job is the same and matches aggregate data, the implied duration will be lower and less volatile than in the data. Prior occupation is a useful margin along which to divide searchers. The unemployment rate and average unemployment duration vary substantially across occupations and their dispersion increases during recession. Moreover, other studies have suggested that occupational skills are quite important, enough to motivate these differences.

The model augmented a standard Mortensen-Pissarides model with occupation-specific skills and shocks. This meant that unemployed workers directed their search, balancing the wage upon matching and their probability of finding that job. Because workers might choose to search in relatively low-finding rate occupations, match efficiency was endogenous and, as we showed pro-cyclical. Generally, this model delivered realistic business cycle fluctuations in unemployment duration and the counter-cyclical dispersion across occupation. In particular, the incidence of long-term unemployment was nearly identical between model and data. Finally, I applied the model to the Great Recession. In this experiment, I took as given the distribution of occupational skills amongst the unemployed and the occupation-specific shocks. This was meant to measure the effect of the skills of the unemployed and observed shocks, which spread in the recession. These factors greatly increased long-term unemployment, nearly 70% of the total observed increase.

In the future, the model can be applied to a number of policy experiments. It is a good laboratory to study the effects of targeted interventions. Much of government spending affects certain occupations much more than others, such as building projects that affect construction workers. In this model, unlike models without a notion of occupations, I can assess the affect of such asymmetric intervention on long-term unemployment. I also plan to assess the Obama administration's answer to long-term unemployment, whereby they encourage employers to hire and train them. In the "Bridge to Work" plan, employers can hire and train workers without paying them. While this would encourage vacancies to be posted for such workers, workers might not want to apply for such jobs, which is an effect my directed search model incorporate. Its mechanism can help disentangle these effects and assess the policy's effects.

References

ABRAHAM, K. G. AND L. F. KATZ (1986): "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?" *Journal of Political Economy*, 94, 507–22.

- ALTONJI, J. G. AND R. A. SHAKOTKO (1987): “Do Wages Rise with Job Seniority?” *Review of Economic Studies*, 54, 437–59.
- ALTONJI, J. G., A. SMITH, AND I. VIDANGOS (2009): “Modeling Earnings Dynamics,” NBER Working Papers 14743, National Bureau of Economic Research, Inc.
- ALVAREZ, F. AND R. SHIMER (2011): “Search and Rest Unemployment,” *Econometrica*, 79, 75–122.
- ANDOLFATTO, D. AND G. MACDONALD (2004): “Jobless Recoveries,” Macroeconomics 0412014, EconWPA.
- BARNICHON, R. AND A. FIGURA (2011): “What drives matching efficiency? a tale of composition and dispersion,” Finance and Economics Discussion Series 2011-10, Board of Governors of the Federal Reserve System (U.S.).
- BEAUCHEMIN, K. AND M. TASCI (2008): “Diagnosing labor market search models: a multiple-shock approach,” Working Paper 0813, Federal Reserve Bank of Cleveland.
- BHAT, C. R. (1995): “A heteroscedastic extreme value model of intercity travel mode choice,” *Transportation Research Part B: Methodological*, 29, 471 – 483.
- BLINDER, A. S. (2009): “How Many US Jobs Might be Offshorable?” *World Economics*, 10, 41–78.
- BOSKIN, M. J. (1974): “A Conditional Logit Model of Occupational Choice,” *Journal of Political Economy*, 82, 389–98.
- BROWNING, M. AND T. F. CROSSLEY (2009): “Shocks, Stocks, and Socks: Smoothing Consumption Over a Temporary Income Loss,” *Journal of the European Economic Association*, 7, 1169–1192.
- CARRILLO-TUDELA, C. AND L. VISSCHERS (2013): “Unemployment and Endogenous Reallocation over the Business,” Working Paper.
- CHEREMUKHIN, A. A. AND P. RESTREPO-ECHAVARRIA (2010): “The labor wedge as a matching friction,” Working Papers 1004, Federal Reserve Bank of Dallas.
- CLARK, K. B. AND L. H. SUMMERS (1979): “Labor Market Dynamics and Unemployment: A Reconsideration,” *Brookings Papers on Economic Activity*, 10, 13–72.
- DAGANZO, C. (1979): *Multinomial probit: the theory and its application to demand forecasting*, Economic theory, econometrics, and mathematical economics, Academic Press.

