IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15017

Worker-Firm Screening and the Business Cycle

Jake Bradley

JANUARY 2022

DISCUSSION PAPER SERIES

IZA DP No. 15017

Worker-Firm Screening and the Business Cycle

Jake Bradley
University of Nottingham and IZA

JANUARY 2022

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Worker-Firm Screening and the Business Cycle*

There has been a substantial body of work modeling the co-movement of employment, vacancies, and output over the business cycle. This paper builds on this literature, and informed by empirical investigation, models worker and firm search and hiring behavior in a manner consistent with recent micro-evidence. Consistent with empirical findings, for a given vacancy, a firm receives many applicants, and chooses their preferred candidate amongst the set. Similarly, workers in both unemployment and employment, can evaluate many open vacancies simultaneously and choose to which they make an application. Business cycles are propagated through turbulence in the economy. Structural parameters of the model are estimated on U.S. data, targeting aggregate time series. The model can generate large volatility in unemployment, vacancies, and worker flows across jobs and employment state. Further, it provides a theoretical mechanism for the shift in the Beveridge curve after the 2008 recession - a phenomenon often referred to as the jobless recovery. That is, persistently low employment after the recession, despite output per worker and vacancies having returned to pre-crisis levels.

JEL Classification: J63, J64
Keywords: Beveridge curve, screening, jobless recovery

Corresponding author:
Jake Bradley
University of Nottingham
School of Economics
University Park
Nottingham, NG7 2RD
United Kingdom
E-mail: jake.bradley@nottingham.ac.uk

* I would like like to thank seminar audiences at: the University of Edinburgh; the University of Nottingham; the Millennium Institute on Market Imperfections and Public Policy, Diego Portales University, Chile; Trinity College Dublin; the University of Essex; the University of Geneva; and the University of Sheffield.

1 Introduction

This paper starts from the premise that the main impediment in finding a job is not a *black-box* matching function that approximates other frictions, as is typically assumed in the literature. Rather, a job seeker fails in their search for a job because a more suitable candidate is hired instead. A quantitative model is set up in which unemployed and employed workers can observe many job openings and are discerning in where to apply. On the other side of the market, firms receive many applicants and screen all candidates, hiring the most suitable. The business cycle is propagated through turbulence in the economy. Turbulence is defined as variation in the probability that worker-firm matches are hit by idiosyncratic productivity shocks. The model is estimated on U.S. data and is able to reconcile a number of empirical regularities regarding labor market dynamics and the behavior of job vacancies over the cycle.

Firm hiring and worker search practices are informed by recent empirical evidence and embedded into an otherwise standard model of the labor market. In particular, the paper documents that a given job opening is likely to receive many applications and that screening, of some form, is ubiquitous. Indeed, using administrative Swiss data, Blatter et al. (2012) show that the majority of a firm's hiring cost is devoted to worker screening – approximately two thirds. On the worker side, using survey evidence provided by the New York Fed. Faberman et al. (2017) show that the unemployed devote much more time to job search activities and apply to jobs more frequently than their employed counterparts. However, despite this, outcomes for the employed, as measured by offered wages, are better than those received by unemployed workers. A model with these features incorporated is able to generate levels and dynamics of worker transitions across employment states and jobs that are consistent with the data. Further, the model simultaneously matches these flows with a downward sloping Beveridge curve that exhibits an outward shift following the great recession.

Informed by the evidence provided by Faberman et al. (2017), the model is set up so that the frequency to which workers actively engage in job search varies by employment status and the number of open vacancies visible to them is governed by a Poisson distribution that again depends on whether they are employed or unemployed. The estimated parameters suggest that although the unemployed search for jobs more frequently, conditional on looking, the employed are presented with a greater number of opportunities, on average. The interpretation of this search behavior is that the unemployed are endowed with more time to engage in job search and hence apply more frequently. However, being out of work disengages them from the labor market and as such are '*out of the loop*' of new potential job opportunities. Hence, when the employed apply for work they will in expectation sample more job offers. In a typical search model, even with on-the-job search, the pool of unemployed provide a greater opportunity for open vacancies than the employed as they search with a greater efficiency. If one models both the in-and-out flow from unemployment, this can lead to an upward sloping Beveridge curve; as in downturns, unemployment increases which leads to a rise in the return of posting a vacancy. This is not necessarily the case in the mechanism discussed here however as the searching employed are more discerning in that they choose the best match from a larger set of offers than the unemployed. Given that firms too are discerning in who they match with, it is not clear whether the frequent searching unemployed or the more discerning employed offer greater opportunities for firms. The estimated model thus generates a downward sloping Beveridge curve with pro-cyclical flows out of unemployment and counter-cyclical flows into unemployment.

When deciding whether to post vacancies, as in a model absent of on-the-job search, firms care about the level of turbulence in the economy and the unemployment rate. Since approximately half of all hires come from the pool of employed workers (Nagypal, 2008) the distribution of matches amongst the employed will also matter. Well matched workers are less likely to leave their current employers than poorly matched ones, therefore this distributional object is also consequential to the aggregate level of vacancies. Unlike the unemployment rate which adjusts relatively quickly, the distribution of match quality is a slow moving object. In the model, it is this channel that generates the high vacancy rate following the great recession, the so-called *jobless recovery.* The dramatic downturn sees a quick drop in unemployment and vacancies, but the distribution of matches takes longer to adjust. As the worker-firm matches deteriorate there is more incentive to post vacancies and the Beveridge curve shifts outwards as employers hire poorly matched workers leaving the unemployment rate relatively stable.

It seems quite counterintuitive that the rate the employed find alternative employers is so consequential for the behavior of unemployment over the business cycle. The job ladder is not only responsible for the slow employment recovery post-crisis but also for the movement in output and wages over the cycle. Since neither aggregate productivity nor wages are hit directly by aggregate shocks, their movement is determined endogenously by workers moving up and down rungs of the job ladder.

2 Related literature

There have been numerous models featuring on-the-job search attempting to understand the dynamics of the labor market.[1] Of these, the most related paper is Engbom (2021). As in this paper, workers search on the job and firms screen workers. Contrary to this paper, screening is modeled as producing an informative signal as to whether the match is good or bad. Aggregate fluctuations are driven by shocks to the separation rate and the model is able to reconcile large movements in employment with smaller movements in labor productivity, the so-called Shimer puzzle, Shimer (2005b). Engbom (2021) like here stresses the importance of the search behavior of the employed in determining the cyclicality of unemployment. Eeckhout and Lindenlaub (2019) make the same counterintuitive conjecture. In Eeckhout and Lindenlaub (2019), a worker hired from employment is more productive than one from the pool of unemployed. Vacancy posting firms thus benefit from on-the-job search and employed searchers benefit from vacancy posting, leading to multiple equilibria. As will be discussed, a similar strategic complementarity exists here which also generates a downward sloping Beveridge curve and a shifting of said curve following an extreme downturn.

[1] Menzio and Shi (2010, 2011) develop models in which firms post wages and workers direct their search to a specific submarket. From a random search perspective, Robin (2011) and Lise and Robin (2017) explore labor market dynamics where wages are bargained over ex post of meeting and firms can make counteroffers in an attempt to retain the worker. Finally, Coles and Mortensen (2016) and Moscarini and Postel-Vinay (2013) consider a random search environment in which wages are posted ex ante, and in the case of Moscarini and Postel-Vinay (2013) are committed to.

Worker search behavior. Crucial to the mechanism of this paper is the differential search behavior across employment state. The unemployed search more frequently but each time are exposed tom on average, fewer opportunities than the employed. This leads to the employed seemingly sampling from a better wage distribution. The mechanism is supported empirically by recent survey data provided by the New York Fed, see Faberman et al. (2017) and goes against the predictions of a canonical sequential random search model with on-the-job-search.

There are many classes of models in which the employed sample jobs from a different distribution from the unemployed. In a directed search model with on-the-job-search as in Delacroix and Shi (2006) the jobs workers direct their search to is a function only of their current wage. Conditional on a worker's place in the job ladder the sampling distribution of jobs will be degenerate at the targeted rung. Variance in the sampling distribution can be introduced by assuming jobs are targeted with some imprecision, Lentz et al. (2019). In a random search environment, Bradley and Gottfries (2021) develop a model in which workers accrue opportunities and apply to their set of opportunities at differing rates, depending on their employment state. The sampling of jobs thus depends on their entire labor market history and they show large quantitative differences in these distributions between the employed and unemployed.

Firm screening. Typically, in the labor search literature firms are modeled as passive. They post vacancies and negotiate or post wages. Hiring of workers in this model is slightly more involved. After posting a vacancy, the firm screens all those who apply at a cost, and then chooses their preferred candidate to fill the position. Evidence from firm hiring data suggests that screening constitutes a large share of the frictional impediment in finding workers, see Blatter et al. (2012), Davis and Samaniego de la Parra (2017) and Andrews et al. (2017).

In the model, workers can observe many vacancies and ultimately apply to one. This results in an urn-ball matching function whereby firms receive a Poisson number of applications.[2] The

[2]This follows Moen (1999), Burdett et al. (2001), Shi (2001, 2002, 2006), and Shimer (2005a) who all model such a co-ordination friction in the labor market.

framework can be extended to consider workers making multiple applications as in Albrecht et al. (2006), Kircher (2009) and Galenianos and Kircher (2009). Multiple applications within a period introduces an additional co-ordination friction, in that a firm must internalize the probability that a worker receives an offer from an alternative firm. Finally, Villena-Roldan (2012) and Wolthoff (2018) go a step further still, and introduce an intensive margin into the screening process. The firm chooses a subset of applicants to screen, something that this paper abstracts from. Instead, in this paper screening is treated as a fixed cost for the firm and therefore all or no candidates are screened.

The aforementioned papers are solved in a stationary environment. However, firm screening has also been modeled in the presence of aggregate shocks. Following the work of Pissarides (2009) and Silva and Toledo (2009), Christiano et al. (2016) and Moscarini and Postel-Vinay (2018) introduce screening costs as a labor turnover cost incurred ex post of meeting a worker. Like here, while vacancy costs are paid upfront, a screening cost is paid conditional on meeting a worker. A high unemployment rate increases the probability of a firm hiring. However, since some of the hiring costs are paid once a firm has met a worker the models are able to generate a larger volatility in unemployment for the same aggregate shock process. Quantitatively, it is shown that such a cost helps to mitigate the lack of employment volatility as pointed out by Shimer (2005b).

3 Empirical Motivation

This section documents a number of empirical regularities regarding the hiring behavior of firms and search behavior of workers. These facts will inform the modeling choices made in section 4 and be used to evaluate the quantitative performance of the model in section 6. Put succinctly, the facts relating to firm hiring are:

1. Screening is ubiquitous, to some degree a firm will screen every candidate it ends up hiring.

2. A typical job opening can expect to receive many job applicants.

3. Screening costs constitute the lion's share of total hiring cost, approximately two thirds.

Regarding the search behavior of workers:

4. Relative to the employed, the unemployed apply more frequently to jobs. But, receive a similar number of job offers.

5. Conditional on receiving a job offer, the employed and unemployed appear to sample wages from different distributions.

