RESEARCH DIVISION

Working Paper Series

Long-term Unemployment: Attached and Mismatched?

David Wiczer

Working Paper 2015-042A
https://doi.org/10.20955/wp.2015.042

December 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Long-Term Unemployment: Attached and Mismatched?

David Wiczer *

Federal Reserve Bank of St. Louis, wiczerd@stls.frb.org

December 2, 2015

Abstract

In this paper, I quantify the contribution of occupation-specific shocks and skills to unemployment duration and its cyclical dynamics. I quantify specific skills using microdata on wages, estimating occupational switching cost as a function of the occupations' difference in skills. The productivity shocks are consistent with job finding rates by occupation. For the period 1995-2013, the model captures 69.5% of long-term unemployment in the data, while a uniform finding rate delivers only 47.2%. In the Great Recession, the model predicts 72.9% of the long-term unemployment that existed in the data whereas a uniform finding rate would predict 57.8%.

JEL Codes: E24, J24, J64, E32

Keywords: Occupational Choice, Human Capital, Unemployment Duration

*Special thanks to Victor Ríos-Rull and Fatih Guvenen. Thanks also to Alex Bick, Jonathan Heathcote, Dirk Kreuger, Iourri Manovski, Ellen McGrattan, Amanda Michaud, Josep Pijoan-Mas, Kjetil Storesletten, Laura Sunder-Plassmann, the participants of the Chicago Fed Search and Matching Workshop, and seminars at Arizona State University, FRB Richmond, FRB St Louis, Haifa University, Hebrew University, Tel Aviv University, University of California at Santa Cruz, University College London, University of Illinois, University of Wisconsin, Yeshiva University. James Eubanks provided excellent research assistance. All remaining errors are my own. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve System, the Boards of Governors, or the regional Federal Reserve Banks.

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, a uniform finding rate implies only $\frac{3}{4}$ of the time-series standard deviation of long-term unemployment and half the standard deviation in mean duration. Since the start of the Great Recession, the fall in the average finding rate predicts a rate of long-term unemployment of only about 20% while the average during this period has been nearly twice that. 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 difficult search: Jobs in their own occupation may be difficult to get, but they will benefit from their occupation-specific skills and earn a higher wage if they find one. But switching is not a guarantor of fast job finding either because their skills do not perfectly transfer to another occupation and so they are a relatively expensive hire if they reallocate. I show that the data is consistent with this mechanism: Unemployed workers transiting across occupations experience wage losses on average and also there are sizeable differences in the job finding rate across occupation and this dispersion is counter-cyclical. I estimate the degree to which skills can transfer across occupations and the underlying productivity shocks that drive these changes in finding rate. I then allow the model to follow the panel of shocks in the data and observe how it predicts unemployment duration.

As in the data, the model generates finding rates between occupations that are quite variable over the time-series and counter-cyclically dispersed. I target these cyclical properties with a parametrization of the stochastic process such that 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 longer unemployment spells compared to searchers skilled in other occupations. This makes the tail of the distribution of unemployment duration extend by even more than would be implied by the fall in the average finding rate. The central question of the paper is then, quantitatively, how much does this contribute to the rise in long-term unemployment observed during the Great Recession and recessions generally? To put this differently, if recessions affect searchers differently, how much of the rise of long-term unemployment can be accounted for by searchers who are worst affected by shocks during a recession?

To answer this quantitative question, I estimate the wage loss associated with switching occupations after a job loss. Rather than a fixed cost, I estimate the loss as a function of the distance over skill space between the original and destination occupations. This quantifies the motive for workers to search within their own occupation and for employers not to post vacancies for workers from another occupation. On the other side, I measure the occupational productivity that drives fluctuations in the hiring demand in each occupation. On top of these factors, I ensure that I match average statistics on the gross flow of workers across occupations by including preference shocks over occupations. The model then predicts how much difficulty searchers will face in finding another job.

My primary results are that the model implies a long tail to unemployment duration, increasing both the average duration and propensity to long-term unemployment over a benchmark in which workers find jobs at a uniform rate. Over the period 1995-2013, a uniform finding rate under-predicts the mean duration of unemployment—only 66.76% of its value in the data—and the propensity to long-term unemployment—only 47.24% of its value in the data. My model has the same short-term finding-rate dynamics but increases duration of unemployment and the fraction in the tail experiencing long-term unemployment—79.86%

of the mean duration and 69.46% of the long-term unemployment. In 2008-2010,[1] when unemployment duration and long-term unemployment rose precipitously, the model predicts a rate of long-term unemployment that is 72.87% of the value in the data while a uniform fall in the finding rate would predict 57.78%.

I use data on workers employment, occupation and wage histories in the Survey of Income and Program Participation (SIPP) from 1996-2012 to estimate the fall in earnings workers will face if they begin a job in a new occupation. To characterize the skills used by each occupation, which will govern the magnitude of the fall, I use data from the O*NET. However, the change in earnings is observable only if the switch actually occurs, so to deal with this endogeneity, I take a structural approach to this estimation. I use indirect inference to estimate the structural parameters governing wage loss as a function of the difference in skills between the two occupations.

To estimate the process of occupation-specific shocks, I infer the shocks based on the observed finding rate in a destination occupation. I assume that there is some productivity shock in each occupation and this would imply some process for their probability of matching, and to fit the observed matching rates in each occupation I choose a sequence of productivity shocks. I use this tactic for two main reasons. First, occupation-specific productivity is not directly observable because the output of an occupation is not directly measured. Second, even if we could observe output per unit of labor, as with the aggregate statistics, it is well known that these fluctuations do not move the finding rate sufficiently to match the data (see Shimer (2005)). I use finding rate data to solve for the joint process for exogenous shocks to occupational productivity. The specification of these shocks allows aggregate fluctuations to which occupations differ in their exposure.

An important factor to consider for realistic long-term unemployment is that the cross-

[1] Throughout, I will present statistics from the period from 2008-2010 to describe the experience of the Great Recession, though I will also present the actual time-series. The reason is that unemployment duration lags the cycle. A large increase in separations reduces unemployment duration greatly and, by definition, it takes time for these workers to become long-term unemployed. One could reasonably extend the window even further than 2010, but my choice is meant to keep a window mostly coincident with the NBER definition of the Great Recession.

sectional standard deviation of duration is significant and counter-cyclical. The model gets partly there: In the data the correlation between average duration and the standard deviation across occupations is 0.86. In the model, the correlation is 0.94, though it understates the amount of cross-sectional standard deviation is lower in the model than the data, only 77.6% of it. On the other hand, any model with a single rate generates no dispersion. Dispersion is countercyclical because occupations' productivity differs in cyclical 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.

Beyond the baseline scenario, I can use the model to analyze policy changes to the duration of unemployment benefits. In the baseline, unemployment benefits are extended from 6 to 24 months, as happened in the Great Recession. While this probably had risk-sharing motives, within the model it also adversely affects unemployment duration and the overall finding rate. Within the model's structure, I can undo these extensions and see how much less unemployment duration would have risen. I find that without the extensions, unemployment duration in this period would be 0.31 months lower and long-term unemployment would be 2.29 percentage points lower.

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 6.3 and then conclude in Section 7.

2 Related Literature

To understand the underlying shocks that compose a recession, I borrow from a long literature on countercyclical 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 countercyclical, Abraham and Katz (1986) and followers attribute countercyclical dispersion to differences in the cyclical sensitivity. 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. The heterogeneity in cyclical sensitivity in my model 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 coming from the structure of the model.

To understand the Great Recession, several studies have addressed the degree to which the shock was uneven across sectors. Hobijn (2012) posits that the composition of new job postings across occupation and industry has significantly slowed the number of successful new matches in the recession and recovery. Underlying that analysis on job postings, Mehrotra and Sergeyev (2012) use factor analysis to isolate shocks that differentially affect certain sectors. Sahin et al. (2012) connect differences in the unemployment rate across sectors to differences in vacancy postings across sectors. They create a working definition of "mismatch" as the suboptimal allocation of workers relative to observed vacancies across sectors and then measure its affect on unemployment. In a more structural approach, Pilossoph (2014) studies the degree to which variance shocks can increase unemployment in a frictional labor market. Both Sahin et al. (2012) and Pilossoph (2014) show the difficulty in explaining the rise in unemployment during the Great Recession and provide context for the exercise herein, which does not try to explain aggregate unemployment, but instead looks at the dispersion in its effects.

My model builds on the 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 extend the model like Alvarez and Shimer (2011) and Carrillo-Tudela and Visschers (2014) citeLucasPrescott1974 to include within-islands unemployment. The latter is the closest work to my own. In both, workers may be unemployed because they are misallocated across islands or because of search frictions that exist on all islands. In both this paper and Carrillo-Tudela and Visschers (2014), 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 (2014), the nature of shocks and the 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 idiosyncratic shocks are to match quality. I also model endogenous separations differently, closer to 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. In mine, an island's labor market is always open because there is always positive probability someone will apply there, whereas their islands can shut down, which contributes significantly to unemployment in their model.

In my model, unemployment duration and finding rate are negatively-correlated because of composition effects. 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. Clark and Summers (1979), Machin and Manning (1999) and Elsby et al. (2008) show this relationship 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) and Heckman (1991) describe the conditions to separately identify the forces. Starkly, my model is going to abstract from duration dependence and will study only the composition effect generated by the mechanism. This is similar to the mechanisms in Hornstein (2012) or Ljungqvist and Sargent (1998), where unobservable characteristics or human capital obsolescence make some searchers less effective.

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 bears a strong resemblance to the conditional logit model estimated by Boskin (1974). Finally, Altonji et al. (2013) 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, a feature already present in the structure of my model.

3 Descriptive Data on Unemployment Duration

In this section, I will present several features of the data on unemployment duration which the model will then attempt to recover or which should provide insight into its mechanisms. First, I present data on the evolution of unemployment duration. These are not targeted statistics, but I will assess the model's success as it can reproduce them. Then, I will describe the variation in finding rate and duration that is tied to prior occupation and how this is not entirely sufficient without a model to infer the direction workers chose to search.

In the Great Recession, the rate of long-term unemployment and unemployment duration increased exceptionally, as shown in Figures 1 and 2.[2] Notably, this rise went beyond that expected by a uniform fall in the job finding rate, which we plot along-side the observed time series. If all workers find a job at the same rate, the implied unemployment duration

[2] To define unemployment duration actually requires some nuance. The BLS defines unemployment duration as the average time unemployed of the current pool of unemployed people instead of the *expected* time before a match of a currently unemployed person. Throughout this paper, we keep to this definition, to be consistent with our treatment of the data and model. This definition 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 this measure.

is always lower than the true duration because some workers have exceptionally low finding rates which implies long durations and pulls up the average.