- DEN HAAN, W. J., G. RAMEY, AND J. WATSON (2000): “Job Destruction and Propagation of Shocks,” *American Economic Review*, 90, 482–498.
- ELSBY, M., B. HOBIJN, AND A. SAHIN (2008): “Unemployment Dynamics in the OECD,” NBER Working Papers 14617, National Bureau of Economic Research, Inc.
- FERNNDEZ-VILLAYERDE, J. AND J. RUBIO-RAMREZ (2010): “Macroeconomics and Volatility: Data, Models, and Estimation,” NBER Working Papers 16618, National Bureau of Economic Research, Inc.
- FREDRIKSSON, P. AND B. HOLMLUND (2001): “Optimal Unemployment Insurance in Search Equilibrium,” *Journal of Labor Economics*, 19, 370–99.
- FUJITA, S. (2011): “Declining Labor Turnover and Turbulence,” Working Papers 11-44, Federal Reserve Bank of Philadelphia.
- GRUBER, J. (1997): “The Consumption Smoothing Benefits of Unemployment Insurance,” *The American Economic Review*, 87, 192–205.
- HAGEDORN, M. AND I. MANOVSKII (2008): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited,” *American Economic Review*, 98, 1692–1706.
- HECKMAN, J. (1991): “Identifying the hand of past: Distinguishing state dependence from heterogeneity,” *The American Economic Review*, 81, 75–79.
- HECKMAN, J. J. AND B. SINGER (1984): “Econometric duration analysis,” *Journal of Econometrics*, 24, 63–132.
- HOBIJN, B. (2012): “The Industry-Occupation Mix of U.S. Job Openings and Hires,” Working Paper Series 2012-09, Federal Reserve Bank of San Francisco.
- HORNSTEIN, A. (2012): “Accounting for Unemployment: The Long and Short of It,” Working paper, Federal Reserve Bank of Richmond.
- HOSIOS, A. (1994): “Unemployment and vacancies with sectoral shifts,” *The American Economic Review*, 124–144.
- HOSIOS, A. J. (1990): “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies*, 57, 279–98.
- KAMBOUROV, G. AND I. MANOVSKII (2009): “Occupational Specificity Of Human Capital,” *International Economic Review*, 50, 63–115.

- LAZEAR, E. AND J. SPLETZER (2012): “The United States Labor Market: Status Quo or A New Normal?” Tech. rep., National Bureau of Economic Research.
- LILIEN, D. M. (1982): “Sectoral Shifts and Cyclical Unemployment,” *Journal of Political Economy*, 90, 777–93.
- LJUNGQVIST, L. AND T. J. SARGENT (1998): “The European Unemployment Dilemma,” *Journal of Political Economy*, 106, 514–550.
- LOMBARDO, G. AND A. SUTHERLAND (2007): “Computing second-order-accurate solutions for rational expectation models using linear solution methods,” *Journal of Economic Dynamics and Control*, 31, 515–530.
- LUCAS, R. J. AND E. C. PRESCOTT (1974): “Equilibrium search and unemployment,” *Journal of Economic Theory*, 7, 188–209.
- MACHIN, S. AND A. MANNING (1999): “The causes and consequences of longterm unemployment in Europe,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 3 of *Handbook of Labor Economics*, chap. 47, 3085–3139.
- MALIAR, L., S. MALIAR, AND S. VILLEMOT (2011): “Taking Perturbation to the Accuracy Frontier: A Hybrid of Local and Global Solutions,” Dynare Working Papers 6, CEPREMAP.
- MARIMON, R. AND F. ZILIBOTTI (1999): “Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits,” *Economic Journal*, 109, 266–91.
- MCGRATTAN, E. R. AND E. C. PRESCOTT (2012): “The labor productivity puzzle,” Working Papers 694, Federal Reserve Bank of Minneapolis.
- MEHROTRA, N. AND D. SERGEYEV (2012): “Sectoral Shocks, the Beveridge Curve and Monetary Policy,” Tech. rep., Columbia University.
- MILLER, R. A. (1984): “Job Matching and Occupational Choice,” *Journal of Political Economy*, 92, pp. 1086–1120.
- MOEN, E. R. (1997): “Competitive Search Equilibrium,” *Journal of Political Economy*, 105, 385–411.
- NAKAJIMA, M. (2012): “A quantitative analysis of unemployment benefit extensions,” *Journal of Monetary Economics*, –.