Firm side. The Multi-City Study of Urban Inequality (MCSUI) coupled with recent evidence from Blatter et al. (2012) are exploited to document a number of empirical phenomena to inform the modeling choices made in the paper.

The MCSUI employer survey is a survey of establishments in Atlanta, Boston, Detroit and Los Angeles, conducted over the telephone between the springs of 1992 and 1995. Employers were questioned about the specific firm's characteristics, hiring and recruiting practices in general, as well as specific information pertaining to the recruitment of the last employee hired. Firms were asked about their hiring protocol and whether they '*always*', '*sometimes*' or '*never*' implemented a variety of practices. The 15 practices questioned about have been split, in a fairly *ad hoc* way, into either a screening or recruitment motive. Screening is defined as an attempt to select the best applicant from all who applied. The recruitment motive is defined as maximizing the total number of applicants. Creating binary indicators based on whether or not a firm '*always*' implements a specific practice and running a principal components analysis, to some extent, supports the *ad hoc* categorization. Factor loadings are depicted in Figure A.1 of Appendix A.1. The principal component analysis is not only supportive of the *a priori* categorization but also is indicative of two separate classes of hiring practices.

The left panel of Table 1 shows the responses of firms to the questions deemed to be related to screening practices. Every responding firm, of which there were 3,510, stated they '*always*' check at least one of the five criteria listed. This finding is interpreted as screening, although the specific practice varies, is universally carried out. It is easy to see why. Respondents were asked about the number of applications received for the last position they hired. The distribution exhibits a very

Table 1: Screening and Applications

(a) Screening Intensity

	Always	Sometimes	Never
do you check?			
written application	82%	12%	7%
personal interview	87%	11%	2%
references	74%	22%	4%
education/training	32%	39%	29%
criminal record	32%	17%	52%
any of the above	100%	0%	0%

(b) Applicants and Interviews

	# applicants	# interviewed
Median	10	5
Unweighted mean	17.7	6.2
Weighted mean	18.4	6.4

Notes: Data are taken from the MCSUI. Panel (b) refers to the last hire the firm made. For these statistics the top 5% of the distribution is trimmed. Weights are constructed as one divided by the number of months between hiring and the survey interview date. The idea is that this should proxy for the overall hiring intensity of a firm.

long left tail. Even after trimming the top five percent of the distribution a job opening can expect to get almost twenty applicants. Of these applicants, they are likely to interview more than six.

Whilst the MCSUI is informative about the screening and recruiting practices of firms, it tells us little about the resources attributed to these practices. To get a better understanding of this, work by Blatter et al. (2012) is relied upon. The upper panel of Table 2 is computed from Blatter et al. (2012). Hiring costs have been differentiated into three broad margins: costs of job posting; costs of interviewing; and costs in hiring external advisors. As before, earmarking costs towards screening and hiring is quite *ad hoc*. It appears uncontroversial to attribute job posting costs to recruitment and interview costs to screening. However, it is not clear whether the hiring of external advisors was to aid in increasing the number of applicants to interview or to determine the best candidate from the set of applicants. In order to approximate the relative costs of screening and recruitment the bottom panel of Table 2 computes two extremes. One in which external advisors only help with recruitment and one in which they only aid with screening. The results suggest that in aggregate, screening accounts for somewhere between 57% and 69% of total hiring costs.

Worker side. Data from the Survey of Consumer Expectations (SCE) are used and in particular the job search supplement. The main data follows workers for a year and the supplement is annual and conducted in October each year. All data is based on 2013 to 2017, inclusive and is restricted to

Table 2: Recruiting Costs

	Number of employees				Total
	1-9	10-49	50-99	100+	
Costs measured relative to mean monthly wages					
Costs for job postings $\overline{v}$	11%	25%	37%	51%	18%
Costs for interview per applicant $\overline{c_a}$	5%	8%	9%	12%	6%
No. of interviewed applicants $\overline{J}$	4.7	4.9	4.7	5.4	4.8
Personnel costs for interviews $\overline{Jc_a}$	25%	41%	44%	70%	32%
Costs for external advisors $\overline{e}$	4%	9%	18%	24%	7%
Total cost: $\overline{Jc_a} + \overline{v} + \overline{e}$	41%	75%	99%	145%	56%
Implied share of hiring cost attributed to screening					
LB: $\overline{Jc_a}/(\overline{Jc_a} + \overline{v} + \overline{e})$	62%	54%	45%	48%	57%
UB: $(\overline{Jc_a} + \overline{e})/(\overline{Jc_a} + \overline{v} + \overline{e})$	72%	67%	63%	65%	69%

Data are taken from Blatter et al. (2012) and based on administrative Swiss data between 2000 and 2004, inclusive. Costs are benchmarked to mean wages using OECD data on mean annual wages over the same window, implying a mean monthly wage of 6,299.7 CHF (Swiss Francs).

workers between the ages of 18 and 65. For a comprehensive discussion of the data, see Faberman et al. (2017). The data contains information not typically observable in labor market data. In particular the survey contains information regarding a worker's failed job search and on job offers rejected by a worker. Here the focus will be on the number of applications made and job offers received in the last four weeks, as well as the wage of the best offer received in the last six months. Note, the best offer is determined by the respondent and may not coincide with the highest wage offered.

Table 3 presents differences in applications made, offers received and wages of best offers by employment status. Using the main sample of the SCE a variety of worker level controls have also been included. The results of these regressions represent the supporting argument for facts four and five listed previously. First, note that a typical employed individual, will be 60 percentage points less likely to send out a job application in a given month than an unemployed individual. These differences persist even after conditioning on the employed looking for work. However differences in job offers are much lower across employment state. With the employed only three percentage

Table 3: Worker Search Behavior

	Applied: last four weeks		**Received offer: last four weeks**		**Log wage of best offer**	
Employed (all)	-0.599^{***} (0.030)		-0.032^{**} (0.017)			0.220 (0.175)
Employed (looking)		-0.103^{***} (0.025)		0.017 (0.018)		
Employed (not looking)		-0.749^{***} (0.023)		-0.047^{***} (0.017)		
In work					0.221^{***} (0.056)	
N_e	3067	3067	3067	3067	738	292
N_u	186	186	186	186	238	25
R^2	0.130	0.517	0.012	0.026	0.191	0.186

The table shows the coefficients for the variables of interest. Covariates not shown include: age; age squared; and dummies for gender; white ethnicity; hispanic; marital status; calendar year; four regions; and eight educational dummies. Unemployment follows the BLS definition and the employed are those in paid work, regardless of hours, omitting the self-employed. Employed (looking) is defined as the employed who have done '*anything in the last four weeks to look for new work*'. In work is based on the whether at the time of offer a worker was in paid work, it does not discriminate between self- and paid-employment. In this case the control group is being not in work (including the economic inactive) in all other cases the control group are the unemployed. The total sample size is $N_e + N_u$, where N_e is the size of the treatment group (usually the employed) and N_u the control. Only workers aged between 18 and 65 are included and stars represent conventional levels of statistical significance.

points less likely to be the recipient of a job offer in a given month. Once one conditions on the employed to be looking for work they become, albeit insignificantly, more likely to receive a job offer.

The final columns report coefficients from a Mincer wage regression on the '*best*' offer received by a worker in the last six months. The '*best*' offer is self-reported and need not be the highest wage offered. The employment status is at the time the offer was received. Respondents were asked if they were in or out of work at the time. However, it is not clear if this means paid or self-employment if in work, or unemployed or inactive, if out. In some instances, if the offer falls in a window where respondents were questioned in the main sample, a concurrent employment status consistent with the BLS definition can be inferred. The employment status is then cross-referenced with the response regarding work. In both cases there appears to be a large premium for the employed, who on average sample wages more than 20% higher than their unemployed counterparts.

4 Model

Environment. Time is discrete and the economy is populated by risk neutral, infinitely lived firms and workers. Both firms and workers are ex ante homogeneous. Workers can be either employed or unemployed and have access to a search technology that allows them to evaluate multiple job openings simultaneously. The number of vacancies they assess is governed by a Poisson distribution, with parameter λ_u for the unemployed and λ_e for the employed. The match productivity z, is specific to the worker-firm match, and drawn from a distribution with probability density function $g(z)$. After evaluating all matches, workers have the ability to apply to just one per period. The productivity level z is transitory, and at rate γ_t, a new productivity z' is redrawn from the sampling distribution $g(z'|z)$. The duration of the match quality is governed by γ_t and is thought of as a measure of turbulence in the economy. The parameter γ_t evolves stochastically over time and is the source of aggregate fluctuations in the economy.

Firms post vacancies at a cost κ and a vacancy can be matched with multiple workers. If workers apply to a vacancy a firm then screens its applicants at a cost σ. A firm then chooses which candidate to hire. The same screening cost is paid irrespective of the number of applicants.[3] After picking the ideal candidate the firm and worker negotiate wages according to a Nash bargaining protocol without the threat of permanent separation. Wages are then re-negotiated every period.

Timing. The within period timing is as follows:

1. New γ_t is realised.

[3] One can imagine a more sophisticated screening cost as a function of the number of applicants. The simple setup is motivated by inspection of Table 2 where recruiting costs are compared across the firm size distribution. While the number of candidates interviewed varies little with firm size. There is large variation in the personnel costs for interviews. The interpretation of this result is that the number of candidates screened plays little role in the total cost of screening.

2. Wages in all jobs are re-bargained over.

3. Produce output and exchange wages according to contractual agreement.

4. A productivity shock hits the match with probability γ_t. The new productivity z' is drawn from the sampling distribution $g(z'|z)$, which is a function of the current match productivity z.

   - Existing matches may separate.

5. Firms post vacancies at a cost, κ.

6. Worker's actively engage in search with probability α_u for the unemployed and α_e for the employed.

7. The active searching unemployed and employed (who have not been laid off) workers sample vacancies at random.

   - The number of vacancies they look at depends on a draw from a Poisson distribution, with rate parameter λ_s, $s \in \{u, e\}$.
   - For each vacancy they draw a match specific productivity z, from an exponential distribution, with parameter λ_z.