To construct these figures and our model targets, I build a sample using individually-linked data from the CPS in the period 1994-2013.[3] Workers who are unemployed but do not report an unemployment duration are dropped as are those who do not report an occupation or cannot be consistently linked (see Flood et al. (2015)). I correct for time aggregation using the method of Elsby et al. (2009), similar to Shimer (2012), and use the finding rate for a monthly interval.[4] To compute duration in the case with a single finding rate, the population with 0 months is computed using the aggregation-corrected separation rate. Then each period, t, the unemployed population with duration d is $u_{d,t} = (1 - F_{t-1})u_{d-1,t-1}$, where F_t is the monthly job finding rate in period t. After aggregating observations, all of the series are seasonally adjusted by taking out the multiplicative monthly factors.

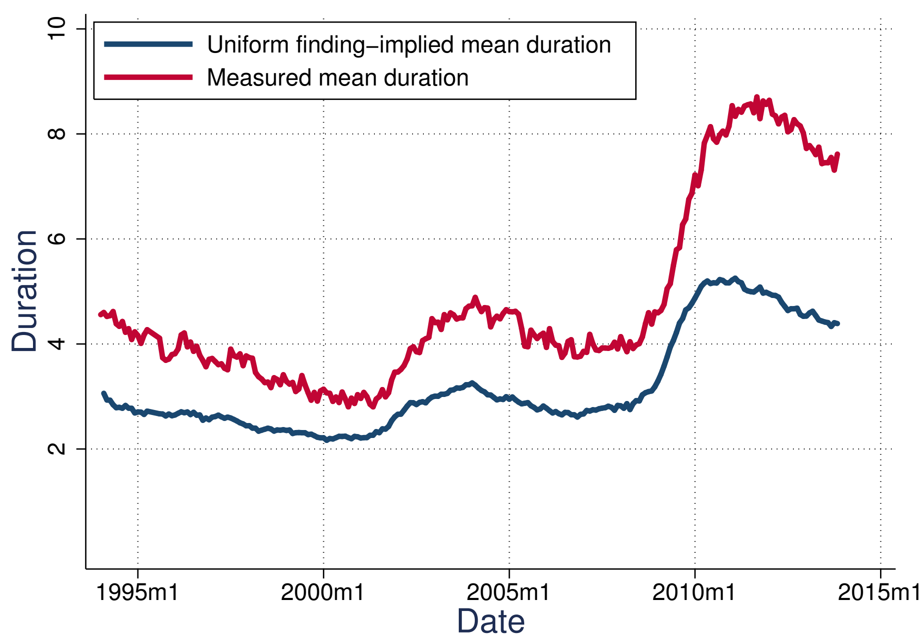

Figure 1: Mean Unemployment Duration, the mean duration of unemployment within the pool of unemployed and that which is implied by a uniform finding rate.

[3]For many of the figures, to construct the unemployment pool's duration we will need a several observations, so we will often start series from 1995 instead of 1994.

[4]In both Elsby et al. (2009) and Shimer (2012) the authors correct for survey design differences before and after 1994. Because my sample begins after the redesign, I do not follow any of these steps.

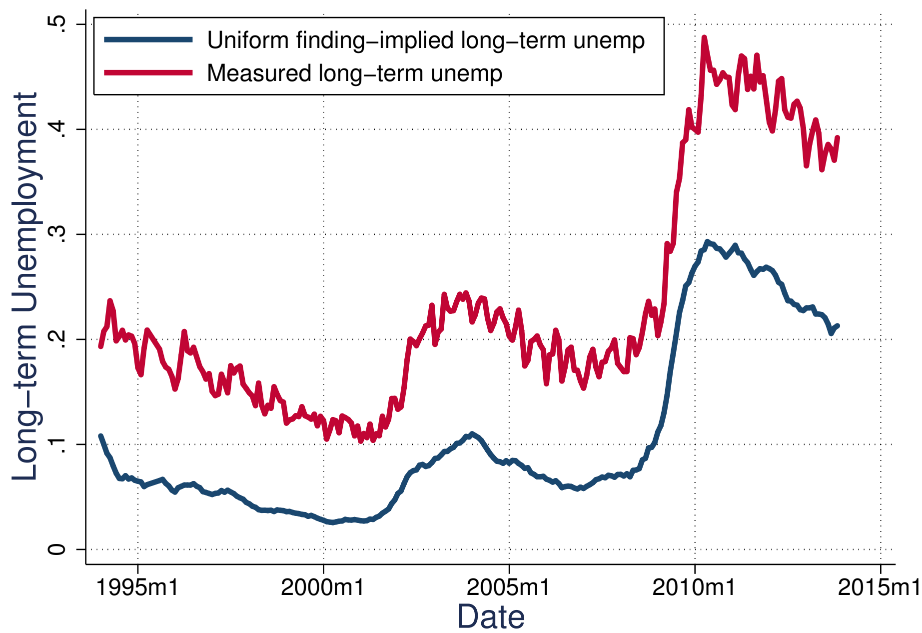

Figure 2: Long-term Unemployment, the fraction of unemployed whose duration is ≥ 6 mo. and that which is implied by a uniform finding rate.

What we are observing in the gap between the statics on unemployment duration and those generated by statistics on job finding is a result of the long-tail of unemployment duration. This is a well-known phenomenon, that the distribution of unemployment duration has a longer tail than the exponential that would be implied by a uniform finding rate. Hence, the average and fraction at long durations are higher than we would observe with a uniform finding rate.

This is statistically and conceptually closely related to duration dependence, in which workers at longer unemployment durations have a lower observed probability of transitioning into employment and this pushes more workers into long unemployment durations, lengthening the distribution's tail (see Clark and Summers (1979), Machin and Manning (1999) or Elsby et al. (2008) for empirical support). Thus, when we look for factors that will skew upwards the distribution of unemployment duration, we should also be looking for how it affects the relationship between finding rate and unemployment duration.

In Figure 3 I plot the change in the monthly finding rate at various unemployment durations in my data sample which shows that indeed the rate is declining. I will use this as a consistency check later with results from the model. The literature on duration dependence cites two potential sources of this slope: composition or treatment. The composition effect is that if different types of workers find jobs at different rates, then their composition will change with duration and gradually shift to more slow job finders. The treatment effect is that being unemployed for longer reduces the worker's ability to find a job. In this paper, duration dependence is exclusively due to composition, though only some of the differences are observable, as we will explain below.

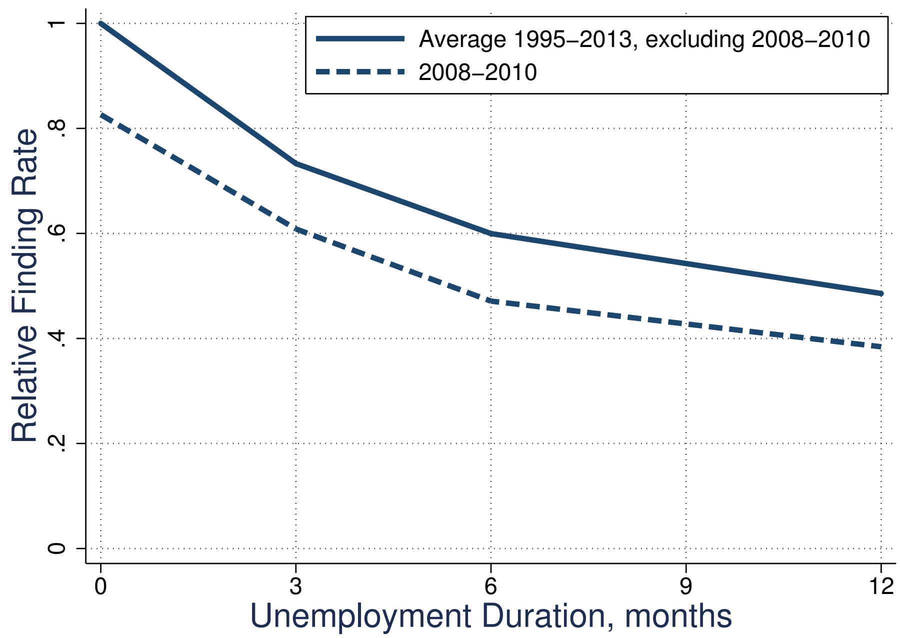

Figure 3: The job finding rate normalized to the average rate at zero months duration.

3.1 Heterogeneity Across Occupations

This paper will link the apparent heterogeneity in finding rates to differences in prior occupation. Once I introduce the structual model, we will see why simple statistical models that condition on past occupation are insufficient, but this section is meant to introduce some of the ideas. I will focus especially on the way that this heterogeneity interacts with business cycles. While other studies, e.g. Hornstein (2012), attribute differences in finding rate to inherent and unobservable heterogeneity, the connection that I make to prior occupation has several advantages. Most importantly, my tactic imposes additional structural by measuring the observably different incentives that guide agents' choices. If observed dispersion in finding rate is due to attachment to one's prior occupation, then this heterogeneity cannot be summarized by some arbitrary factors that are assumed to be policy neutral. Moreover, unemployment duration evolves quite differently in different cycles. With a more structural interpretation, we can understand what conditions affect this. In particular, I highlight how 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 of scholarship has been devoted to the importance of occupation-specific experience and skills. Kambourov and Manovskii (2009) very influentially highlights 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%, results which were amended by Sullivan (2010) but which reconfirmed the overall importance of occupation specific skills. If specific skills are important for wage growth, then potentially they are also important to a worker's experience in unemployment.

Job finding rates and duration are observably different depending on one's prior occupation. From here on, I use the two-digit standard occupational classification (SOC) definition of occupations and make these consistent using the cross-walks provided by Flood et al. (2015). The list of occupations is given in Appendix A.[5] As can be seen in Figures 4 and 5, the exit from unemployment differs substantially and counter-cyclically across occupations. In Figure 4 I show the dispersion in finding rate across occupations. I present this in logs because, as suggested by Elsby et al. (2009), percentage fluctuations are actually more informative than levels about their impact on labor flows. In Figure 5 I plot the differences in

[5] I combined several occupations when there fewer than 1% of the sample. These are 45 combined with 47, 17 combined with 19 and 15, and 21 combined with 23.

unemployment duration across occupations.[6] Again this dispersion is countercyclical.