- PISSARIDES, C. A. (2000): *Equilibrium Unemployment Theory, 2nd Edition*, vol. 1 of *MIT Press Books*, The MIT Press.
- SAHIN, A., J. SONG, G. TOPA, AND G. VIOLANTE (2012): “Mismatch Unemployment,” NBER Working Papers 18265, National Bureau of Economic Research, Inc.
- SCHAAL, E. (2012): “Uncertainty, Productivity and Unemployment during the Great Recession,” Working Paper.
- SHIMER, R. (2005): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies,” *American Economic Review*, 95, 25–49.
- (2012): “Reassessing the Ins and Outs of Unemployment,” *Review of Economic Dynamics*, 15, 127–148.
- SMITH, A. A. J. AND M. KEANE (2004): “Generalized Indirect Inference for Discrete Choice Models,” Econometric Society 2004 North American Winter Meetings 512, Econometric Society.
- STOLE, L. A. AND J. ZWIEBEL (1996): “Intra-firm Bargaining under Non-binding Contracts,” *Review of Economic Studies*, 63, 375–410.
- SULLIVAN, P. (2010): “Empirical evidence on occupation and industry specific human capital,” *Labour Economics*, 17, 567–580.
- ZHANG, H., A. CONN, AND K. SCHEINBERG (2010): “A derivative-free algorithm for least-squares minimization,” *SIAM Journal on Optimization*, 20, 3555–3576.

A Choice probabilities

Exponentially distributed shocks are somewhat non-standard in additive random utility models, like my own. The issue is that, my additive random utility is not strictly iid, because the matching probability makes the shocks heteroskedastic. With heteroskedasticity and non-zero mean, other common distributional assumptions are quite algebraically inconvenient. In this section I derive the distribution in Equation 30 by induction. Note that this is not exactly the same as the problem discussed in Daganzo (1979) because the mean is not necessarily zero.

We wish to show that the probability of location d being optimal given choices $j \in \{1, \dots, J\}$ where return in j is $R_j + \tilde{\psi}_j$ where $\Pr[\psi \leq \tilde{\psi}_j] = 1 - e^{-\tilde{\psi}_j/\sigma_j}$. To set notation, let the density of the shock in direction j be $f_j(\tilde{\psi}_j)$. In this proof, I will show that

$$\Pr[d = \operatorname{argmax}_{j \in \{1, \dots, J\}} (R_j + \psi_j)] = \prod_{j=1, \dots, J} \frac{\sigma_d}{\sigma_d + \sigma_j} e^{(R_d - R_j)/\sigma_d}$$

Basis step:

Suppose there are two options d and j with return $R_d + \tilde{\psi}_d, R_j + \tilde{\psi}_j$ and distribution parameters σ_d, σ_j .