8. Workers choose which single vacancy to apply for out of the set available to them.

9. Firms who receive applications, screen said applicants at a cost σ.

10. Screening firms offer their first choice candidate a job and reject all others.

Evolution of γ. At the start of a period t, γ_t, the probability an existing match is hit with a productivity shock is revealed and is public information. It is assumed that γ_t follows a first order autoregressive process of the form

$$\gamma_{t+1} = \gamma_0 + \rho\gamma_t + \epsilon_t \quad \text{where, } \epsilon_t \sim N(0, \sigma_\gamma^2).$$

For brevity, the Markov transition probability of γ evolving to γ' is denoted as $\pi(\gamma', \gamma)$. After a shock hits, the new productivity z' is drawn from the same distribution as it is initially sampled from, bounded by the current match productivity level z. It is assumed that $g(z)$ follows an exponential distribution, with parameter λ_z, and therefore,

$$g(z'|z) = \begin{cases} \frac{\lambda_z \exp(-\lambda_z z')}{1-\exp(-\lambda_z z)} & \text{for, } 0 \leq z' \leq z \\ 0 & \text{otherwise.} \end{cases}$$

It is the evolution of γ that propagates movements in the business cycle. Prolonged periods of high γ are associated with shorter match duration and lower future levels in productivity, this will consequently correspond with economic downturns in this setting. Similarly, the booms are the result of prolonged periods of low γ. One can either interpret the changing γ as turbulence in the economy. An alternative interpretation is since all shocks reduce the match productivity, γ represents the speed of technological depreciation.

Wage Determination. Wages are determined via Nash bargaining, but echoing Binmore et al. (1986) and Hall and Milgrom (2008), with threats of temporary disruption. That is a firm cannot credibly let the worker return to unemployment and leave the match vacant. This environment has been explored in continuous time, with large firms and decreasing returns to scale by Elsby and Gottfries (2019). Following Elsby and Gottfries (2019), the threat point of the worker is not returning to unemployment but instead a temporary disruption during which time the worker receives a per period payoff ω_w. To ensure a worker would always prefer a temporary lapse in the negotiation phase it is assumed that the flow benefit associated with unemployment is sufficiently small. The firm in the event of negotiation breaking down incurs a cost ω_f. The total per period surplus of a match producing output z is given by $z + \omega_f - \omega_w$. Assuming the worker takes a share $\beta \in (0, 1)$ of the total surplus, the negotiated wage solves the equality

$$\beta(z + \omega_f - \omega_w) = w - \omega_w.$$

Rearranging gives a simple linear wage equation increasing in the productivity of the match.

$$w = \omega_0 + \beta z \quad \text{where,} \quad \omega_0 := \beta\omega_f + (1 - \beta)\omega_w \tag{1}$$

Negotiation will breakdown if the joint period surplus is negative. Although the worker would prefer to remain idle in the match, no wage can credibly be agreed upon. Negotiation thus breaks down when the per period surplus is negative, or when the match draw is less than the threshold productivity $z^\star$, where

$$z^\star = \omega_w - \omega_f. \tag{2}$$

It will be shown later that matches will separate when productivity falls below this threshold value which is independent of the state of the economy. A further virtue of wage bargaining in this setup is that it can accommodate Nash bargaining when workers look for jobs in employment. Typically, in this scenario, as discussed in Shimer (2006) and Gottfries (2019) the bargaining set becomes non-convex. A higher wage deters would be poachers and thus increases the expected duration of the match, thereby increasing the total surplus that is being bargained over. Akin to the timing in Krause and Lubik (2007), since negotiation, payment and production happens before turnover decisions, bargaining sets remain convex as the wage has no bearing on the duration of a match.

Idiosyncratic productivity shock. After production in a period, the idiosyncratic productivity shock shuffles the distribution of matches in the economy. The measure of unemployment at the start of period t is denoted as u_t and the measure of employed workers in productivity z is $e_t(z)$. The '+' subscript is used to denote these measures in the immediate aftermath of the productivity shock.

$$e_{t+}(z) = (1-\gamma_t)e_t(z) + \gamma_t \int e_t(z')g(z|z')dz' \quad \text{for, } z > z^\star \tag{3}$$

The new measure of employed workers of productivity z are those previously in match quality z not hit by a shock plus those, previously of higher match quality who draw a new productivity z following an idiosyncratic shock.

Firm value. At the time of bargaining, in period t, the value of a match of productivity z is denoted at $\Pi_t(z)$. The Bellman equation of which is given below, where $f_t(z)$ is the probability a worker of match quality z leaves for an alternative employer and V_t is the value of an unfilled vacancy.

$$\Pi_t(z) = z - w(z) + \frac{1}{1+r}\mathbb{E}_t\Bigg[(1-\gamma_t)\Big[f_t(z)V_{t+1} + (1-f_t(z))\,\Pi_{t+1}(z)\Big] \\ + \gamma_t \int^{z^\star} g(z'|z)V_{t+1}dz' + \gamma_t \int_{z^\star}^{z} g(z'|z)\Pi_{t+1}(z')dz'\Bigg]$$

Given free entry, the value of an unfilled vacancy will equal zero. Free entry, coupled with the the wage equation (1) gives (4).

$$\Pi_t(z) = (1-\beta)z - \omega_0 + \frac{1}{1+r}\Bigg[(1-\gamma_t)\,(1-f_t(z))\int \pi(\gamma_{t+1},\gamma_t)\Pi_{t+1}(z)d\gamma_{t+1} \\ + \gamma_t \int \pi(\gamma_{t+1},\gamma_t)\int_{z^\star}^{z} g(z'|z)\Pi_{t+1}(z')dz'd\gamma_{t+1}\Bigg] \tag{4}$$

Notice, this value function still retains a t subscript. The probability a worker exits the match $f_t(z)$ will be endogenized later and will be a function of the distribution of match quality in the economy at large. How the firm approximates this infinitely dimensional state variable will be returned to.

Active searchers. A worker engages in search behavior with probability α_u if they start the period unemployed or with probability α_e if employed and have not been hit by a productivity shock making them redundant. Given one engages in search, they actively search - send out an application with probability $(1-\exp(-\lambda_s))$, $s \in \{u,e\}$. At time t, the mass of active searchers is defined as

$$a_t := \alpha_u\,(1-\exp(-\lambda_u))\,u_t + \alpha_e\,(1-\exp(-\lambda_e))\int e_{t+}(z)dz.$$

Since a given worker applies to a given vacancy with probability $1/v$. The number of applicants per job $k \sim \text{bin}(a, \frac{1}{v})$, where a is the active number of job searchers. For a continuum of job searchers and vacancies, the number of applicants per vacancy is distributed with a Poisson distribution with parameter θ^{-1}, where θ is defined as the effective market tightness, $\theta := v/a$, see Appendix A.2.

Firm hiring. A firm's decision on whether and which candidate to hire after screening is straightforward. They hire a worker if they generate positive value for the firm, the match quality exceeds the threshold $z^\star$. If multiple candidates apply to a vacancy a firm chooses the highest match quality draw from the set. So, given k applicants, the probability that the match draw z is the best candidate and offered the job is given by $q(z|k)$.

$$q(z|k) = 0 \qquad \text{for, } z < z^\star$$

$$q(z|k) = k \times \text{Prob(no better } z' \text{ arrives from } k-1 \text{ trials)} \times \text{Prob(} z \text{ is drawn)} \qquad \text{for, } z \geq z^\star$$

At period t, this probability can be expressed as

$$\begin{aligned} q_t(z|k) &= 0 && \text{for, } z < z^\star \\ q_t(z|k) &= k\left(1 - \frac{1}{a_t}u_t\int_z g_u(z')dz' - \frac{1}{a_t}\int^z e_{t+}(z'')dz''\int_z g_e(z')dz'\right)^{k-1} \\ &\quad \frac{1}{a_t}\left(u_t g_u(z) + \int^z e_{t+}(z')dz' g_e(z)\right) && \text{for, } z \geq z^\star. \end{aligned} \tag{5}$$

Contained in equation (5) are the measures $g_s(z)$, $s \in \{u, e\}$. These represent the measure of actively searching workers in each state who apply to a vacancy with match quality z. These objects will be endogenized shortly. Since the number of applications per vacancy follows a Poisson distribution, shown in Appendix A.2, the unconditional probability $q_t(z)$ is given as

$$q_t(z) = \sum_{k=0}^{\infty}\left(\frac{\theta_t^{-k}}{k!}e^{-1/\theta_t}q_t(z|k)\right). \tag{6}$$

Finally, the decision to post a vacancy for a firm is determined by the firm's expected return to a match (equation (4)), the probability of matching (equation (6)), and the cost of posting and screening, κ and σ, respectively. The posting cost κ is always paid by a recruiting firm, and the screening cost is paid conditional on meeting a suitable candidate. Firms continue to enter the market until the value of posting a vacancy V_t is zero.

$$V_t = -\kappa + \int_{z^\star}^{\infty} q_t(z)\left(\Pi_t(z) - \sigma\right)dz \tag{7}$$

Equation (7) pins down the the level of vacancies posted in period t, v_t.

Job finding rates. Since firms only make offers to workers they are confident of employing we know the worker's job filling rate and job contact rate are identical and they will thus be used interchangeably.

$$q_t(z) = \frac{m_t(z)}{v_t} \tag{8}$$

The number of matches of type z at time t are defined as $m_t(z)$. Further, $m_t(z, z')$ is the number of matches of match quality z hired from match quality z', the absence of z', written as $\varnothing$, denotes a hire from unemployment.

$$\begin{aligned} m_t(z, \varnothing) =& m_t(z). \frac{u_t g_u(z)}{u_t g_u(z) + \int^z e_{t+}(z')dz' g_e(z)} \\ m_t(z, z') =& m_t(z). \frac{e_{t+}(z')g_e(z)}{u_t g_u(z) + \int^z e_{t+}(z')dz' g_e(z)} \quad \text{for, } z' < z \end{aligned} \tag{9}$$

Finally, the probability a worker leaves a match of quality z for better suited employment is given below. This equation is required for computation of the value of a match to the firm, equation (4).

$$f_t(z) = \frac{\int m_t(z', z)dz'}{e_t(z)}$$

Worker's application strategy. Since wages are monotonic in match productivity, equation (1), and firms strictly prefer higher match quality, a worker's optimal strategy is to apply to the maximum match quality they are exposed to. Define $\tilde{G}_s(z)$ as the cumulative distribution function of the best match quality an actively searching worker in state $s \in \{u, e\}$ samples. Recall, a worker is actively searching with probability α_s. Since the number of offers are governed by a Poisson distribution, the probability an actively searching worker gets no offers is $\exp(-\lambda_s)$. Thus the probability a worker gets no offers above $\tilde{G}_s(z)$ will follow a Poisson distribution with a reduced rate equal to $\lambda_e(1 - G(z))$. Where $G(z)$ is the primitive sampling cumulative distribution function.

$$\tilde{G}_s(z) = \exp(-\lambda_s(1 - G(z))) \quad \text{for } z \geq 0$$

Match quality is assumed to be drawn from an exponential distribution with parameter λ_z, so

$$\tilde{G}_s(z) = \exp(-\lambda_s \exp(-\lambda_z z)) \quad \text{for } z \geq 0.$$

Notice that $\tilde{G}_s(0) = \exp(-\lambda_s)$, the probability an actively searching worker looks at zero open vacancies. For $z > 0$ the distribution corresponds to a truncated type 1 extreme value distribution (Gumbel) with scale parameter equal to $1/\lambda_z$ and location parameter equal to $\log(\lambda_s)/\lambda_z$.[4] The probability any worker, actively searching or not, in state s, has no $Z > z$ amongst all the their visible vacancies, is given by $G_s(z)$.

$$G_s(z) = 1 - \alpha_s + \alpha_s \exp(-\lambda_s \exp(-\lambda_z z)) \quad \text{for } z \geq 0.$$

The measure of quality z a worker in state s applies to is the derivative of $G_s(z)$ and given by

$$g_s(z) = \alpha_s \lambda_s \lambda_z \exp(-\lambda_s \exp(-\lambda_z z) - \lambda_z z) \quad \text{for, } z > 0.$$

This gives the different distributions that unemployed and employed workers seemingly sample vacancies from. Notice they are independent of the aggregate state of the economy, and the degree of difference will depend on differences in search behavior. Further, this is a measure rather than a probability density as it does not integrate to one. The definite integral is equal to $\alpha_s(1-\exp(-\lambda_s))$, which is the probability a worker observes at least one vacancy.