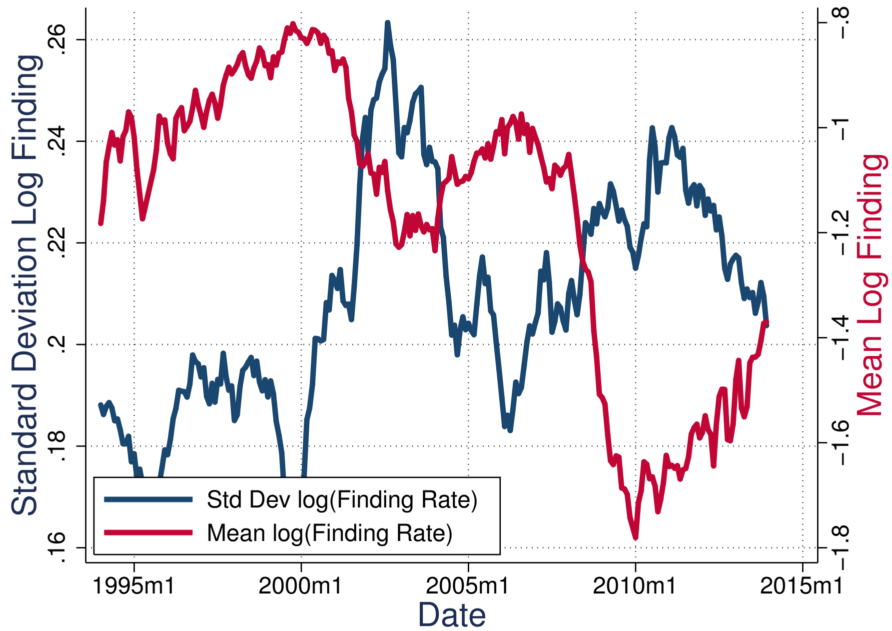

Figure 4: The variation in finding rate across occupations

How much of the tail in unemployment duration can be explained purely by occupation-level heterogeneity? Figure 6 takes the mean finding rate within a given occupation and then computes the implied unemployment duration. Notice that it is indeed higher than if there were a uniform finding rate, but only barely. In this exercise, we have taken occupations as fixed entities, with the same workers leaving and entering employment. Because these occupations have different flow rates, some longer than others, it does introduce some heterogeneity in the finding rate and hence a tail to unemployment duration that is longer than the exponential implied a uniform finding rate. However, the effect is fairly small, and duration only rises moderately. With occupation-specific finding rates, the average duration is 3.32 months instead of 3.18 months with a perfectly equal finding rate across occupations. Both are significantly lower than the time-series average of 4.78 months in our sample of the data.

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

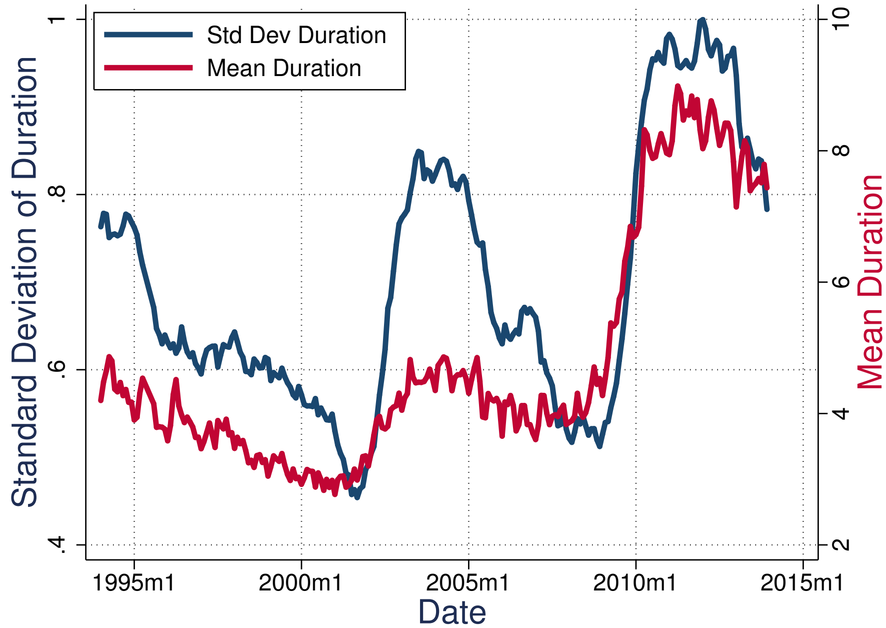

Figure 5: The variation in unemployment duration across occupations

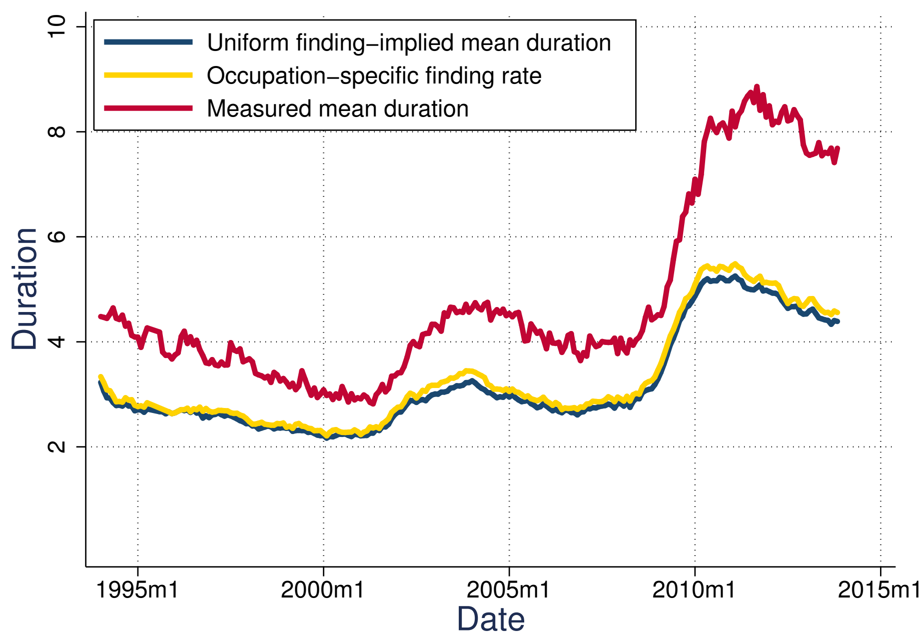

Figure 6: The average duration implied by differences in finding rate at the occupation level.

Taking the mean finding rate within an occupation misses much of the heterogeneity. But, it still incorporates the downward slope because some occupations have, on average, slower finding rates. In this exercise we are, however, ignoring the endogenous reallocation across occupations that also may occur among unemployed workers. On average, about half of unemployed workers will match in another occupation. A worker who switches occupations, however, does not simply find a job at the rate of this other occupation, instead they generally have longer unemployment durations (see e.g. Pilossoph (2014) and Carrillo-Tudela and Visschers (2014)).

There is evidently additional heterogeneity among searchers that leads some to find a job in their same occupations and others to switch. Those who switch lose their occupation-specific human capital, and potentially face an entirely different finding rate compared to others who originate in the same occupation. Moreover, switchers do not necessarily go to the occupation with the highest finding rate, depending on a number of factors. In what follows, the model will try to incorporate this additional heterogeneity and the ability to switch occupations.

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 tied directly to data.

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 $\omega_{\ell d}$ will be determined by the data, potentially greater than or less than 1. Labor in an occupation is aggregated linearly in 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 comovements. Aggregate productivity follows a simple AR(1) process and the vector of factors follows a VAR. Productivity shocks are described by the system

$$Z_t = \rho_Z Z_{t-1} + \epsilon_t \tag{1}$$

$$z_{d,t} = \lambda_{f,d}\mathbf{f}_t + \lambda_{Z,d} Z_t + \rho_z z_{d,t-1} + (1-\rho_z) + \zeta_{d,t} \tag{2}$$

$$\mathbf{f}_t = \Gamma \mathbf{f}_{t-1} + \eta_t \tag{3}$$

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 comovements among some occupations beyond their relationship through the aggregate cycle. The form is meant to be parsimonious because there are J occupations and so an estimate for $E[\zeta\zeta']$ is infeasible.

Several additional shocks govern the effects of worker transitions. Workers become experienced at rate τ, meaning they provide productivity $\omega_{dd} = 1$. If experienced workers separate, they can keep that experience if they match with the same occupation. Those who are separated while inexperienced become unattached to any occupation, and upon matching with 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 modelling 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.

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 the first period of unemployment, 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. This preference shock is crucial as it drives gross flows that are far larger than net flows. Its implication is that workers with observably similar characteristics, namely the same occupational work history, have different job finding rates. This is a very similar form to the additive random utility in Boskin (1974) or Miller (1984).

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 40% replacement of their former salary.[7] 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 is similar to Fredriksson and Holmlund (2001) and saves unemployment duration from becoming a state. I denote δ here as a parameter, but in reality it actually fluctuates over the cycle (see Mitman and Rabinovich (2015)).

Note that unemployed workers do not enjoy leisure utility, which Hagedorn and Manovskii

[7] I give 40% 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.

(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.

4.3 Search and Market Structure

Each (ℓ, d, e) has its own labor search market $m \in M$, where $M = \{(\ell, d, e) \ : \ \ell \in \{0, \ldots, J\}, d \in \{1, \ldots, J\}, e \in \{0, 1\}\}$. 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 $p\left(\theta^m\right)$ 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. $M_{\ell*, e^*}$ is the set of markets for which this worker can apply, $M_{\ell^*, e^*} = \{(\ell, d, e) | \ell = \ell^*, d \in \{1, \ldots, J\}, e = e^*\}$. 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 $\sum_{m \in M_{\hat{\ell}, \hat{e}}} g^m p(\theta^m)$.

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.

4.4 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

4.5 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 newly unemployed workers enjoying benefits is $U_b(\ell, 0, 1, x, \mathcal{Z})$ and once they choose a search direction is becomes $U_b(\ell, 0, 1, x, \mathcal{Z})$. The value function for those with expired benefits is $U_b(\ell, d, 0, x, \mathcal{Z})$, and $U_w(\ell, d, 1, x, \mathcal{Z})$ for employed workers. These composite value functions are slightly more convenient, but the overall value function, in which I include employment status as an explicit state is $U(k, \ell, d, 1, x, \mathcal{Z})$.