Then

$$\begin{aligned} \Pr[R_d + \psi_d \geq R_j + \psi_j] &= \int_0^\infty \int_{R_j + \tilde{\psi}_j - R_d}^\infty f_d(\tilde{\psi}_d) f_j(\tilde{\psi}_j) d\tilde{\psi}_d d\tilde{\psi}_j \\ &= \int_0^\infty \frac{1}{\sigma_j} e^{-\tilde{\psi}_j/\sigma_j} \left(e^{-(R_j + \tilde{\psi}_j - R_d)/\sigma_d} \right) d\tilde{\psi}_j \\ &= e^{(R_d - R_j)/\sigma_d} \left(\int_0^\infty \frac{1}{\sigma_j} e^{-\tilde{\psi}_j/\sigma_j} e^{-\tilde{\psi}_j/\sigma_d} \right) d\tilde{\psi}_j \\ &= e^{(R_d - R_j)/\sigma_d} \left(\int_0^\infty \frac{1}{\sigma_j} e^{-\tilde{\psi}_j(\sigma_j + \sigma_d)/\sigma_d \sigma_j} \right) d\tilde{\psi}_j \\ &= e^{(R_d - R_j)/\sigma_d} \frac{1}{\sigma_j} \frac{\sigma_j \sigma_d}{\sigma_d + \sigma_j} \left(\int_0^\infty \frac{\sigma_d + \sigma_j}{\sigma_j \sigma_d} e^{-\tilde{\psi}_j(\sigma_j + \sigma_d)/\sigma_d \sigma_j} \right) d\tilde{\psi}_j \\ &= e^{(R_d - R_j)/\sigma_d} \frac{\sigma_d}{\sigma_d + \sigma_j} \end{aligned} \tag{33}$$

Induction:

Suppose that for $j = 1, \dots, J - 1$, the probability that $\Pr[d] = \operatorname{argmax}_{j \in \{1, \dots, J-1\}} (R_j + \psi_j)] = \prod_{j=1, \dots, J-1} \frac{\sigma_d}{\sigma_d + \sigma_j} e^{(R_d - R_j)/\sigma_d}$ then for the J instance

$$\begin{aligned} \Pr[R_d + \tilde{\psi}_d \geq R_j + \tilde{\psi}_j \ \forall j = 1, \dots, J] &= \Pr[R_d + \tilde{\psi}_d \geq R_j + \tilde{\psi}_j \ \forall j = 1, \dots, J - 1] * \Pr[R_d + \tilde{\psi}_d \geq R_J + \tilde{\psi}_J] \\ &= \prod_{j=1, \dots, J-1} \frac{\sigma_d}{\sigma_d + \sigma_j} e^{(R_d - R_j)/\sigma_d} * \Pr[R_d + \tilde{\psi}_d \geq R_J + \tilde{\psi}_J] \\ &= \prod_{j=1, \dots, J-1} \frac{\sigma_d}{\sigma_d + \sigma_j} e^{(R_d - R_j)/\sigma_d} * \frac{\sigma_d}{\sigma_d + \sigma_J} e^{(R_d - R_J)/\sigma_d} \\ &= \prod_{j=1, \dots, J} \frac{\sigma_d}{\sigma_d + \sigma_j} e^{(R_d - R_j)/\sigma_d} \end{aligned} \tag{34}$$

The first line is true because $\tilde{\psi}$ is independent for each direction.

B Planner's Problem

To establish efficiency results, I will set up the Planner's Problem, show that it has a unique allocation and then show that the competitive equilibrium with Nash bargained wages solves that. The Social Planner's value function is W and he chooses search direction, tightness and separation policy in each labor market.

$$W(x, \mathcal{Z}) = \max_{\{\theta^m, g^m, \xi^m\}_{m \in M}} \sum_{l=0}^J \sum_{d=1}^J x'_{ld1} \omega_{ld} z_d + \left(x_{\ell d1} (1 - \tau \mathbb{I}_{\ell \neq d}) + \tau \mathbb{I}_{\ell \neq d} \sum_{j \neq d} x_{jd1} \right) \int_{\xi_{\ell d1}}^0 \xi h(\xi) d\xi \\ + b_e x'_{l0e} - \sum_m \int_{\psi} g^m a^m (\kappa \theta^m - p(\theta^m) \psi) df(\psi) + \beta E[W(x', \mathcal{Z}')] \quad (35)$$