Labor dynamics. At the very end of the period employment transitions occur. The evolution of unemployment from period t to period $t+1$ is given by equation (10). That is, the unemployed tomorrow are those previously unemployed, net of those who found a job, plus those previously employed who received a productivity shock sending them to unemployment.

$$u_{t+1} = u_t - \int m_t(z, \varnothing) dz + \gamma_t \int^{z^\star} \int e_t(z') g_t(z|z') dz' dz \tag{10}$$

The measure of employed workers in matches of productivity z evolves period to period according to equation (11). The first term on the right hand side are those of productivity z after the productivity shock hits. The final two terms are those who move up the job ladder to productivity

[4]The cumulative distribution function of a Gumbel distribution with location parameter μ and scale parameter $\beta > 0$ is given by $F(x; \mu, \beta) = \exp(-\exp(-(x-\mu)/\beta))$.

z net of those who move to better matches.

$$e_{t+1}(z) = e_{t+}(z) + \int m_t(z, z')dz' - \int m_t(z', z)dz' \tag{11}$$

Match formation and dissolution. To understand why a threshold productivity $z^\star$ is sufficient in determining the feasibility of matches, consider the value function of a matched firm, equation (4). All future productivity shocks to the match are drawn from a distribution bounded by the current match productivity. Thus if the match produces no output today, because of negotiation breakdown, there is no associated option value in retaining it. In the context of the model therefore, separations should be interpreted as being initiated by the firm rather than the worker. It is still to be established that a worker would be willing to accept a match of quality $z^\star$. The value function of an unemployed worker is given by the equation below.

$$U_t = b + \frac{1}{1+r}\left[\int \pi(\gamma_{t+1}, \gamma_t)\left(U_{t+1} + \int_{z^*} \mu_t(z, \varnothing)\left(W_{t+1}(z) - U_{t+1}\right)dz\right)d\gamma_{t+1}\right]$$

The parameter b is the per period payoff a worker receives while unemployed, and encompasses both pecuniary and non-pecuniary aspects. The following period, after the realization of the new aggregate state, one of two things can happen. Either the worker stays in unemployment, or with probability $\mu_t(z, \varnothing)$ finds a job with match quality z, yielding present value $W_{t+1}(z)$. The probability measure $\mu_t(z, \varnothing)$ can be computed by dividing the expression for $m_t(z, \varnothing)$ (equation (9)) by the stock of unemployed u_t. The Bellman equation for an individual employed at match productivity z at time t is presented below. If the worker is not hit by an adverse productivity shock they either stay in their current job or move to a higher productivity job z' with probability measure $\mu_t(z', z)$. The probability measure is computed by dividing the expression for $m_t(z', z)$ by the measure of employed at productivity z, given by $e_t(z)$. If the match is subject to a productivity shock, either the worker stays in their job at a lower productivity, in the interval $[z^\star, z]$. Or, if it is sufficiently bad and the new match productivity is less than $z^\star$, the worker separates from the match, becoming unemployed.

$$W_t(z) = \beta z + \omega_0 + \frac{1}{1+r}\Bigg[\int \pi(\gamma_{t+1}, \gamma_t)\Bigg((1-\gamma_t)\left(W_{t+1}(z) + \int_z \mu_t(z', z)\left(W_{t+1}(z') - W_{t+1}(z)\right)dz'\right) + \gamma_t \int_{z^\star}^{z} g(z'|z) W_{t+1}(z') dz' + \gamma_t \int_0^{z^\star} g(z'|z) U_{t+1} dz'\Bigg) d\gamma_{t+1}\Bigg]$$

For an unemployed worker to accept a job of match quality $z \geq z^\star$ it must be, that for all states of the world, summarized by the index t, $W_t(z^\star) \geq U_t$. Comparing the value functions for the employed and unemployed, for sufficiently low b this can be guaranteed for any state of the economy. In the estimation, the flow payoff associated with unemployment is treated as a free parameter and it is assumed it is sufficiently low that acceptance decisions are invariant to the business cycle. While attractive from a tractability perspective having worker's search decisions invariant to the cycle may omit other interesting mechanisms.[5]

Aggregation. When deciding on whether to post vacancies, the state variables facing a firm are the parameter γ and the entire distribution of match quality in the economy. The distributional state variable makes the solution somewhat problematic. A solution relies on something somewhat similar to the approach taken by Krusell and Smith (1998).[6] Rather than observing the entire distribution a firm observes three moments of the distribution of match quality and makes a parametric assumption. The firm perfectly observes the unemployment rate and the mean and standard deviation of productivity amongst those in employment, m_1 and m_2, respectively. Further, firms assume the match quality distribution follows a log-normal distribution, truncated at $z^\star$. That is, rather than observe $e(z)$ perfectly, they observe u perfectly and assumes $e(z) \approx \tilde{e}(z; u, \check{m}_1, \check{m}_2, z^\star)$, equation (12). The parameters of the distribution, $\check{m}_1$ and $\check{m}_2$ can be backed out to rationalize the

[5]See for example Barnichon and Zylberberg (2019), who suggest screening as an explanation for the abundance of 'over-qualified' workers employed in a recession.

[6]Similar solution algorithms in this context are implemented in Krolikowski (2019) and Walentin and Westermark (2018).

mean and standard deviation of match productivity, m_1 and m_2.

$$\tilde{e}(z;u,m_1,m_2,z^\star) = \begin{cases} (1-u).B(z^\star,\check{m}_1,\check{m}_2).\frac{1}{z\check{m}_2\sqrt{2\pi}}.\exp\left(\frac{(\log(z)-\check{m}_1)^2}{2\check{m}_2{}^2}\right) & \text{for, } z \geq z^\star \\ 0 & \text{for, } z < z^\star \end{cases} \tag{12}$$

where, $B(z^\star,\check{m}_1,\check{m}_2) = \left(\frac{1}{2} - \frac{1}{2}\text{erf}\left[-\frac{(\log(z^\star)-\check{m}_1)}{\sqrt{2}\check{m}_2}\right]\right)$ and $\text{erf}(\cdot)$ is the Gauss error function.

Given this parameterization of firms' beliefs, the value of a match of productivity z can be written as a function of four aggregate state variables. They are: γ, the idiosyncratic probability of the match being hit by an adverse productivity shock; u, the unemployment rate; and m_1 and m_2, the mean and standard deviation of match productivity amongst matches in the economy.

$$\begin{aligned}\Pi(z,\gamma,u,m_1,m_2) = (1-\beta)z - \omega_0 + \\ \frac{1}{1+r}\Bigg[(1-\gamma)\left(1-f(z,\gamma,u,m_1,m_2)\right)\int \pi(\gamma',\gamma)\Pi(z,\gamma',u',m_1',m_2')d\gamma' \\ +\gamma\int \pi(\gamma',\gamma)\int_{z^\star}^{z} g(z'|z)\Pi(z',\gamma',u',m_1',m_2')dz'd\gamma'\end{aligned}$$

Included in the value function are the beliefs firms have about the evolution of the aggregate state variable and the probability that a worker leaves for an alternative employer $f(\cdot)$. These are computed by replacing equation (3) with the approximation

$$e_{t+}(z) = (1-\gamma_t)\tilde{e}(z;u,m_1,m_2,z^\star) + \gamma_t\int \tilde{e}(z';u,m_1,m_2,z^\star)g_t(z|z')dz' \quad \text{for, } z > z^\star.$$

The solution algorithm is written up in full in Appendix A.3.

5 Estimation

Data. To discipline the model, parameters are estimated to fit aggregate time series of the U.S economy from 2001 to 2014, inclusive. Monthly employment transitions: the rate employed workers leave for unemployment; and the rates the unemployed and employed find new jobs are computed from the Current Population Survey (CPS). Additionally, the aggregate unemployment rate is also computed from the CPS. Attention is restricted to 18-65 year old workers. The monthly vacancy rate is taken from the Job Openings and Labor Turnover Survey (JOLTS) and a composite labor market tightness is also computed as the computed vacancy rate divided by unemployment. The

JOLTS data begins in December 2000, at the start of the period examined. From the six labor market monthly series, the mean, standard deviation, and cross- and autocorrelations are computed. Finally, the mean labor share is taken from the Bureau of Labor Statistics (BLS). All moments are reported in the upper panel of Table 5.

Parameterization. The model is calibrated at a monthly frequency. Certain parameters are calibrated ex ante of estimation. The discount rate is set at $r = (1.05)^{1/12} - 1$, implying an annual discount rate of 5%. To be consistent with the empirical definition of unemployment it is assumed that the unemployed actively engage in search in every period, $\alpha_u = 1$.[7] To provide a numeraire λ_z is normalized to one. Implying that the mean match quality of a random worker and job opening is one. The per period outside option of the firm is assumed to be zero, $\omega_f = 0$. Finally, embedded in the model is an endogenous urn-ball matching function. The urn-ball matching function is as documented by Petrongolo and Pissarides (2001) "*too naive to be empirically a good approximation*". Without additional parameters the model struggles to replicate simultaneously the level of unemployment and vacancies seen empirically. Consequently, it is assumed that not all vacancies are empirically observable, rather a fraction η. So, given υ vacancies posted, the empirical counterpart is

$$\upsilon^{\star} = \eta \upsilon \quad \text{where, } \eta \in (0, 1).$$

Revisiting the microevidence. The novel parameters of the model presented relate to the search and screening behavior of workers and firms. In particular, the relative number of jobs sampled by the actively searching employed and unemployed (λ_e/λ_u) and the relative cost in screening compared to recruiting intensity (σ/κ). In a pre-estimation step these ratios are fixed relative to other parameters to ensure that they are consistent with the microevidence presented earlier. The remaining parameters will be estimated to target aggregate data. To quantify the strength of the mechanism, the baseline model will be compared to a model in which employed and unemployed workers have the same search technology $\lambda_u = \lambda_e$ and there is no screening cost $\sigma = 0$. Fixing

[7] Notice, this does not imply the unemployed always sample a positive number of job openings. But rather every period sample a number determined by a Poisson distribution governed by parameter λ_u.

these parameters ex ante has the advantage that the model's mechanism can be quantified relative to a model with the same degrees of freedom in fitting the data.