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

$$U(k, \ell, d, e, x, \mathcal{Z}) = \mathbb{I}_{k=w} U_w(\ell, d, 1, x, \mathcal{Z}) + \mathbb{I}_{k=b} U_b(\ell, 0, e, x, \mathcal{Z}). \tag{4}$$

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

$$
\begin{aligned}
U_w(\ell,d,1,x,\mathcal{Z}) = &\max_{\bar{\xi}^{\ell d1}}(1-\tau\mathbb{I}_{\ell\neq d})\left[\left(1-s^{\ell d1}\right)\left(\int_{\bar{\xi}^{\ell d1}}^{0} w^{\ell d1}(\xi,0)+\xi dH(\xi)+\beta EU(\ell,d,1,x',\mathcal{Z}')\right)\right.\\
&\left.+\, s^{\ell d1}U_b(0,0,1,x,\mathcal{Z})\right]+\tau\mathbb{I}_{\ell\neq d}U_w(d,d,e,x,\mathcal{Z}) \qquad (5)\\
U_b(\ell,0,e,x,\mathcal{Z}) = &\int_\psi \max_{\{g^m\}_{m\in M_{\ell,e}}} \sum_{m\in M_{\ell,e}}^{J} p(\theta^m)g^m\left(w^{\ell d1}(0,\psi_m)+\psi_m+\beta E[U(\ell,d\in m,1,x',\mathcal{Z}')]\right)\\
&+\left(1-\sum_m p(\theta^m)g^m\right)\left(b(\ell,0,e)+\beta E[U_b(\ell,b,e',x',\mathcal{Z}')]\right)dF(\{\psi_m\}) \qquad (6)\\
U_b(\ell,d,e,x,\mathcal{Z}) = &\int_\psi p(\theta^{\ell de})\left(w^{\ell de}(0,\psi_m)+\psi_m+\beta E[U(\ell,d,1,x',\mathcal{Z}')]\right)\\
&+\left(1-p(\theta^{\ell de})\right)\left(b(\ell,d,e)+\beta E[U_b(\ell,d,e',x',\mathcal{Z}')]\right)dF(\{\psi_m\}). \qquad (7)
\end{aligned}
$$

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

$$
\begin{aligned}
U_w(d,d,1,x,\mathcal{Z}) = &\max_{\bar{\xi}^{\ell d1}}\left(1-s^{dd1}\right)\left(\int_{\bar{\xi}^{dd1}}^{0} w^{dd1}(\xi,0)+\xi dH(\xi)+\beta EU(d,d,1,x',\mathcal{Z}')\right)\\
&+\, s^{dd1}U(d,0,1,x,\mathcal{Z}). \qquad (8)
\end{aligned}
$$

All workers take as given equilibrium conditions on θ^m, w^m, which are described in Appendix B. Expectations for the law of motion of $\mathcal{X}(x) = x'$ are consistent with the actual law of motion. The policy function that determines the search direction g^m is a function of the state at that stage $\ell, 0, 1, x, \mathcal{Z}$ and also the J-dimensional vector of shocks, ψ. Note also that s^m already includes the integration over ξ, that is $s^m = H(\bar{\xi}^m)$. For notational convenience in the definition of equilibrium, it will be helpful to denote the number of applicants in a labor market a^m. The definition is just accounting, given in Appendix B.

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 change nonmonotonically over a worker's tenure. Wages 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 his own occupation. In subsequent periods, if he matched in a new occupation the outside option is lower because he will enter unemployment as an untenured, unskilled worker. He is penalized both by a lower unemployment benefit and a lower continuation value that reflects lost human capital. To avoid this problem, I would have to define the problem of workers who are displaced with labor history ℓ, d. When they then matched with a new occupation, j, I would need to define the value of a worker with labor history ℓ, d, j. 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 Firms

The representative multiworker firm produces using 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 earnings, $\{w^m\}$ 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 associated with changing the scale of the firm.

$$\Pi(\mathbf{L}, x, \mathcal{Z}) = \max_{\{v^m\}_{m\in M}} \sum_{\ell=0,d=1}^{J} \left[\omega_{\ell d} z_d L'_{\ell d} - \left(L'_{\ell d} - \sum_{e\in\{0,1\}} v^{\ell de} q(\theta^{\ell de}) \right) \int_{\bar{\xi}^{\ell d}}^{0} w^{\ell d1}(\xi, 0) dh(\xi) \right]$$
$$- \sum_{m} \int_{\psi} v^m q(\theta^m) w^m(0, \psi) d\tilde{f}(\psi) + \kappa v^m + \beta E[\Pi(\mathbf{L}', x', \mathcal{Z}')] \quad (9)$$

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

$$L'_{\ell d} = (1-\tau)(1-\hat{s}^{\ell d1}) L_{\ell d} + q(\theta^{\ell d0}) v^{\ell d0} + q(\theta^{\ell d1}) v^{\ell d1} \quad (10)$$

$$L'_{dd} = (1-\hat{s}^{dd1}) L_{dd} + q(\theta^{dd1}) v^{dd1} + q(\theta^{dd0}) v^{dd0} \quad (11)$$

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

I define the equilibrium conditions in the Appendix, Section B.

4.7 Cyclical Dynamics of Finding Rate Heterogeneity

A crucial result of the model is that the dispersion of unemployment duration is countercyclical. To deliver this, there must be a mechanism such that occupation switching takes longer in recession. In a recession, the occupations that are most cyclically sensitive have the worst productivity and the workers attached to this occupation are unemployed for an extended period whether they find a job in their own or other occupations.

Two channels 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 $b + \beta E[U_b(\cdot)]$.[8] 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.

5 Quantitative Strategy

Crucially, workers in the model have some attachment to their prior occupation and there are aggregate forces that push them away. I 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 finding rates.

5.1 Occupation Specific Shocks

To estimate the shock process described by Equations 1-3, we will impute it from fluctuations in the occupation specific job finding rate. This addresses two potential pitfalls: Worker composition affects occupation's productivity and, in models such as this one, observed volatility in productivity will not deliver realistic volatility in job finding unless I constrain the movement of wages. Instead, by selecting the productivity process to match the finding rate process, we are able to match the volatility in job finding at both aggregate- and occupation-level while still having equilibrium fluctuations in wages.

First, even with an observation on output in an occupation, we could not directly estimate the process for productivity in an occupation because we would also need workers per

[8] This statement holds so long as $\omega_{\ell d} < 1$. Though we let the data dictate these values, it will prove to generally be the case.

occupation and their working history. Even with output within an occupation d and the number of workers there, $\sum_\ell L'_{\ell d}$, this would be insufficient because a worker's productivity depends on his work history as well, because $\omega_{\ell d} < 1$ if $\ell \neq d$. Thus, the observed productivity per worker in an occupation d is $z_d \frac{\sum_\ell \omega_{\ell d} L'_{\ell d}}{\sum_\ell L'_{\ell d}}$. Where we could potentially have output data we do not have worker composition and where we have workforce composition we do not have proper output data.

The second reason to use match finding rate fluctuations is to address a common problem with search models. Search models with flexible wages generally yield insufficient variation in finding rates for productivity shocks that correspond to those in the data, a finding often associated with Shimer (2005). Even if we could properly measure output per worker in an occupation and quality-adjust depending on that worker's history, we would still have the problem that search models generally do not deliver much variation in the finding rate. At the aggregate level, this point is shown very clearly in Shimer (2005). Solutions generally involve constraints on wages, either as explicitly sticky wages or with a calibration that yields a very small surplus in which wages may vary as in Hagedorn and Manovskii (2008). Neither of these are good solutions in this context because wage variation in response to productivity shocks is necessary. In the model, wages are a signal to workers to chose which occupation in which to search in this period. And in my empirical strategy, I will use wage variation to infer the the human capital loss associated with different types of occupation switches.

To back out productivity from finding rates, will still require a numerical solution to the model, as the model's implied finding rate incorporates endogenous decisions on where to search. To see this, we can directly construct the model-implied finding rate in destination occupation d, which we will call p_d, as

$$p_d = \frac{\sum_{\ell,e} a^{\ell de} p(\theta^{\ell de})}{\sum_{\ell,e} a^{\ell de}}$$

Though $\theta^{\ell de}$ only depends on z_d, its expectations and technology parameters, $a^{\ell de}$ is a function of the whole vector of $\mathcal{Z}$. Introducing some notation, let Λ be the set of parameters that govern the process for z_d, $\Lambda = \{\rho_Z, \rho_z, \{\lambda_{f,d}\}_{d=1}^J, \{\lambda_{Z,d}\}_{d=1}^J, \Gamma, var(\epsilon), var(\zeta), var(\eta)\}$ and $\{p_{d,t}^{data}\}$ and $\{p_{d,t}^{model}\}$ be the sequence of occupation specific finding rates. The task is to choose Λ minimize the distance between the distribution of $\{p_{d,t}^{data}\}$ and $\{p_{d,t}^{model}(\Lambda)\}$.

I will use data from the CPS that was described in Section 3. To choose Λ in the model, I will find the panel of shocks $\{z_{d,t}\}$ such that the endogenous finding rates in the model exactly match the panel of finding rates in the data $\{p_{d,t}^{data}\}$. I then estimate Λ on this sequence and solve the model with these governing expectations and repeat the procedure until convergence. The result is that the process of finding rates in the model exactly matches the process for finding rates in the data because the panels themselves match observation-wise.

In Appendix C I provide estimates for finding rates as well as the counter-parts that govern the model's productivity process. The difference between finding rates in model and data are very small. The total difference over the whole panel of finding rates in the data and model is 0.28458. This is only an average disparity of 0.004743 per parameter of Λ.

5.2 The 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 O*NET. Then, I assume $\omega_{\ell d}$ is a 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 1996-2012.

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 several aspects of the requirements 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 this data both from surveys and specialized analysts. I use the final Analyst Database, also titled O*NET 4.0 and released in 2002.

To characterize occupations, I use the "skills" measure because, in my model, workers are attached to their occupation by human capital learned on the job. O*NET's skills are the closest representation to this sort of on-the-job learning. 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 to three principal components. This is because I will eventually have to match coefficients on each of these skill dimensions and also because there is a high correlation across dimensions meaning that some of it can be considered redundant. The first three components explain about 80% of the variation between all of the O*NET occupations. I also will have to map the O*NET occupations, which are classified more finely than my two-digit SOC occupations. 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 Survey of Income and Program Participation data

The data on wage histories that I will use comes from the Survey of Income and Program Participation (SIPP). This survey consists of panels from 1996-2000, 2001-2003, 2004-2007 and 2008-2012. Each starts with a new set of respondents and then follows them for the duration of the survey. Data is presented at a monthly frequency though it is collected in four month intervals and the intervening months are filled by recall. Importantly, it has evidence on monthly earnings, employment status, occupation and occupational tenure. Occupational tenure is given as the answer to a retrospective question rather than calculated directly from the survey because the panels are relatively short.

The dataset is the same as set up in Eubanks and Wiczer (Forthcoming). To clean reporting errors in the earnings data, we take a light touch. In the sample there are about 4.6% whose earnings are less than would be implied working for 10 hours per week at minimum wage. We leave these in the data sample to preserve potentially meaningful variation, though they can be excluded without significantly affecting the estimates. At the top of the distribution, we use the top-code adjustment developed by the Center for Economic and Policy Research (2014). There are fewer than 1% of the sample with earnings that have large spikes within a month that quickly revert. We drop these observations, which we define as a monthly change exceeding 200% but which reverts such that the two-month change is less than 10%.

Before using earnings in the data, I first regress them on sex, age, age squared and college education. I use residual earnings because all of these complicating factors in wage determination do not exist in my model, but are quite significant in the data.[9] We use earnings rather than wages because earnings are reported directly in the SIPP by most of the respondents and then wages are computed by combining hours data, which is often measured with error.

The SIPP has several advantages compared to other common datasets that might be used for labor market histories. In particular, it contains many unemployment transitions and occupation switches, 24,413 with usable data. It also has data on earnings immediately after arriving in the new occupation instead of other datasets such as the PSID in which I can only observe earnings annually. With monthly rather than annual data, time aggregation

[9]Results from these first stage regressions are available upon request.

does not obscure the difference in earnings between new and experienced hires.