$$x'_{dd1} = (1 - s^{dd1}) \left(x_{dd1} + \tau \sum_{\ell=0, \ell \neq d}^J x_{\ell d1} \right) + p(\theta^{dd1}) g^{dd1} a^{dd1} + p(\theta^{dd0}) g^{dd0} a^{dd0} \quad (36)$$

$$x'_{\ell d1} = (1 - s^{\ell d1}) (1 - \tau) x_{\ell d1} + p(\theta^{\ell d1}) g^{\ell d1} a^{\ell d1} + p(\theta^{\ell d0}) g^{\ell d0} a^{\ell d0} \quad (37)$$

$$x'_{\ell 01} = \left(1 - \sum_d g^{\ell d1} p(\theta^{\ell d1}) \right) \left((1 - \delta) x_{\ell 01} + s^{\ell \ell 1} \left(x_{\ell \ell 1} + \tau \sum_{j \neq \ell} x_{j \ell 1} \right) \right) \quad (38)$$

$$x'_{001} = \left(1 - \sum_d g^{0d1} p(\theta^{0d1}) \right) \left((1 - \delta) x_{001} + (1 - \tau) \sum_{d=1}^J \sum_{l=0, l \neq d}^J s^{\ell d1} x_{\ell d1} \right) \quad (39)$$

$$x'_{\ell 00} = \left(1 - \sum_d g^{\ell d0} p(\theta^{\ell d0}) \right) x_{\ell 00} + \delta \left(1 - \sum_d g^{\ell d1} p(\theta^{\ell d1}) \right) x_{\ell 01}, \quad \ell = 0, \dots, J \quad (40)$$

$$1 = \sum_{m \in M_{\ell e}} g^m \quad \forall \ell, e \quad \text{and } g^m \geq 0 \quad (41)$$

$$a^m = \begin{cases} x_{\ell 01} + s^{\ell \ell 1} (x_{\ell \ell 1} + \tau \sum_{j \neq \ell} x_{j \ell 1}) & \ell > 0, e = 1 \\ x_{001} + (1 - \tau) \sum_{j=0}^J \sum_{k \neq j} s^{jk1} x_{jk1} & \ell = 0, e = 1 \\ x_{\ell d0} & e = 0 \end{cases} \quad (42)$$

And the process for \mathcal{Z} follows Equations 2-4

Let the multiplier on the law of motion for $x_{\ell de}$ be $v^{\ell de}$. The first order conditions are:

$$\{\theta^m\} : 0 = -\kappa a^m g^m + p'(\theta^m) a^m g^m \left(v^m - (1-\delta)v^{n1} - \delta v^{n0} - \int \psi d\tilde{f}(\psi) \right) \quad (43)$$

$$\{g^m\} : 0 = p(\theta^m) \left(\psi_d + v^m - (1-\delta)v^{n1} - \delta v^{n0} + \kappa \frac{\theta^m}{p(\theta^m)} \right) - \gamma^{\ell e} \quad (44)$$

$$\{\bar{\xi}^m\} : 0 = \left(x_{\ell d 1} (1 - \tau \mathbb{I}_{\ell \neq d}) + \tau \mathbb{I}_{\ell \neq d} \sum_{j \neq d} x_{j d 1} \right) h(\bar{\xi}^m) (\bar{\xi}^m + v^{\ell 0 1} - v^m) \quad (45)$$

$$n1 = (\ell \in m, 0, e \in m), \quad n0 = (\ell \in m, 0, 0)$$

Condition 43 puts the marginal return to a posting, the expected value of a new match against the cost of posting it κ . Condition 44 states that, for any locations to which an applicant applies, the marginal value of an application has to be equalized. Of course, this will not generally be the case, given the linearity of the choice g^m . However, integrating over ψ will give us a distribution of applications, as in the competitive case. And Condition 45 gives the condition on separations, which again weighs the marginal value to the planner of a match against the value of a searcher.