Evidence from the SCE suggests that an employed worker's best job offer is 22% better paid than their unemployed counterparts, see the final two columns of Table 3. The mean of the best match productivity offer a worker in state $s \in \{u, e\}$ samples from is denoted as $\overline{z}_s$, and given by the upper expression in equation (13). Where, $E_1(\cdot)$ is the exponential integral and γ the Euler-Mascheroni constant.[8] The expression is derived in Appenddix A.4. Since wages are linear in productivity, the 22% premium can be pinned down by the first equality in (13). In the baseline model, λ_e is fixed such that this equality always holds. On the firm side, evidence from Switzerland suggests that screening costs are between 57% and 69% of total hiring costs, depending on which costs are attributed to screening. The parameter σ is fixed such that the screening cost share is the midpoint of this range. Where $\overline{h}$ is the mean level of hiring and $\overline{v}$, the mean level of vacancies adjusted by the fraction observed, in the data. It is necessary to weight the costs by these margins since in the model the screening cost is only paid when a firm successfully hires and the vacancy cost is paid irrespective of filling the position.

$$
\begin{aligned}
\overline{z}_e &= \frac{0.22\omega_0 + 1.22\beta\overline{z}_u}{\beta} && \text{where,} && \overline{z}_s = \frac{1}{\lambda_z(1-e^{-\lambda_s})}\left(\log(\lambda_s) + E_1(\lambda_s) + \gamma\right) \\
0.63 &= \frac{\sigma\overline{h}}{\sigma\overline{h} + \kappa\overline{v}} && \text{where,} && \overline{v} = \frac{\overline{v^\star}}{\eta} \quad \text{and} \quad \overline{h} = \overline{f_0 u + f_1(1-u)}
\end{aligned} \tag{13}
$$

The vector of parameters to be estimated is defined as the (9×1) vector $\vec{\theta}$, given below. Note the exclusion of the parameters λ_e and σ which are calibrated ex ante to meet the constraints in (13).

$$
\vec{\theta} := (\alpha_e, \lambda_u, \kappa, \eta, \rho, \gamma_0, \sigma_\gamma, \beta, \omega_w)'
$$

Estimation protocol. The model is simulated at a monthly frequency and moments are computed

[8] The Euler-Mascheroni constant is equal to 0.577 to three decimal places and the exponential integral is given by

$$
E_1(z) = \int_z^\infty \frac{e^{-t}}{t} dt.
$$

from the simulated series in an identical way to the empirical series. In practice simulated moments are computed by simulating a series for a given sequence of γ drawn using the Tauchen (1986) method. The series begins with a burnin period of fifty years and the moments are computed from 14 subsequent years (the same length as the data). Since the simulated moments depend on the sequence of γ as well as the vector of structural parameters $\vec{\theta}$ this process is repeated one hundred times and the mean of each moment is computed. Formally the estimated vector $\hat{\vec{\theta}}$ is the solution to

$$\hat{\vec{\theta}} := \arg\min_{\vec{\theta}\in\vec{\Theta}} \left(\frac{1}{M} \sum_{i=1}^{M} m_i(\vec{\theta}) - m \right)' \Omega \left(\frac{1}{M} \sum_{i=1}^{M} m_i(\vec{\theta}) - m \right).$$

Where, $m_i(\vec{\theta})$ is a (34×1) column vector of simulated moments of the i^{th} sequence of γ's. M is the number of re-simulated series, set at 100. The (34×34) weighting matrix $\Omega = \text{diag}(\omega * \frac{1}{m})$, where m is the column vector of moments taken from the data and ω is a vector of additional fixed weights. In order to make sure the simulated series matches the levels well. Moments that are means are given tenfold more weight. Two models are estimated. One in which employed workers, on average, sample wages that offer 22% more than unemployed workers and resolve equation (13). The other, absent the screening mechanism, in which workers sample from the same distribution ($\lambda_u = \lambda_e$) and screening is costless ($\sigma = 0$).

Parameter estimates are presented in Table 4. The estimates suggest a different paradigm between the two models in how the unemployed and employed search for jobs. In the baseline model, the unemployed actively search almost eight times as frequently as the employed. However, conditional on searching the employed are exposed to more than three times as many opportunities. Although this paper does not take a stand on what drives the large difference in search behavior across employment status, Faberman et al. (2017) provides empirical support to the phenomenon.[9] In the comparison model, the unemployed and employed have the same search technology. In this

[9] Potential explanations are explored theoretically by Bradley and Gottfries (2021) and Arbex et al. (2019). In Bradley and Gottfries (2021) opportunities arrive stochastically to the employed and unemployed and if one goes to the market less frequently (the employed) they are likely to have a greater number of opportunities at hand. Arbex et al. (2019) model search in a network structure and those in employment are on average better placed through their exposure to peers to search more efficiently.

scenario, conditional on actively searching, the employed are less likely to find employment and thus the job finding rate of the employed is rationalized by them applying to jobs more frequently.

Table 4: Parameter estimates

Baseline Model			
Worker search parameters		Firm hiring parameters	
λ_u	0.753	κ	7.035
α_e	0.127	σ	17.569*
λ_e	2.353*	η	0.547
Wage bargaining parameters		Aggregate volatility parameters	
β	0.451	ρ	0.883
ω_w	0.752	γ_0	0.003
		σ_γ	0.003
No Screening			
Worker search parameters		Firm hiring parameters	
λ_u	0.883	κ	13.608
α_e	0.248	σ	0*
λ_e	0.883*	η	0.433
Wage bargaining parameters		Aggregate volatility parameters	
β	0.318	ρ	0.877
ω_w	1.082	γ_0	0.003
		σ_γ	0.002

The two models differ in their calibration of the parameters λ_e and σ. In the baseline model these are restricted to match two micro-moments related to search behavior, see equation (13). Under the no screening model they are fixed such that $\sigma = 0$ and $\lambda_u = \lambda_e$. Restricted parameters are denoted by a $\star$.

In the baseline model, hiring cost parameters are difficult to interpret in isolation. Since the vacancy cost is paid on every vacancy and the screening cost is not, a comparison of the values of κ and σ alone is uninformative. A more informative assessment is performed in section 6.2. Unsurprisingly, since there is only one hiring cost absent screening, this cost is larger than in the baseline model. The rent share parameter β in both models is typical of a macro calibration, but arguably larger than the majority of estimates in the empirical micro literature, see Table 4 of Manning (2011).

Table 5: Fit of the Baseline Model

Empirical Moments

	s	f_0	f_1	u	v	θ
Mean	0.014	0.219	0.019	0.066	0.027	0.447
Standard dev.	0.002	0.048	0.003	0.019	0.004	0.17
Correlations						
f_0	-0.485					
f_1	-0.322	0.641				
u	-0.73	-0.668	0.86			
v	-0.913	0.651	0.876	-0.78		
θ	0.48	0.698	-0.745	-0.921	0.886	
Autocorrelation	0.666	0.893	0.824	0.993	0.929	0.979
Labor Share	Mean = 0.59					

Simulated Moments

	s	f_0	f_1	u	v	θ
Mean	0.013	0.218	0.021	0.063	0.026	0.456
Standard dev.	0.003	0.017	0.9×10^{-3}	0.013	0.003	0.136
Correlations						
f_0	-0.818					
f_1	-0.573	0.665				
u	-0.844	-0.816	0.579			
v	-0.818	0.904	0.915	-0.179		
θ	0.79	0.387	-0.532	-0.912	0.754	
Autocorrelation	0.852	0.844	0.766	0.982	0.697	0.953
Labor Share	Mean = 0.593					

Data covers the U.S economy from 2001 to 2014, inclusive. Employment dynamics and the unemployment rate are computed from the CPS and vacancy rates from the JOLTS. All are computed at a monthly frequency. The labor share comes from the BLS and is reported at a quarterly frequency, it is logged and HP filtered as described in the main body.

The fit of the estimated model to the data can be assessed by comparing the two panels of Table 5. As a result of extra weight being placed on fitting levels in the weighting matrix Ω mean moments are fitted particularly well. The model contains a trade-off in matching the second moments of the series. A direct separation shock is a special case of a shock to match productivity and it is this that drives aggregate fluctuations. The simulated series therefore can comfortably match the volatility in separation rates but since the other series are a consequence of that shock, the volatility of other series are under-predicted. The under-prediction is particularly pronounced for the job-to-job transition rate series.

As has been discussed, it is the search behavior of workers and hiring behavior of firms that allow the model to simultaneously generate a downward sloping Beveridge curve and match workers transitional dynamics. This can be seen in the fit of the model absent screening, presented in Table 9. This model generates an upward sloping Beveridge curve, exhibited by a positive correlation between unemployment and vacancies. In the comparison model, from a firm's perspective, a high unemployment rate is desirable. Since the two types of workers have the same search technology, a high unemployment rate means vacancies are more likely to be filled. In the baseline model, a higher unemployment rate also increases the probability of hiring a worker, motivating more vacancies to be posted. However, because of the different search technologies, conditional on hiring a worker, an employed worker will on average have drawn a higher match productivity. The baseline model is thus able to generate correlations between labor market tightness and unemployment of the correct sign.

Aggregation. In the model, the free entry condition depends, as is typical, on the aggregate state and the unemployment rate. Additionally however, since firms also hire from the pool of employed workers, the distribution of match quality across employed workers also features in a firm's decision making process. To reduce the dimensionality of the problem, it is assumed that rather than observing the full distribution firms can observe the first two moments and assume match quality follows a truncated log-normal. In Appendix A.6, the accuracy of this approximation is explored. Simulated transition rates of the model are compared with synthetic transition rates, computed assuming the firm's projection is accurate.

6 Results

6.1 Aggregate Series

To better understand the mechanism of the model the main series used in estimation are fitted directly. To do this the baseline model is simulated for a thousand years following the same burnin as is used in estimation. The starting point is the point of time that bears most resemblance to the first period in estimation, January 2001.[10] Following this initialization the realization of γ is backed out to best match the vacancy series following another two year burnin. The fit of the model is presented in Figure 1 along with the corresponding γ series.

Propagation. Aggregate fluctuations are driven by the turbulence in the economy – the exogenous Markov process for γ. Following an adverse shock, an increase in the value of γ, there are two immediate impacts. From a firm's hiring perspective, the expected duration of a match reduces as any potential employee is more likely to be hit by a negative productivity shock. Consequently, the value to a firm of hiring a given worker reduces and so therefore does the value of posting a vacancy. A reduction in vacancies reduces the flow out of unemployment, increasing the stock of unemployed, and slows movement up the job ladder. The second immediate response concerns already matched agents. Following an increase in γ a greater proportion are hit by negative productivity shocks, increasing the speed workers fall down the job ladder and increasing the pool of unemployed.