5.2.3 Estimating occupational skills and productivity

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

$$\omega_{\ell d} = e^{\sum_{i=1}^{3} \beta_i (k_{i,d} - k_{i,\ell}) + \beta_0}$$

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

The form is partly motivated by considering a linear approximation of log earnings of new hires around the average wage of the experienced worker, $\bar{w}^{dd}$. Using the bargained earnings in Equation 22, this yields a convenient linear equation relating $\{\beta\}$ to the wage gap between experienced and inexperienced workers:

$$\log\left(\frac{w^{\ell d}}{\bar{w}^{dd}}\right) \propto \sum_i \beta_i \left(k_{i,d} - k_{i,\ell}\right) + \beta_0 - \mu\psi_d \tag{12}$$

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 regress the experience premium on the skill difference in both SIPP and model-generated data.

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 from 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 $\bar{w}^{dd}$ because most of the experienced workers are not near the threshold.

As seen in Table 1, the auxiliary model is quite precisely estimated and skill differences are strong predictors of differences in the wage premium for experience. To interpret these coefficients is difficult as each dimension is really a composite of skills. However, it is showing that a beginner suffers a larger wage gap in an occupation using a skill much more intensively than his old occupation.

	Coefficient	T-statistic	Model	Structural
Skill 1	-0.331	-6.88	-0.35	-0.276
Skill 2	-0.294	-6.20	-0.32	-0.557
Skill 3	-0.186	-4.33	-0.19	-0.275
Constant	-0.289	-23.98	-0.35	-0.207

Table 1: 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.

5.3 Calibration of the Economy

I will calibrate the rest of the parameters in the model using data from the CPS and SIPP. The setup of these was described previously.

5.3.1 Random utility, matching function and separations

In this paper, I model random deviations in search behavior through Gumbel-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. Even with standard distributional assumptions on the shock itself, the worker faces random utility with a payoff that is endogenous and dependent on model tightness.

The choice is *almost* like the standard additive random utility model like Boskin (1974) except that the random component is scaled by $p(\theta^{\ell de})$. 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 $\left(p(\theta^{\ell de})(1-\mu)\right)^2 \sigma_\psi^2$. As is standard, preference shocks are still Gumbel-distributed but are now now heteroskedastic. Therefore, the direction policy is chosen according to a heteroskedastic logit, as described in Bhat (1995). Unfortunately, policies for $\{g^{\ell de}\}$ no longer have the closed form description they had in the homoskedastic case.

The choice of matching function is nontrivial in this model because shocks may be quite dispersed across occupations and so some may experience very large realizations. To avoid realizations with a matching probability of 1, I use the form $m(u,v) = \frac{vu}{(v^\phi + u^\phi)^{1/\phi}}$. This is still homogeneous 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," this elasticity is too high.[10] I have both ϕ_0, ϕ_1 to match both level and elasticity of the job finding rate with the matching function to

$$m(u,v) = \phi_0 \frac{uv}{(v^{\phi_1} + u^{\phi_1})^{1/\phi_1}}, \quad p(\theta) = \phi_0 \frac{\theta}{(1+\theta^{\phi_1})^{1/\phi_1}} \tag{13}$$

For the distribution of disutility to govern shocks, I use a negative-exponential as the

[10] As 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.

tail with a mass point at a disutility of zero. Most workers experience no disutility from work, but some fraction can, if they would stay in the match, experience a large amount of disutility.

$$Pr[\xi < x] = \begin{cases} \lambda_0 e^{\lambda_1 x} & \text{if } x < 0 \\ 1 & \text{if } x = 0 \end{cases} \tag{14}$$

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.[11]

5.3.2 Choosing targets

The parameters then that we have to match are $\phi_0, \phi_1, \lambda_0, \lambda_1, \tau$, and σ_ψ. There are three other parameters I set exogenously. $\kappa = 0.26$ corresponds to the measured posting cost, as used in Hagedorn and Manovskii (2008). The firms' share, μ, is not easily observable and the literature uses a wide range of values. The reasons behind the values used in Shimer (2005) and Hagedorn and Manovskii (2008) are not applicable here, so I choose a relatively conservative 0.5 in line with den Haan et al. (2000). The results are not particularly sensitive to values anywhere between 0.5 and 0.8, though at very high levels other moments of the data become difficult to match. The unemployment replacement rate is 40% of the prior working wage, which is approximately average across US states and the same as chosen by Shimer (2005).

τ, which governs the speed at which workers become experienced, and the wage helps the model match the average returns to occupational tenure. I use the value from Kambourov and Manovskii (2009), who estimate the five year return using PSID data and correcting for the endogeneity of tenure by instrumenting as in Altonji and Shakotko (1987).[12]

[11]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 productivity shocks affect inexperienced workers by too much and few workers are able to become experienced.

[12]Kambourov and Manovskii (2009) use data from 1968-1993, but the results are not much different when

σ_ψ, the standard deviation of occupation specific preference shocks, determines the rate of switching occupations when unemployed. For a very low σ_ψ, searchers go to the single occupation with the largest return and a high σ_ψ increases the spread of $g^{\ell d}$ regardless of $\{z_d\}$. As described in Section 5.3.1, the standard deviation of preference shocks, as perceived by searchers, actually depends on the finding rate in that occupation so σ_ψ just acts as a baseline.

I choose the average separation rate so that the average unemployment rate is 5.7%, the average in my CPS sample. 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, though either could do. This is to emphasize the role of separations in determining the distribution of unemployed people.

Table 2 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.

Target	Model	Data	Source
Change Probability	0.46	0.46	SIPP 1996-2012
Job Finding Rate	0.31	0.33	CPS, 1994-2013
Returns to tenure	0.006	0.006	Kambourov and Manovskii (2009) (monthly rate)
Match Elasticity	0.37	0.48	Barnichon and Figura (2015)
Unemployment Rate	0.064	0.057	CPS, 1994-2013
sd(Separation)	0.01	0.01	CPS, 1994-2013
Posting Cost	0.26		Hagedorn and Manovskii (2008)
Firms' share	0.5		den Haan et al. (2000)
Replacement Rate	0.4		Shimer (2005)

Table 2: Summary of calibration targets

extended to 2006, although changes in the survey make this later data somewhat more suspect.

6 Results

The model provides insight into how unemployment duration evolves over the business cycle and in steady state. My model, with occupation-level heterogeneity, delivers a longer duration overall and more of the rise during the Great Recession than a model with a uniform finding rate. The model generates countercyclical dispersion of duration and unemployment rates across occupations. Summarily, this is because with heterogeneous searchers in a downturn some are affected more greatly than others. This fall in their finding rate pulls out the whole duration distribution. This is akin to the logic in Figures 3 and 1, 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 is lower than others and these workers' long duration pulls out the distribution of duration enough to generate levels of long-term unemployment that are closer to the data. As discussed in the Section 4.7, in a recession, those who are attached to an occupation with a slow finding rate are doubly in trouble: if they stay, their occupation is not hiring much and if they try to switch they have an especially slow finding rate because unskilled workers are particularly unprofitable in a recession. This exacerbates countercyclical dispersion. With this heterogeneity in recession, the workers from affected occupations constitute a greater part of the tail of the duration distribution.

In this section, to discuss the baseline results I will first present statistics that represent how duration looks within my model relative to the data, then I study its cross-sectional dispersion and finally I focus particularly on the Great Recession. In each subsection, I show data generation by the model in which it is fitting the exact sequence of finding rates from the data. In the baseline results, I extended unemployment benefits from 6 to 24 months from June 2008 through the end of my sample. I explore the impact of leaving these fixed in the final subsection.

6.1 Unemployment Duration

Unemployment duration was not a target of the calibration or estimation. It is instead an endogenous outcome given fluctuations in the finding rate and a set of skills that are imperfectly transferable. To the extent that the model generates unemployment duration beyond a benchmark implied by a uniform finding rate, which it does, this is a result of these additional features.

As we have tried to emphasize earlier, realistic levels of unemployment duration require that workers find jobs at different rates. With a uniform finding rate that matches the estimate from the average flows out of unemployment, duration and long-term unemployment will be too low. In Figure 7, I show that the finding rate declines with duration in my model. This is purely a result of composition: Those left still searching at longer durations have low finding rates in my model.

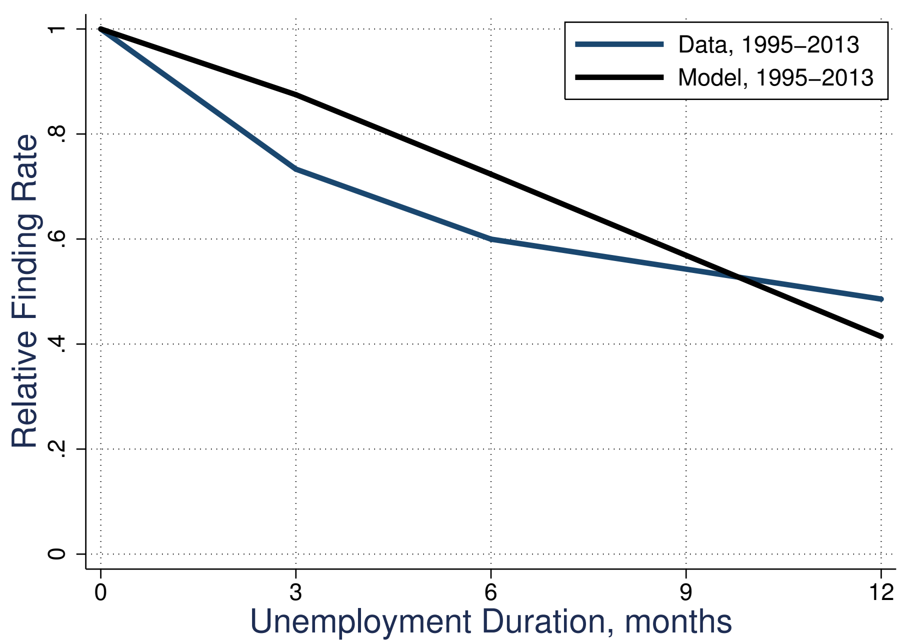

Figure 7: The change in the finding rate with duration in the data and in the model. Both model and data are normalized to the level at zero months.

Notice that the model introduces a lot of heterogeneity across individuals in the finding rate. But Figure 6 suggested that the finding rates were not sufficiently different across occupations to generate a long tail on the duration distribution. The difference is that the model has an additional source of heterogeneity based on the preference shocks, ψ. Not all workers from the same occupation will search in the same place and crucially, a large fraction will choose to move occupations. This squares the number of finding rates in the data and hence introduces much more heterogeneity into the set of types.