C Efficiency

To show efficiency, I will show that the social planner's policies will be the same as those of the competitive market, conditional on $\mu = \frac{\theta^m}{p(\theta^m)} \frac{\partial p(\theta^m)}{\partial \theta^m} \forall m \in M$. **Note:** This section needs to be completed.

C.1 Market tightness

The crucial aspect here is to show the two first order conditions for posting are equal:

$$\begin{aligned} \kappa &= p'(\theta^m) \left(v^m - (1-\delta)v^{n1} - \delta v^{n0} - \int \psi d\tilde{f}(\psi) \right) \\ &= q(\theta^m) \left(\omega_{\ell d} z_d - \int w^m(0, \psi) d\tilde{f}(\psi) + \beta \Pi_{\ell d}(\cdot) \right) \end{aligned}$$

Which we can separate into two pieces. For $p'(\theta^m) = q(\theta^m)\mu$ I use the beginning assumption. Then, it leaves to be shown that

$$\mu \left(v^m - (1-\delta)v^{n1} - \delta v^{n0} - \int \psi d\tilde{f}(\psi) \right) = \left(\omega_{\ell d} z_d - \int w^m(0, \psi) d\tilde{f}(\psi) + \beta \Pi_{\ell d}(\cdot) \right)$$

Which is essentially a result of Nash bargaining, where the left hand side is a fraction of the total value of the match to the social planner and the right hand side, because of Nash bargaining is the same fraction of the social value of the match.

C.2 Search direction

For any realized match in market $m = (\ell, d, 1)$, let the value of the surplus be $R_m(\psi_m, z, \omega)$, it depends on the utility shock ψ , the long-run productivity ω and productivity shock z . From the previous subsection, we know that separations are efficient so the total value of the surplus is the same whether we consider social planner or competitive settings.

Then for an individual, the first order condition is

$$p(\theta^d)R_w(\psi_d, z_d, \chi^d) = p(\theta^j)R_w(\psi_j, z_j, \chi^j) \quad \forall j, d : g^j, g^d > 0$$

Integrating over ψ , the search direction decision, \bar{g}^m is: $\Pr[p(\theta^m)(R_m + \psi_m) \geq p(\theta^k)(R_k + \psi_k) \quad \forall k \in M_{\ell e}]$ and with Nash bargaining, the probability of applying to market m becomes $\Pr[p(\theta^m)\mu(R_m + \psi_m) + W_b \geq p(\theta^k)\mu(R_k + \psi_k) + W_b \quad \forall k \in M_{\ell e}]$ and this simplifies to the same condition.

D Computation

D.1 Solving the model

In this section I briefly describe the computational methods and considerations to solve the model itself. There are two crucial insights to ease computational burden: 1) without risk aversion, most of the model is linear and 2) the distribution of workers across occupations is not a payoff-relevant state variable for the households or firms. The difficulty is still that all of the shocks, $\{Z, \{z_j\}, \mathbf{f}\}$, are required for every problem. Discretization would be infeasible because even with only 2 values per shock, that would mean that have a total of $2^{N_d + N_f + 1}$ values, where N_d are the number of occupations, N_f the number of unobservable factors and there is one more for the average productivity shock, Z .

Hence, I use a hybrid-approach: a second-order perturbation to approximate the expectations for the value functions and then the actual non-linear decision rules using these approximations. The technique for second-order approximation is described in Lombardo and Sutherland (2007) and the hybrid method is described in Maliar et al. (2011). In the baseline model, I take as dynamic states the total value of the match, that is $U_w(\ell, d, 1, \cdot) + \Pi_{\ell d}(\cdot)$ and the value of the unemployed worker $U_b(\ell, 0, e, \cdot)$. To solve for these approximations, I also

need to approximate the policy functions. Once the value functions have been perturbed around their steady state, to perform simulations I need only have the expectation of these values. I take expectations from the approximated versions of $U_w(\ell, d, 1, \cdot) + \Pi_{\ell d}$ and U_w and then evaluate the true non-linear decision rules for the simulations.