The immediate response would be isomorphic in any standard search model with this shock structure – that is an increase in the pool of unemployed through increased separations and less hiring. Where the mechanism in this model differs from the literature is in the propagation of these shocks. In standard random search models, with/without on the job search (assuming that the

[10]The criterion for this is the log difference between each monthly series in the simulated and empirical series.

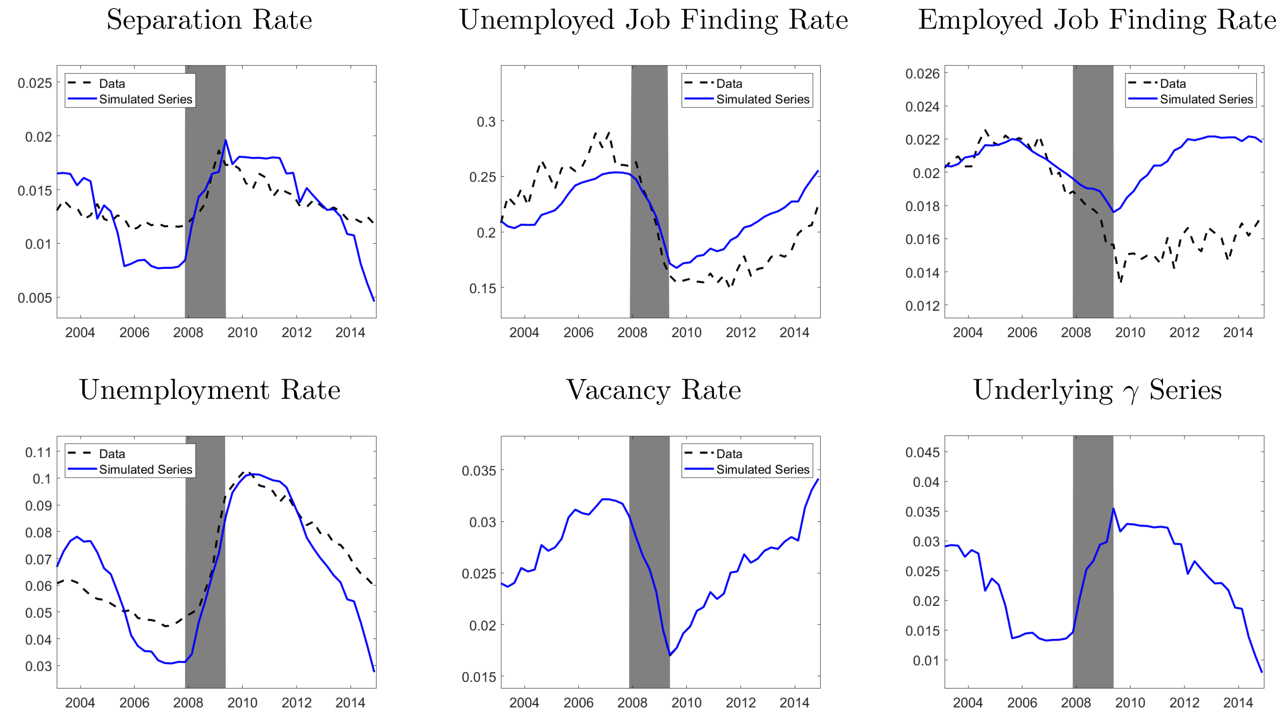
employed search less) a higher unemployment rate will feedback into a greater incentive to post vacancies. A greater number of active searchers means that vacancies are more likely to be filled increasing the returns in posting. It is exactly this mechanism that tilts the Beveridge curve to become upward sloping in the comparison model. It turns out, that when shocks generate movement in the hiring and separation rates this feedback is particularly strong. However, in the model presented here, although the employed search less often, conditional on searching they sample from a better match specific productivity distribution. From the perspective of a hiring firm, a large pool of unemployed workers implies the vacancy is more likely to be filled, but filled by on average a worse match than if there were fewer unemployed. Although the model is quite different there is a similar strategic complementarity between vacancy posting and on-the-job search in Eeckhout and Lindenlaub (2019). In Eeckhout and Lindenlaub (2019) when an employed worker finds a job they are more productive than an unemployed worker and thus the economy benefits from greater search intensity of the employed. In this model, the employed are not inherently more productive, but by sampling from more independent draws they are likely to move to more productive matches.

Shift in the Beveridge Curve. Inspection of Figure 1 shows that the 2008 downturn was caused by a rapid expansion of γ over the recession years. The high value then stayed high beyond the recession dates. Unemployment responded quickly and continued to grow following the recession whereas there was an upturn in the vacancy rate immediately following the end of the crisis. Plotting this on the vacancy-unemployment space, as is done in Figure 2 shows the shift in the Beveridge curve following the crisis.

If one models both the in- and outflow from unemployment, the response of unemployment to a shock is extremely rapid.[11] However, as is documented empirically by Moscarini and Postel-Vinay (2016) the distribution of matches amongst the employed is a slower moving object. In the context of the model, after the rapid expansion of unemployment the new hires from unemployment are on average in worse matches. This comes through a standard job ladder mechanism but also

[11] This can be seen by the accuracy of a steady-state unemployment rate in predicting the contemporaneous unemployment rate as is documented by Jolivet et al. (2006)

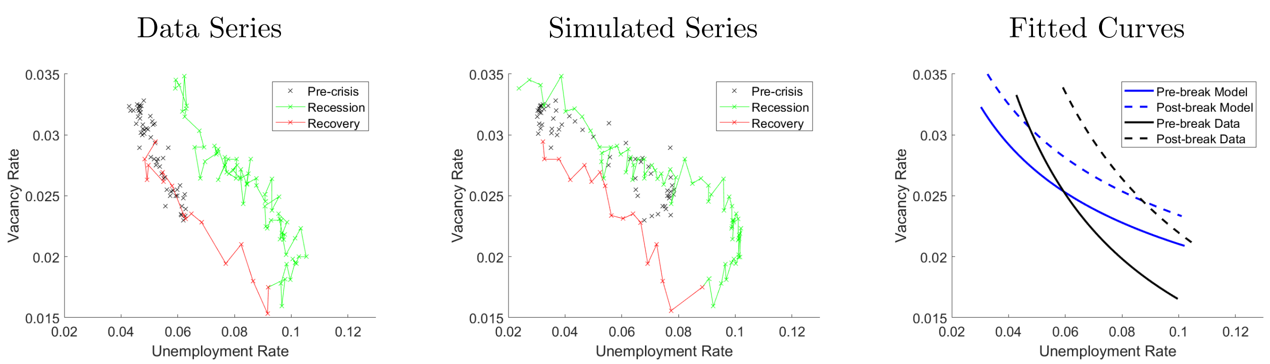
Figure 1: Fit of Series



The simulated series are backed out to replicate the observed vacancy rates in the data.

because they sample matches from a stochastically dominated distribution. These poorly matched workers provide a different opportunity to recruiting firms. They are less likely to hire them than an unemployed person, but conditional on hiring them they can expect a higher match quality and longer match duration. Consequently, vacancy posting increases but rather than hiring the unemployed and reducing unemployment more vacancies are left unfilled or filled by the already employed.

Figure 2: Beveridge Curve



To formally demonstrate that there is a structural break in the unemployment-vacancy relation in the simulated series the relationship is estimated. The following relationship is estimated on the model's simulated series and on the unemployment rate taken from the CPS and vacancy rate from the JOLTS, where α, β, γ and τ are estimated and $\mathbf{1}\{X > Y\}$ represents an indicator function taking the value one if the statement is $X > Y$ is true and zero otherwise.

$$\log(v_t) = \alpha\{t < \tau\} + \beta\{t \geq \tau\} + \gamma \log(u_t) + \epsilon_t \tag{14}$$

The parameter τ is estimated by estimating all other parameters conditional on every possible value of τ and selecting the value with the smallest residual sum of squares. Parameter estimates are given in Table 6 and the fitted values are plotted in the final panel of Figure 2. The model predicts a breakdown in the unemployment-vacancy relationship almost a year later than what is predicted by the data. The size of the shift, the difference between α and β, is also smaller. Suggesting the model generates approximately a third of the observed shift in the Beveridge curve.

Table 6: Fitted Beveridge Curves

	τ	α	β	γ
Simulated series	August 2009	-4.68	-4.58	-0.36
Empirical series	September 2008	-6.01	-5.72	-0.83

This table displays the least square estimates of the parameters in equation (14), for simulated and real data.

The matching function. Through modeling the explicit application and screening decisions of workers and firms, the model has an endogenous urn-ball matching function. The majority of the search and matching literature assumes the matching function to be Cobb-Douglas. Evaluating the market through the lens of a Cobb-Douglas matching function, the matching efficiency term has been shown to be pro-cyclical, Borowczyk-Martins et al. (2013) and Barnichon and Figura (2015). Here it is shown that the simulated model appears well approximated by a Cobb-Douglas matching function with pro-cyclical matching efficiency.[12]

[12] Firm screening is not the only mechanism that can generate such a phenomenon. The same pro-cyclicality has been shown to be generated with: firm recruitment intensity, Gavazza et al. (2018); hiring standards Sedláček (2014); and occupational mismatch Sahin et al. (2014).

Assume a constant returns to scale Cobb-Douglas matching function of the form

$$M(U_t, V_t) = A_t U_t^{1-\eta} V_t^{\eta}.$$

Dividing by U_t so the left hand side becomes the job finding rate of the unemployed, taking logs and including month dummies τ_t gives

$$f_t = \mu + \eta\theta_t + \tau_t + \epsilon_t. \tag{15}$$

Where lower case denotes logs, $\theta_t := \log(V_t/U)_t)$ and the matching efficiency is $A_t = \exp(\mu + \tau_t + \epsilon_t)$. Estimates of the regression on simulated and actual data are presented in Table 7. While the parameters differ, the elasticity of matches with respect to vacancies is more than twice as high in the data, the behavior of matching efficiency is qualitatively similar. The model is able to generate 30% of the dispersion in efficiency as in the data and both exhibit pro-cyclicality. Finally, despite being knowingly misspecified the constant returns to scale Cobb-Douglas matching function appears to explain the simulated series extremely well.