Essentially, there are workers from each occupation fanning out into many other occupations. When they do this, they are faced with other occupation-specific shocks and they are imperfectly transmitting their human capital. The result is that many of these switchers have very slow finding rates and very long durations.

We see this in Table 3 and Table 4 which analyze the time-series of average duration and of long-term unemployment in my model, the data and that which is implied by the average finding rate.

	Mean	Standard Deviation	Skewness
Data	4.85	1.84	1.01
Model	3.87	0.98	0.89
Uniform	3.24	0.90	0.93

Table 3: Duration: Summary statistics on the time-series, comparing the data, model output and that implied by a uniform finding rate.

	Mean	Standard Deviation	Skewness
Data	0.237	0.113	0.887
Model	0.164	0.087	0.840
Uniform	0.112	0.080	0.986

Table 4: Long-term unemployment: Summary statistics on the time-series, comparing the data, model output and that implied by a uniform finding rate.

The average duration in my model is 79.8% of the duration in the data, whereas with a uniform finding rate it implies only 66.8% of the average duration. The difference is more stark when focusing directly on the tail of the duration distribution, when looking at the long-term unemployment. There the uniform finding rate delivers only 47.2% of the long-term unemployment in the data. On the other hand, my model with occupational heterogeneity delivers 69.5%. In both statistics, the uniform finding rate understates the length of the tail of unemployment duration. Introducing occupation-level heterogeneity increases it significantly.

Both unemployment duration and the rate of long-term unemployment also vary more in the data than with a uniform finding rate. Introducing the elements of this model increases the amount of variation over time. If heterogeneity in finding rates yields the higher mean duration, that also seems to vary over time. Thus a single uniform finding rate not only misses on average, it also understates the variation. In this model, counter-cyclical dispersion in finding rates makes the duration more volatile. Partly this is mechanical, because there is simply a wider dispersion of shocks that are hitting the economy. However, these shocks are behaving in a particular way: they generate counter-cyclical dispersion, which is a feature I explore in the next Section 6.2.

The model, however, seems to understate the peaks in duration and long-term unemployment. We can see this in the fact that its skewness is not as great as the data or even that which is predicted by a single finding rate. Section 6.3 will discuss the reasons for this further: essentially separations are too volatile, particularly too high during the Great Recession which pulls down the unemployment duration at just the time that the very bad shock is reducing the finding rate.

Figures 8 and 9 plot the time-series data for long-term unemployment and average unemployment duration in the model, data and a benchmark using a uniform finding rate. The figures generally reconfirm what the summary statistics above showed: the model creates additional unemployment duration beyond that which is predicted by the change in the average finding rate. Visually, the effect is more prominent as a change in the level of unemployment duration than a large increase in its cyclicality. As noted, the volatility of duration does increase but it is actually less skewed than would be predicted by a single finding rate.

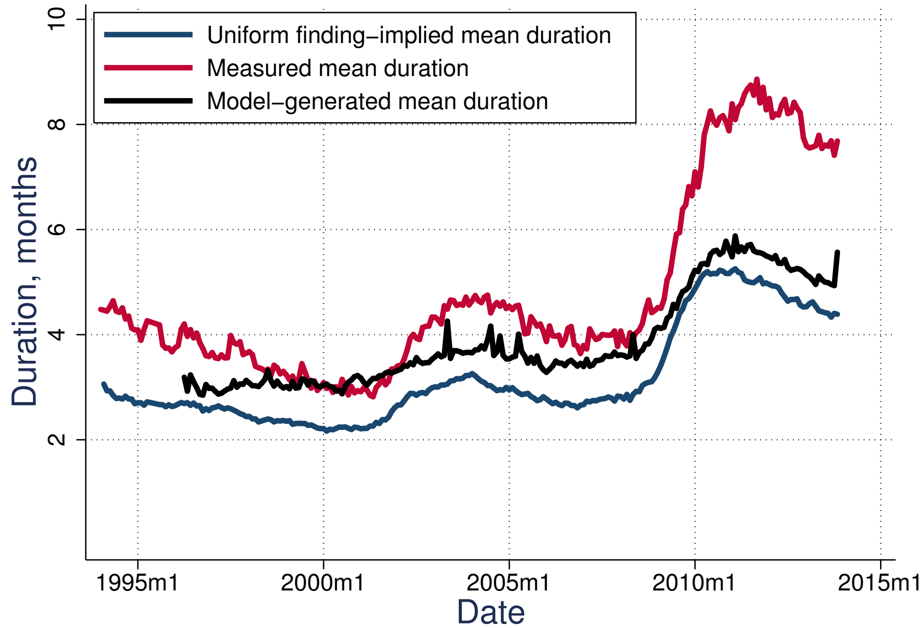

Figure 8: Mean duration as generated by the model with the data and that implied by a uniform finding rate for comparison.

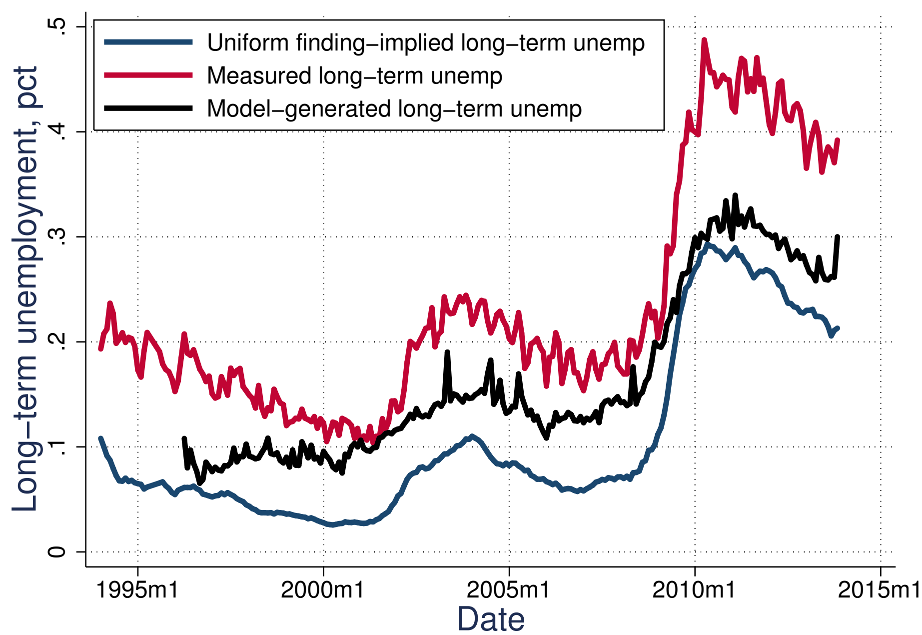

Figure 9: Long-term Unemployment as generated by the model with the data and that implied by a uniform finding rate for comparison.

6.2 Counter-Cyclical Dispersion

Turning to the cross-section over the business cycle, I will discuss the strong countercyclical dispersion in the model. As I have discussed, the literature has pointed out two mechanisms by which the economy can generate countercyclical variance in outcomes. Rather than the variance shocks of Lilien (1982), I generalized Abraham and Katz (1986) so that occupations differ in their cyclical sensitivity. In this framework, the variance in productivity increases symmetrically for positive and negative aggregate shocks to Z. Cross-occupation dispersion in unemployment and duration can be countercyclical, however, 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 can increase instantaneously in response to a negative shock.

Table 5 shows the countercyclical dispersion in log finding rate, unemployment rate and duration. For each statistic, I take the panel of occupations and compute the cross-sectional standard deviation and the average in each period using population weights. I then compute the correlation between these time-series, the level and cross-sectional dispersion in each period. For reference, I also provide the time-series average of the cross-sectional standard deviation.

Notice a subtle point here in the definition of finding rate. Though I targeted the finding rate of destination occupations, this is not the same as the finding rate in the origin occupation. Though the finding rates across destination occupations are the same between model and data, the finding rates conditioning on prior occupation are expected to be different.

For each series, the dispersion in the model is even more cyclical 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. Despite these potential additional considerations, the model is very close to the data in the relationship between the dispersion across occupations' finding rate and the cycle.

	Model	Data
corr(sd($u_{\ell,t}$),u_t)	0.97	0.86
average sd($u_{\ell,t}$)	0.015	0.028
corr(sd($\log p_{\ell,t}$),$\log p_t$)	-0.865	-0.529
average sd($\log p_{\ell,t}$)	0.071	0.212
corr(sd($d_{\ell,t}$),d_t)	0.940	0.862
average sd($d_{\ell,t}$)	0.539	0.695

Table 5: Business cycle properties of dispersion in the model, $p_{\ell,t}$ is the finding rate of those coming from occupation ℓ, $u_{\ell,t}$ is the unemployment rate and $d_{\ell,t}$ is the average duration

Indeed, for the model to be reproducing unemployment duration, it must be creating dispersion in unemployment duration both within and between occupations. On the other hand, a search model with uniform matching would not have any difference across occupations in any of these statistics nor would they be cyclical. The model delivers dispersion across occupations, both at the level and its cyclical volatility. Table 5 compares these to the data and the magnitudes are all realistic. However, the model is also creating the dispersion within occupations because, conditional on their prior occupation, some workers will choose to search in different occupation. This additional heterogeneity proves to be quite important in creating realistic movements of unemployment duration and long-term unemployment.

6.3 The Great Recession

I will now focus on the experience in the Great Recession, in which average unemployment duration spiked as did long-term unemployment. It rose and then did not begin to fall even years into the recovery. This was also a time when occupation's finding rate and duration diverged, as first presented in Figures 4 and 5. To the model, this is a time when the average finding rate falls, which implies there must also be a fall in the aggregate productivity Z_t and the increased dispersion in finding rates should imply that idiosyncratic productivity, $z_{d,t}$ also diverged. As we have seen in Figures 8 and 9, the model generates a rise in unemployment duration but not one that is commensurate with that seen in the data.

The crucial aspect to the divergence between unemployment duration in the model and in the data is the separation rate is higher in the model during this period. The very large productivity shock that affects the observed finding rate also implies a large increase in separations among the currently employed. A high separation rate implies a dip in the unemployment duration, even though the decline in the finding rate would tend to increase duration.

Table 6 describes the differences between the separation rate in the data and the model.[13] Though I did not directly target the average separation rate, by targeting the unemployment rate and the finding rate I ensured the model's average would be very close. I also targeted the cross-sectional dispersion in the separation rate, but not the time-series variation. The time-series variation is instead determined both by the parameterization of my disutility function, $H(\cdot)$ and volatility of the productivity shocks, but both of these are determined by other targets. We see that the variation is an order of magnitude higher in my model than in the data and this is especially borne out by a very high value during the Great Recession when productivity was low.

	Mean	Standard Deviation	Average 2008-2010
Model	0.023	0.012	0.0339
Data	0.023	0.002	0.0231

Table 6: Statistics of the separation rate in the model and data.