D.2 Estimation and calibration

There are six parameters to calibrate, $\psi_0, \psi_1, \lambda_0, \lambda_1, var(\phi), \tau$. To estimate, I have to consider the vector of parameters governing $\{\omega_{\ell d}\}, \{\beta_i\}$ and the parameters of the productivity process, $\rho_Z, \sigma_\epsilon, \{\lambda_{1,j}, \lambda_{2,j}, \lambda_{Z,j}\}_{j=1}^J, \rho_z, \sigma_\zeta, \Gamma, cov(\eta)$.

I take a three-layer approach to the estimation, using stochastic multi-starts and derivative-free minimizers for the calibration parameters and estimating $\{\beta_i\}$. For each set of these parameters, I use an iterative approach to find coefficients of the productivity process. This separation is convenient because there are so many parameters of the productivity process and it would be onerous to estimate the entire Jacobian and the Hessian, which has few exploitable sparsity patterns.

The stochastic multi-start technique is fairly standard. I use a modified version of the multi-stage single-linkage method in which I make a few heuristic adjustments to the prescribed rules for stopping and cluster-size choices. The inner solvers alternate between a Nelder-Mead implementation with a new derivative-free non-linear least-squares method, as described in Zhang et al. (2010).¹⁴

For the inner-most estimation, the crucial observation is that I can directly estimate the factor process for per-capita output, ignoring the endogenous productivity that will modify the underlying, unobservable process. Call this \tilde{Z}^{data} . I use this process as a first guess and simulate the model to get populations that imply productivity. With these populations, I can solve for what should be the underlying productivity process. Maybe not surprisingly, there is relatively little difference between the implied process and the “true” process. I describe the iterative procedure below:

1. From annual industry data, estimate the monthly process \tilde{Z}^{data}
2. Draw M realized histories from \tilde{Z}^{data}
3. Solve and simulate the model around process \tilde{Z}^{data} with realized history m .
4. Solve for $\{z_{d,t}^1\}$ using

$$z_{d,t}^1 = \tilde{z}_{d,t} \frac{\sum_{\ell} x'_{\ell d}}{\sum_{\ell} \chi^{\ell d} x'_{\ell d}} \quad (46)$$

¹⁴Zhang graciously provided his Fortran code, which interfaced with my C code flawlessly.

5. Estimate \mathcal{Z}^1
6. Solve and simulate the model around $\tilde{\mathcal{Z}}^1$
7. Return to Step 4 until the likelihood converges
8. Store these coefficients and return to Step 3
9. Average the coefficients

E Regression tables

Variables	log(earnings)
age	0.159 (0.00101)
age ²	-0.00197 (1.20e-05)
college	0.0554 (0.00338)
Female	-0.372 (0.00305)
Constant	5.784 (0.00283)
Observations	2,238,830
R-squared	0.020
Standard errors in parentheses	

Table 12: Initial regression to set up auxiliary model. Age and age² normalized to mean zero

Variables	(1) Duration	(2) Duration
White	-2.974*** (0.145)	-3.803*** (0.413)
Married	-0.981*** (0.136)	-1.783*** (0.394)
Sex	-4.547*** (0.126)	-2.504*** (0.365)
Age	0.311*** (0.00507)	0.334*** (0.0138)
$D_{occ.t \neq occ.t-1}$	7.946*** (0.301)	16.16*** (0.889)
$D_{ind.t \neq ind.t-1}$	4.672*** (0.300)	6.631*** (0.885)
Year	-0.000882 (0.00730)	
College	0.775*** (0.212)	0.0366 (0.512)
high school	0.820*** (0.137)	1.461*** (0.453)
Constant	12.73 (14.52)	11.38*** (0.827)
Observations	139,785	22,698
R-squared	0.095	0.174
	1978-2008	2008-2010

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 13: Duration regressions