Table 7: Estimates of Cobb-Douglas Matching Function

	η	$\mathrm{E}(A_t)$	$\mathrm{std}(\log(A_t))$	$\mathrm{corr}(A_t, u_t)$	R^2
Simulated series	0.22	0.26	0.03	-0.03	0.95
Empirical series	0.48	0.32	0.10	-0.15	0.79

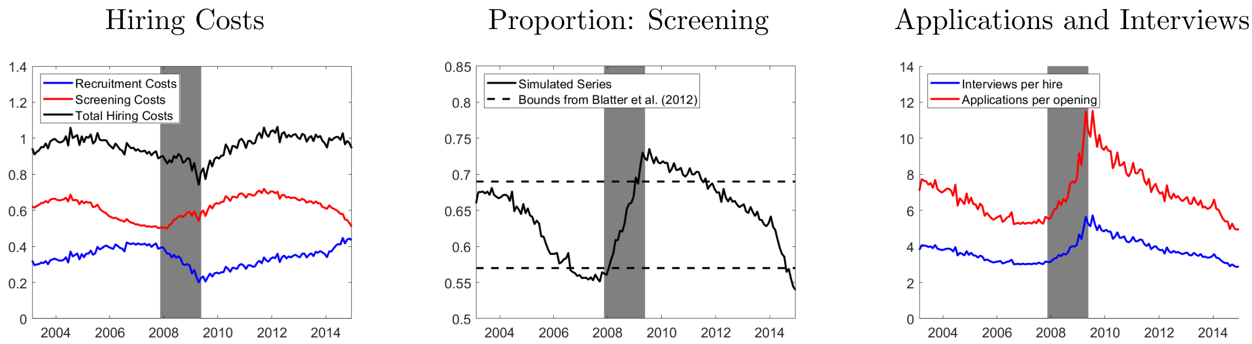
Table presents results of least square regressions of equation (15) for actual and model simulated data.

6.2 The Search Process

In calibration, standard aggregate moments of the labor market were targeted. However, workers and firms follow a quite different search process to as is typically modeled. The goal of this section is to (i) understand the cyclical component of the process and (ii) see how the model compares to empirical studies on recruitment.

Hiring costs. A firm in the model first decides whether or not to post a vacancy. They then screen their set of candidates and decide which applicant to hire. The hiring cost can be decomposed into two distinct parts, a recruitment and a screening intensity. As in section 3 the recruitment intensity is defined as the effort of getting people to apply and the screening intensity is the selection of the candidate. In the model these are interpreted as the total screening costs σ and recruitment cost κ paid by firms per period. The first panel of Figure 3 depicts how these vary over the series previously simulated.

Figure 3: Cost of Hiring and Applications



Screening costs are counter-cyclical. In downturns, there are many more unemployed, who send out applications more frequently than the employed leading to more applicants per vacancy. This is depicted in the third panel. A greater number of applicants thus generates a greater need for screening. Conversely, recruitment costs are pro-cyclical. Since these are defined as the cost of posting a vacancy multiplied by a constant vacancy cost, they like vacancies, exhibit strong pro-cyclicality. Quantitatively, since screening costs are twice as large as recruitment costs, it turns out that the counter-cyclicality of screening costs dominate and total hiring costs are counter-cyclical also. Over the period of simulation total hiring costs and the unemployment rate have a correlation of approximately one half.

Since recruitment costs are pro- and screening costs are counter-cyclical, the share of total costs attributed to screening is extremely counter-cyclical. The share of hiring costs attributed to screening workers is exhibited in the second panel of Figure 3. The share of hiring costs spent on screening rises from around 55% pre-crisis to almost 75% at its height. For the majority of the series the cost share is within the bounds implied by the audit study of Blatter et al. (2012).

Applications and interviews. In the model workers sample many jobs and firms receive many applicants to which one candidate is offered the job. The number of applications sent by workers and the number of screened candidates have not been specifically targeted in estimation. As is shown in Figure 3 when the pool of unemployed is large, in downturns, the number of applications sent out and number of applicants screened is large. While this is probably uncontroversial the volume of applicants and those screened may or may not reflect something close to the number in reality. Table 8 reports recent evidence regarding the number of applications and number of interviews job openings are likely to receive and conduct.

Table 8: Applications and Interviews

	Time period	Mean
Model		
Feasible candidates screened	2003-2015	3.74
Applicants per opening	2003-2015	6.91
Data - number interviewed per hire		
Blatter et al. (2012), Swiss administrative data	2000-2004	4.8
Villena-Roldan (2012), National Employer Survey	1997	2.74-6.48
Wolthoff (2018), Employment Opportunities Pilot Project	1987	5.7
This paper, MCSUI	1992-1995	6.2-6.4
Data - applicants per opening		
Davis and Samaniego de la Parra (2017), DHI Group, Inc.	2012-2017	5.1
Marinescu and Wolthoff (2020), Careerbuilder.com	2011	59
Wolthoff (2018), Employment Opportunities Pilot Project	1987	9.7
This paper, MCSUI	1992-1995	17.7-18.4

The upper panel of Table 8 presents the mean number of feasible applicants screened and the total volume of applications sent out per vacancy. Notice, the former is not necessarily the theoretical counterpart for the number interviewed and it is best interpreted as a lower bound as many infeasible candidates may well be interviewed in reality. The distribution of feasible candidates, given a hire, at time t, is defined as $P_t(k|h)$ and given by

$$P_t(k|h) = \frac{P_t(h|k)P_t(k)}{P_t(h)} = \frac{\int_{z^\star} q_t(z'|k)dz'}{\int_{z^\star} q_t(z')dz'} . \left(\frac{\theta_t^{-k}}{k!} e^{-1/\theta_t} \right) .$$

Recall that embedded in the measures θ_t and firm filling probabilities $q_t(\cdot)$ are only the workers willing to take the job. This therefore omits those who find a better quality match in another firm and those currently employed in better matches. The measure of the mass of applicants used in the model however considers both these groups and is defined as

$$\text{Applicants per opening} = \frac{1}{v_t} \left(\alpha_u \lambda_u u_t + \alpha_e \lambda_e \int e_{t+}(z)dz \right) .$$

Table 8 displays a large disparity in the existing empirical literature regarding the number of applicants and the number interviewed per job. Further, as discussed there does not exist a perfect theoretical counterpart to the latter statistic. This is the motivation for not specifically targeting these moments in estimation.

7 Conclusion

This paper develops a model that incorporates five facts from empirical evidence regarding the search behavior of workers and recruiting behavior of firms. They are: firms always screen; typically a job opening receives many applicants; screening costs constitute the majority of hiring costs; in comparison to the employed the unemployed apply more frequently to jobs, but to on average worse jobs. Embedding these features into an otherwise standard labor search model yields a number of implications. The simulated model simultaneously generates a downward sloping Beveridge curve with cross-state employment dynamics. Further, the model generates a third of the observed shift in the Beveridge curve following the great recession and can rationalize the perceived volatility in matching efficiency.

References

Abramowitz, M. and I. Stegun, "Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables," *9th edition, Dover Publications* (1970).

Albrecht, J., P. Gautier and S. Vroman, "Equilibrium Directed Search with Multiple Applications," *Review of Economic Studies* 73 (2006), 869–891.

Andrews, M., J. Fu and R. Upward, "Meeting and matching: New evidence on search in the labour market," *mimeo* (2017).

Arbex, M., D. O'Dea and D. Wiczer, "Network Search: Climbing the Job Ladder Faster," *International Economic Review* 60 (2019), 693–720.

Barnichon, R. and A. Figura, "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics* 7 (2015), 222–249.

Barnichon, R. and Y. Zylberberg, "Underemployment and the Trickle-Down of Unemployment," *American Economic Journal: Macroeconomics* 11 (2019), 40–78.

Binmore, K., A. Rubinstein and A. Wolinsky, "The Nash bargaining solution in economic modelling," *RAND Journal of Economics* 17 (1986), 176–188.

Blatter, M., S. Muehlemann and S. Schenker, "The costs of hiring skilled workers," *European Economic Review* 56 (2012), 20–35.

Borowczyk-Martins, D., G. Jolivet and F. Postel-Vinay, "Accounting for Endogeneity in Matching Function Estimation," *Review of Economic Dynamics* 16 (2013), 440–451.

Bradley, J. and A. Gottfries, "A Job Ladder Model with Stochastic Employment Opportunities," *Quantitative Economics* 12 (2021), 1399–1430.

Burdett, K., S. Shi and R. Wright, "Pricing and Matching with Frictions," *Journal of Political Economy* 109 (2001), 1060–1085.

Christiano, L., M. Eichenbaum and M. Trabandt, "Unemployment and Business Cycles," *Econometrica* 84 (2016), 1523–1569.

Coles, M. and D. Mortensen, "Equilibrium Labor Turnover, Firm Growth, and Unemployment," *Econometrica* 84 (2016), 347–363.

Davis, S. and B. Samaniego de la Parra, "Application Flows," *mimeo* (2017).

Delacroix, A. and S. Shi, "Directed Search on the Job and the Wage Ladder," *International Economic Review* 47 (2006), 651–699.

Eeckhout, J. and I. Lindenlaub, "Unemployment Cycles," *American Economic Journal: Macroeconomics* 11 (2019), 175–234.

Elsby, M. and A. Gottfries, "Firm dynamics, on-the-job search and labor market fluctuations," *mimeo* (2019).

Engbom, N., "Contagious Unemployment," *mimeo* (2021).

Faberman, R. J., A. Mueller, A. Sahin and G. Topa, "Job Search Behavior among the Employed and Non-Employed," *NBER Working Paper No. 23731* (2017).

Galenianos, M. and P. Kircher, "Directed Search with Multiple Job Applications," *Journal of Economic Theory* 144 (2009), 445–471.

Gavazza, A., S. Mongey and G. Violante, "Aggregate Recruiting Intensity," *American Economic Review* 108 (2018), 2088–2127.

Gottfries, A., "Bargaining with renegotiation in models with on-the-job search," *mimeo* (2019).

Hall, R. and P. Milgrom, "The limited influence of unemployment on the wage bargain," *American Economic Review* 98 (2008), 1653–1674.

Jolivet, G., F. Postel-Vinay and J.-M. Robin, "The Empirical Content of the Job Search Model: Labor Mobility and Wage Dispersion in Europe and the U.S.," *European Economic Review* 50 (2006), 877–907.

Kircher, P., "Efficiency of Simultaneous Search," *Journal of Political Economy* 117 (2009), 861–913.

Krause, M. and T. Lubik, "On-the-job search and the cyclical dynamics of the labor market," *ECB Working Paper Series* (2007).

Krolikowski, P., "Job Heterogeneity and Aggregate Labor Market Fluctuations," *mimeo* (2019).

Krusell, P. and A. Smith, "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy* 106 (1998).

Lentz, R., J. Maibom and E. Moen, "Competitive or Random Search?," *mimeo* (2019).

Lise, J. and J.-M. Robin, "The Macrodynamics of Sorting between Workers and Firms," *American Economic Review* 107 (2017), 1104–1135.

Manning, A., "Imperfect Competition in Labour Markets," *Handbook of Labor Economics* 4 (2011), 973–1041.

Marinescu, I. and R. Wolthoff, "Opening the Black Box of the Matching Function: the Power of Words," *Journal of Labor Economics* 38 (2020), 535–568.

Menzio, G. and S. Shi, "Directed Search on the Job, Heterogeneity, and Aggregate Fluctuations," *American Economic Review* 100 (2010), 327–332.

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

Moen, E., "Education, Ranking, and Competition for Jobs," *Journal of Labor Economics* 17 (1999), 694–723.