This said about the separation rate, the model does capture a good deal of the rise in unemployment duration in the Great Recession and the countercyclical rise in dispersion across occupation. Table 7 compares the rise in unemployment duration between the data and that predicted by the model. The model predicts 72.9% of the long-term unemployment in the data, whereas if we only used the fall in the finding rate to imply a distribution of unemployment duration, we would only see 57.8% of the long-term unemployment. A

[13]This is the separation rate correcting for time aggregation using the method described in Elsby et al. (2009).

similar statement can be said about average duration. If we use a uniform finding rate to imply a mean duration in the pool of unemployment, it is only 4.06 months, or 69.2% of the true duration in the Great Recession. The model, on the other hand predicts average unemployment duration that is 78.2% of the mean duration observed in the data.

Long-term Unemployment, Model	0.234
Long-term Unemployment, Uniform	0.186
Long-term Unemployment, Data	0.321
Mean Duration, Model	4.59
Mean Duration, Uniform	4.06
Mean Duration, Data	5.87

Table 7: 2008-2010: Long-term unemployment and duration in the model and data.

In the model, the rise in unemployment duration is considerably beyond that which would be predicted by a uniform finding rate because of the composition effect, which creates a negative correlation between finding rate and duration as we have seen throughout. Figure 10 plots the finding rate as a function of unemployment duration. In the Great Recession the finding rate shifts down at all durations in both model and data. Though model does not replicate the exact shape of duration dependence in the data, the heterogeneity implied by occupational skills does create a significant downward slope in the finding rate. This downward slope is present in the 2008-2010 subperiod and allows the model to deliver unemployment duration more similar to that in the data.

Table 8 looks at the cross-occupation dispersion in the model relative to the data during the Great Recession. The model picks up only about half of the standard deviation of the finding rate during the Great Recession. However, it matches about 80% of the standard deviation across occupation of the duration.

Just as the discrepancy between the separation rate in the model and data contributes to the model's inability to match the large rise in unemployment duration, the lack of dispersion in finding rates in this period is also a major factor. In the model, the standard deviation in this period is higher than its average by about 2.9 percentage points. However, the model was never capturing sufficiently the dispersion in the finding rates across occupations. On

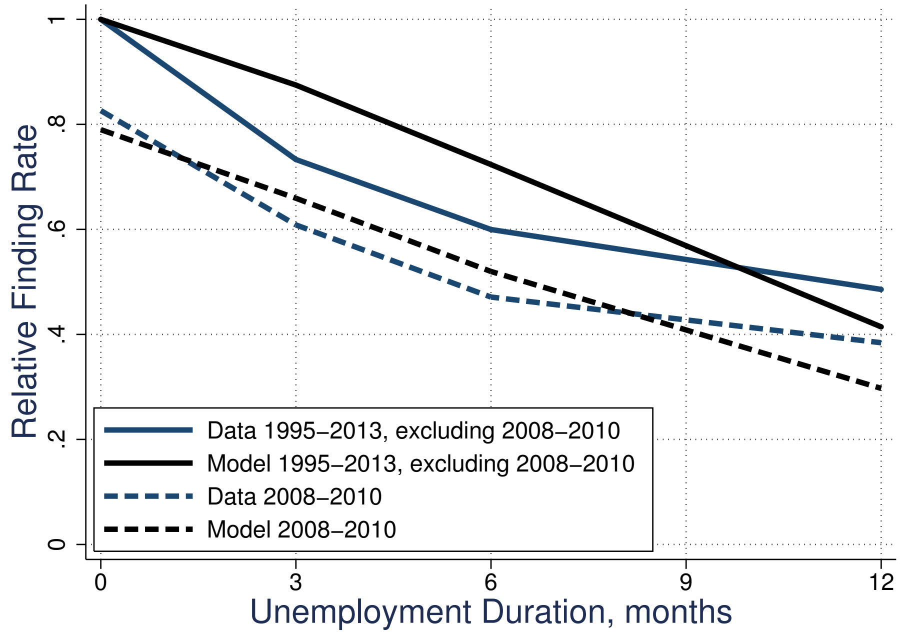

Figure 10: The change in the finding rate with duration, comparing the 2008-2010 period with the rest of the sample. Normalized to the level at zero months during 1995-2013 excluding 2008-2013.

average during the 1995-2013 sample the model's standard deviation was only 33% of that in the data. The rise in dispersion was actually larger in the model than the data, where it only rises by 1.2 percentage points over its sample mean.

Std Dev Duration, Model	0.691
Std Dev Duration, Data	0.861
Std Dev log(finding), Model	0.100
Std Dev log(finding), Data	0.224

Table 8: 2008-2010: Cross-occupation standard deviation of unemployment duration and log finding rate in the model and data.

6.4 Unemployment extensions

Within the model, we can perform a policy experiment by leaving set the unemployment benefits. Within this model, unemployment benefits play an important part in unemployment duration because when they expire, they reduce the outside option of the worker and make his finding rate higher. By extending the benefits, these workers find jobs more slowly. The effect of the policy is especially strong during a recession because more workers are at long-durations when their benefits would have expired and their finding rate would have risen.

To perform this policy experiment, I keep all parameters fixed, which were estimated under the policy with extended benefits as actually happened. However, a new actual sequence of productivity shocks will be implied to fit the data on finding rates that were observed. This way, I maintain the feature that the average finding rate is still the same between model and data and isolate the effect on heterogeneity driving unemployment duration. Just as workers did not anticipate the extension of unemployment benefits, this experiment does not affect their beliefs either.

I find that without the extension of unemployment benefits, the average duration of unemployment is 0.31 months shorter during 2008-2010. The longer-term unemployment rate is also 2.29 percentage points lower. I plot the series in Figures 11 and 12. Though the period from 2008-2010 does show a difference between the two policies, it is clear in the figures that the larger differences appear in subsequent years.

There is, however, an important channel we are probably missing, which is that workers whose unemployment benefits expire might optimally choose to change their search strategy. In my framework, workers set their search direction at the beginning of the period and then stay.

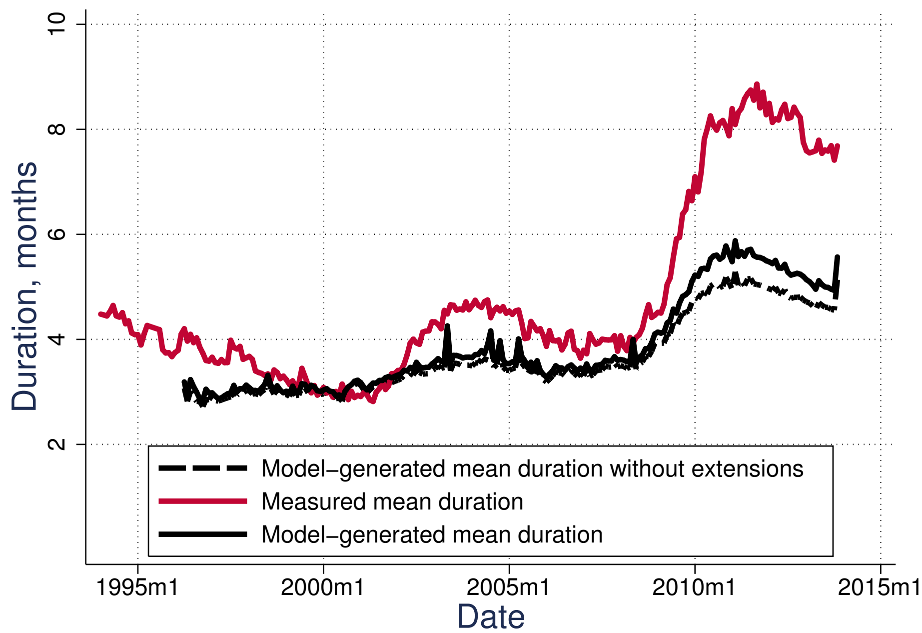

Figure 11: Mean duration comparing the results without unemployment benefit extensions with the data and the baseline model

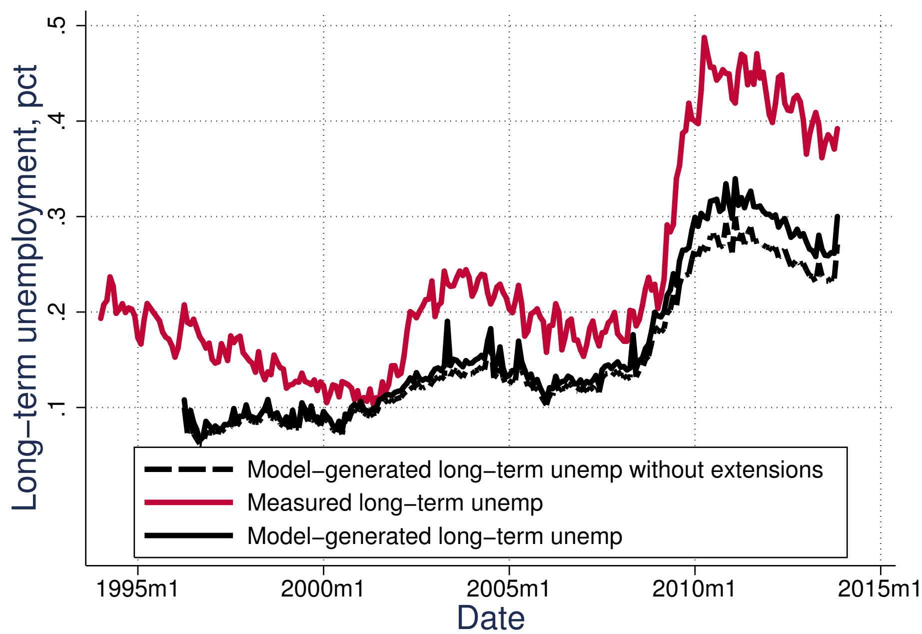

Figure 12: The rate of long-term unemployment comparing the results without unemployment benefit extensions with the data and the baseline model

7 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 in their occupation-specific skills. With discipline from data on the wage premium to occupational experience and the occupation-specific finding rate, it quantified the link between skills, business-cycle fluctuations and unemployment duration. In the model, workers faced a crucial choice of whether to search within their own occupation or try to switch while firms chose whether to post a vacancy for an experienced or inexperienced worker. For workers from hard-hit occupations, their search was particularly slow because their own occupation was unlikely to hire and switching always implies a longer unemployment duration.

I began with a data exercise showing that to understand the level and cyclical variation of unemployment duration, one must allow for differences across the unemployed in their job finding rates are required for. In any model in which the job finding rate is the same for all workers, 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 based on individual earnings dynamics. However, conditioning on prior occupation alone it does not provide enough heterogeneity to replicate unemployment duration series.

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, they might significantly increase their unemployment duration. Generally, this model delivered realistic business cycle fluctuations in unemployment duration and its countercyclical dispersion across occupations. In particular, the model's incidence of long-term unemployment was much closer to the data than a uniform finding rate would predict. Finally, I applied the model to the Great Recession. The model captured a great deal of the duration, about 80%, and also about 80% of the cross-occupation dispersion in duration.

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. Particularly, one could assess a plan like the Obama administration's "Bridge to Work" policy, that encouraged employers to hire and train long-term unemployed. This policy allowed employers can hire and train long-term unemployed 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. A. Smith, and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81, 1395–1454.

Alvarez, F. and R. Shimer (2011): "Search and Rest Unemployment," *Econometrica*, 79, 75–122.

Barnichon, R. and A. Figura (2015): "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics*, 7, 222–49.

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.

Carrillo-Tudela, C. and L. Visschers (2014): "Unemployment and Endogenous Reallocation over the Business Cycle," ESE Discussion Papers 241, Edinburgh School of Economics, University of Edinburgh.

Center for Economic and Policy Research (2014): "SIPP Uniform Extracts, Version 2.1.7," Washington, DC.

Clark, K. B. and L. H. Summers (1979): "Labor Market Dynamics and Unemployemnt: A Reconsideration," *Brookings Papers on Economic Activity*, 10, 13–72.

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.

Elsby, M. W. L., R. Michaels, and G. Solon (2009): "The Ins and Outs of Cyclical Unemployment," *American Economic Journal: Macroeconomics*, 1, 84–110.

Eubanks, J. and D. Wiczer (Forthcoming): "The Earnings Distribution of Occupational Switchers," *Review*.

Flood, S., M. King, S. Ruggles, and J. R. Warren (2015): "Integrated Public Use Microdata Series, CPS: Version 4.0." .

Fredriksson, P. and B. Holmlund (2001): "Optimal Unemployment Insurance in Search Equilibrium," *Journal of Labor Economics*, 19, 370–99.

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.

Kambourov, G. and I. Manovskii (2009): "Occupational Specificity Of Human Capital," *International Economic Review*, 50, 63–115.

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.

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.

Mitman, K. and S. Rabinovich (2015): "Optimal unemployment insurance in an equilibrium business-cycle model," *Journal of Monetary Economics*, 71, 99–118.

Nakajima, M. (2012): "A quantitative analysis of unemployment benefit extensions," *Journal of Monetary Economics*, –.

Pilossoph, L. (2014): "Sectoral Shocks and Move Unemployment," Manuscript, Federal Reserve Bank of New York.

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.

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.

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 Occupations

11	Management Occupations
13	Business and Financial Operations Occupations
16	*Combines*: Computer and Mathematical Occupations (15), Architecture and Engineering (17) and Life, Physical, and Social Science (19)
22	*Combines*: Community and Social Services Occupations (21) and Legal Occupations (23)
25	Education, Training, and Library Occupations
27	Arts, Design, Entertainment, Sports, and Media Occupations
29	Healthcare Practitioners and Technical Occupations
31	Healthcare Support Occupations
33	Protective Service Occupations
35	Food Preparation and Serving Related Occupations
37	Building and Grounds Cleaning and Maintenance Occupations
39	Personal Care and Service Occupations
41	Sales and Related Occupations
43	Office and Administrative Support Occupations
46	*Combines*: 45 Farming, Fishing, and Forestry Occupations (45) and Construction and Extraction Occupations (47)
49	Installation, Maintenance, and Repair Occupations
51	Production Occupations
53	Transportation and Material Moving Occupations
55	*Omitted*: Military Specific Occupations

B 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

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

- Policy functions:
  - $\{g^m\}_{m=0}^{2J(J+2)}$ where $g^m : \mathbb{R}_+^J \times \{0, J\}^2 \times \{0,1\} \times [0,1]^{2(J+1)^2} \times \mathbb{R}^4 \to [0,1]$
  - $\{\bar{\xi}^m\}_{m=0}^{J(J+2)}$ where $\bar{\xi}^m : \{0, J\}^2 \times \{0,1\} \times [0,1]^{2(J+1)^2} \times \mathbb{R}^4 \to \mathbb{R}_-$
  - $\{v^m\}_{m=0}^{2J(J+2)}$ where $v^m : [0,1]^{J(J+1)} \times [0,1]^{2(J+1)^2} \times \mathbb{R}^4 \to \mathbb{R}_+$
- Wages, $\{w^m\}_{m=0}^{(2J(J+2)}$, where $w^m : \mathbb{R}_- \times \mathbb{R}_+ \times [0,1]^{2(J+1)^2} \times \mathbb{R}^4 \to [0,1]$
- Market tightness, $\{\theta^m\}_{m=0}^{2J(J+2)}$, where $\theta^m : [0,1]^{2(J+1)^2} \times \mathbb{R}^4 \to [0,1]$

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

$$a^m = \begin{cases} \bar{g}^m \left(x_{\ell d1} + \bar{s}^{\ell\ell 1}(x_{\ell\ell 1} + \tau \sum_{j \neq \ell} x_{j\ell 1}) \right) & \ell > 0,\ e = 1 \\ \bar{g}^m \left(x_{0d1} + (1-\tau) \sum_{j=0}^{J} \sum_{k \neq j} \bar{s}^{jk1} x_{jk1} \right) & \ell = 0, e = 1 \\ \bar{g}^m \left(x_{\ell d0} \right) & e = 0 \end{cases}$$

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, $\bar{g}^m(\cdot) = \int_{\psi_i} g^m(\psi_i; x, \mathcal{Z}) dF(\psi)$

$$x'_{dd1} = (1-\bar{s}^{dd1})\left(x_{dd1} + \tau \sum_{\ell=0,\ell\neq d}^{J} x_{\ell d1}\right) + p(\theta^{dd1})a^{add1} + p(\theta^{dd0})a^{add0} \tag{15}$$

$$x'_{\ell d1} = (1-\bar{s}^{dd1})(1-\tau)x_{\ell d1} + p(\theta^{\ell d1})a^{\ell d1} + p(\theta^{\ell d0})a^{\ell d0} \tag{16}$$

$$x'_{\ell 01} = \left(1 - \sum_d \bar{g}^{\ell d1} p(\theta^{\ell d1})\right)\left((1-\delta)x_{\ell 01} + \bar{s}^{\ell\ell 1}\left(x_{\ell\ell 1} + \tau \sum_{j\neq l} x_{jl1}\right)\right) \tag{17}$$

$$x'_{001} = \left(1 - \sum_d \bar{g}^{0d1} p(\theta^{0d1})\right)\left((1-\delta)x_{001} + (1-\tau)\sum_{d=1}^{J}\sum_{l=0,l\neq d}^{J} \bar{s}^{\ell d1} x_{\ell d1}\right) \tag{18}$$

$$x'_{\ell 00} = \left(1 - \sum_d \bar{g}^{\ell d0} p(\theta^{\ell d0})\right) x_{\ell 00} + \delta\left(1 - \sum_d \bar{g}^{\ell d1} p(\theta^{\ell d1})\right) x_{\ell 01} \tag{19}$$

- 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:

$$\kappa = q(\theta^{\ell de})\left(\omega_{\ell d} z_d - \int_\psi w^{\ell de}(0,\psi) d\tilde{f}(\psi) + \beta E \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}')\right) \tag{20}$$

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

- Separations are mutual and the cutoff satisfies

$$\omega_{\ell d} z_d + \bar{\xi}^{\ell d1} + \beta E[U_w(\ell, d, 1, x', \mathcal{Z}') + \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}')] = \begin{cases} U_w(l, 0, 1, x, \mathcal{Z}) & \ell = d \\ U_b(0, 0, 1, x, \mathcal{Z}) & \ell \neq d \end{cases} \tag{21}$$

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

$$\begin{aligned} w^{\ell de}(\xi,\psi) = {} & (1-\mu)\left(\omega_{\ell d} z_d + \beta E \Pi_{\ell d}(\{x'_{kj1}\}_{k,j}, x', \mathcal{Z}')\right) \\ & - \mu\left(\xi + \psi - b + \beta E U_w(\ell, d, 1, x', \mathcal{Z}') - \beta E U_b(\ell, 0, e, x', \mathcal{Z}')\right) \end{aligned} \tag{22}$$

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

C Estimates of the stochastic process

Γ, data finding rate		Γ, model productivity	
0.0191	-0.128	0.429	-0.028
-0.0055	0.004	-0.106	0.181

Table 9: Γ for the finding rate process in the data and the productivity process in the model. These govern the persistence of the unobserved factors

Occupation	$\lambda_1, \lambda_2, \lambda_Z$, data finding rate			$\lambda_1, \lambda_2, \lambda_Z$, model productivity		
11	-0.326	0.073	0.857	-0.650	-1.468	0.354
13	-0.174	0.224	0.914	2.268	3.168	0.791
16	-0.186	0.407	0.844	4.941	-5.060	0.147
22	0.686	0.378	0.921	-3.069	-1.303	0.398
25	0.384	0.286	1.088	-2.107	-2.860	0.449
27	-0.190	0.602	1.039	-0.695	-1.575	0.379
29	0.274	0.180	1.031	-0.688	-2.011	0.434
31	0.374	-0.152	0.863	-1.770	-0.378	0.396
33	0.231	0.803	0.987	-0.752	0.219	0.338
35	0.115	0.042	1.044	-1.736	-1.606	0.512
37	0.258	0.108	0.699	-1.053	-1.205	0.351
39	-0.058	-0.320	0.874	-1.769	-1.712	0.425
41	0.018	-0.123	0.963	-1.375	-1.754	0.518
43	-0.129	-0.092	1.013	-1.291	-1.836	0.531
46	0.241	-0.081	1.042	-1.711	-1.962	0.625
49	0.018	-0.193	0.995	-1.819	-4.106	0.690
51	-0.174	-0.055	0.914	-1.417	-2.656	0.601
53	-0.210	0.016	1.004	-1.332	-1.508	0.490

Table 10: Λ for the finding rate process in the data and the shock process in the model. The first two Λ columns correspond to the loadings on unobserved factors and the third is the loading on the aggregate finding rate/productivity

	data finding rate	model productivity
ρ_z	0.049	0.170
ρ_Z	0.890	0.652
σ_ϵ	0.031	0.309

Table 11: Parameters of the process for the finding rate in the data and the productivity shock process in the model.

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-Meade 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 finding rate and that provides a useful first guess for the actual underlying productivity process. Call this $\tilde{\mathcal{Z}}^{data}$. I use this process as a first guess and repeatedly simulate the model and solve for the productivity that replicates the finding rate. I describe the iterative procedure below:

1. From finding rate data, estimate the monthly process $\tilde{\mathcal{Z}}^{data}$
2. Draw M realized histories from $\tilde{\mathcal{Z}}^{data}$
3. Solve and simulate the model around process $\tilde{\mathcal{Z}}^{data}$ with realized history m.
4. Solve for $\{z_{d,t}^1\}$ by inverting $\{\theta^{\ell de}\}$.

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

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