Moscarini, G. and F. Postel-Vinay, "Stochastic Search Equilibrium," *Review of Economic Studies* 80 (2013), 1545–1581.

———, "Did the Job Ladder Fail After the Great Recession?," *Journal of Labor Economics* 34 (2016), 55–93.

———, "On the Job Search and Business Cycles," *mimeo* (2018).

Nagypal, E., "Worker reallocation over the business cycle: The importance of employer-to-employer transitions," *mimeo* (2008).

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

Pissarides, C., "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica* 77 (2009), 1339–1369.

Robin, J.-M., "On The Dynamics of Unemployment and Wage Distributions," *Econometrica* 79 (2011), 1327–1355.

Sahin, A., J. Song, G. Topa and G. Violante, "Mismatch Unemployment," *American Economic Review* 104 (2014), 3529–3564.

Sedláček, P., "Match Efficiency and Firms' Hiring Standards," *Journal of Monetary Economics* 62 (2014), 123–133.

Shi, S., "Frictional Assignment I: Efficiency," *Journal of Economic Theory* 98 (2001), 232–260.

———, "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology," *Review of Economic Studies* 69 (2002), 467–491.

———, "Wage Differentials, Discrimination and Efficiency," *European Economic Review* 50 (2006), 849–875.

Shimer, R., "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy* 113 (2005a), 996–1025.

———, “The Cyclical Behavior of Equilibrium Unemployment and Vacancies,” *American Economic Review* 95 (2005b), 25–49.

———, “On-the-job search and strategic bargaining,” *European Economic Review* 4 (2006), 811–830.

Silva, J. I. and M. Toledo, “Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment,” *Macroeconomic Dynamics* 13 (2009), 76–96.

Tauchen, G., “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions,” *Economic Letters* 20 (1986), 177–181.

Villena-Roldan, B., “Aggregate Implications of Employer Search and Recruiting Selection,” *mimeo* (2012).

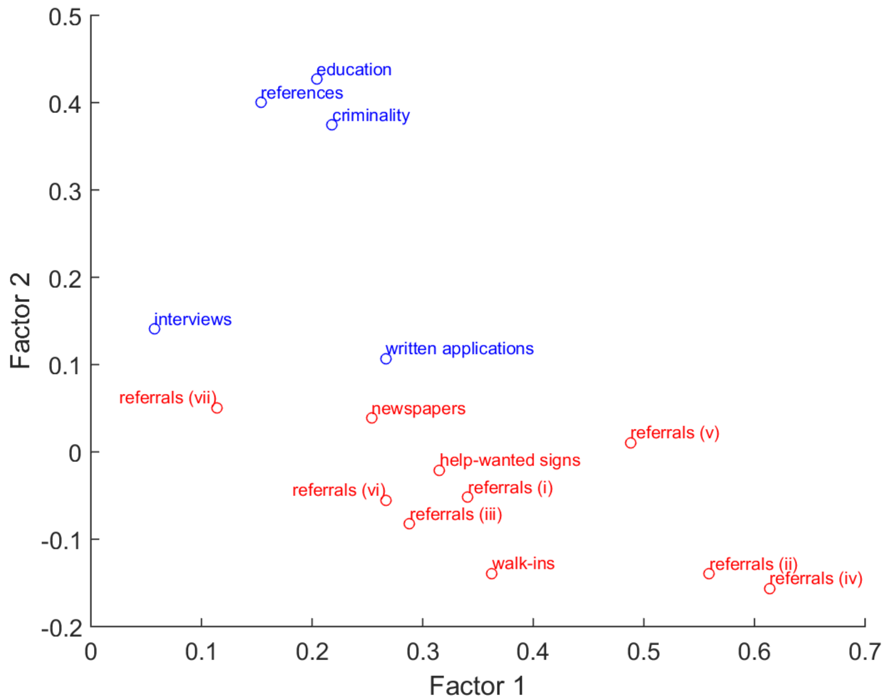
Walentin, K. and A. Westermark, “Learning on the Job and the Cost of Business Cycles,” *mimeo* (2018).

Wolthoff, R., “Applications and Interviews: Firms’ Recruiting Decisions in a Frictional Labour Market,” *The Review of Economic Studies* 85 (2018), 1314–1351.

A Appendix

A.1 Factor Loadings of Principal Component Analysis

Figure A.1: Factor Loadings



Notes: Each point represents a factor loading of a dummy variable. The variables in blue are those that have been pre-categorised as screening variables. They take the value one if the respondents answered they 'always' perform this task in recruitment for the latest recruit hired, and zero otherwise. They were asked how frequently they check: written applications; personal interviews; references; education/training; criminal records. For the recruitment variables in red the respondent was asked whether he administered a specific recruiting practice. The practices are: use of help-wanted signs; newspaper advertisement; consider walk-ins without referrals; ask for referrals from current employees (i); ask for referrals from the state employment service (ii); ask for referrals from a private employment service (iii); ask for referrals from a community agency (iv); ask for referrals from schools (v); ask for referrals from a union (vi); ask for referrals from any other service (acquaintances etc.) (vii).

A.2 Poisson Distribution of Applicants

$$\begin{aligned}
\text{bin}(a, \frac{1}{v}) = P(X = k) &= \binom{a}{k} \left(\frac{1}{v}\right)^k \left(1 - \frac{1}{v}\right)^{a-k} \\
\text{let } \lambda := \frac{a}{v} &\text{ and } a \to \infty \text{ therefore} \\
P(X = k) &= \lim_{a\to\infty} \left\langle \frac{a!}{k!(a-k)!} \left(\frac{\lambda}{a}\right)^k \left(1 - \frac{\lambda}{a}\right)^{a-k} \right\rangle \\
&= \frac{\lambda^k}{k!} \lim_{a\to\infty} \left\langle \underbrace{\frac{a!}{(a-k)!} \left(\frac{1}{a^k}\right)}_{\text{Term 1}} \underbrace{\left(1 - \frac{\lambda}{a}\right)^a}_{\text{Term 2}} \underbrace{\left(1 - \frac{\lambda}{a}\right)^{-k}}_{\text{Term 3}} \right\rangle
\end{aligned}$$

Taking the limit of each term in turn.

$$\begin{aligned}
\lim_{a\to\infty} \left\langle \frac{a!}{(a-k)!} \left(\frac{1}{a^k}\right) \right\rangle &= \lim_{a\to\infty} \left\langle \frac{a(a-1)...(a-k)(a-k-1)...(1)}{(a-k)(a-k-1)...(1)} \left(\frac{1}{a^k}\right) \right\rangle \\
&= \lim_{a\to\infty} \left\langle \frac{a(a-1)...(a-k+1)}{a^k} \right\rangle \\
&= \lim_{a\to\infty} \left\langle \left(\frac{a}{a}\right) \left(\frac{a-1}{a}\right) ... \left(\frac{a-k+1)}{a}\right) \right\rangle \\
&= 1
\end{aligned}$$

Turning to term 2,

$$\begin{aligned}
\lim_{a\to\infty} \left\langle \left(1 - \frac{\lambda}{a}\right)^a \right\rangle &= \lim_{a\to\infty} \left\langle \left(1 + \frac{1}{x}\right)^{-\lambda x} \right\rangle \quad \text{where } x = \frac{-a}{\lambda} \\
&= e^{-\lambda}
\end{aligned}$$

The limit of the third term is one, so

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

A.3 Solving the Model

To solve the model the support of z, γ, u, m_1 and m_2 are discretized. The size of the grids are lengths 50 for productivity, 20 for values of γ and unemployment and the mean and standard deviation of productivity are all length 10. Hence the model is solved for $50 \times 20 \times 10^3 = 1,000,000$ state variable combinations. The grid for z is from $z^\star$ up to $-\frac{1}{\lambda_z} \log(-\frac{1}{\lambda_e} \log(1 - \epsilon))$. The maximum is the value of z for which the employed who actively search will draw a productivity $Z > z$ with probability equal to ϵ. The value ϵ is set to 1e-8. It is important to set the upper bound of the z grid sufficiently high to avoid bunching at the top of the distribution. The grid is then equi-log spaced. The grid of γ is set according to the method proposed by Tauchen (1986) and is equally spaced between $\frac{\gamma_0}{1-\rho} - 3\sqrt{\frac{\sigma_\gamma^2}{1-\rho^2}}$ and $\frac{\gamma_0}{1-\rho} + 3\sqrt{\frac{\sigma_\gamma^2}{1-\rho^2}}$. A sensible grid for u, m_1 and m_2 is more difficult to define and the issue will be returned to in section A.3. For the rest of this section, it is assumed all grids have been defined. For brevity of notation, a state of the world $\mathcal{S} := (u, m_1, m_2) \in \mathbb{R}^3$ is defined.

1. Define distribution of match quality $\tilde{e}(z; \mathcal{S}, z^\star)$ given the aggregate state of the economy as defined by (12).

2. Approximate equation (3) as

$$e_+(z; \gamma, \mathcal{S}) = (1-\gamma)\tilde{e}(z; \mathcal{S}, z^\star) + \gamma \int \tilde{e}(z'; \mathcal{S}, z^\star) g_t(z|z') dz' \quad \text{for, } z > z^\star.$$

3. Define the number of active job seekers as a function of the state variables (γ, u, m_1, m_2) as

$$a(\gamma, \mathcal{S}) := \alpha_u \left(1 - \exp(-\lambda_u)\right) u + \alpha_e \left(1 - \exp(-\lambda_e)\right) \int e_+(z; \gamma, \mathcal{S}) dz.$$

4. Compute the probability a firm hires a worker of type z given they receive k applicants, equation (5). The grid for k takes every integer value from 1 to $K = 20$.

$$\begin{aligned} &q(z|\gamma, \mathcal{S}, k) = 0 && \text{for, } z < z^\star \\ &q(z|\gamma, \mathcal{S}, k) = k\left(1 - \frac{u}{a(\gamma, \mathcal{S})}\int_z g_u(z')dz' - \frac{\int e_+(z''; \gamma, \mathcal{S})dz''}{a(\gamma, \mathcal{S})}\int_z g_e(z')dz'\right)^{k-1} \\ &\quad \frac{1}{a(\gamma, \mathcal{S})}\left(u g_u(z) + \int^z e_+(z'; \gamma, \mathcal{S})dz' g_e(z)\right) && \text{for, } z \geq z^\star. \end{aligned}$$

5. Make a guess at policy function $v(\gamma, \mathcal{S})$ called $v_0(\gamma, \mathcal{S})$ that determines the number of vacancies posted at state (γ, u, m_1, m_2).

   6. Compute the probability a firm hires a type z worker as

   $$q(z; \gamma, \mathcal{S}) = \sum_{k=0}^{K}\left(\left(\frac{v_0(\gamma, \mathcal{S})}{a(\gamma, \mathcal{S})}\right)^{-k} \exp\left(-\frac{a(\gamma, \mathcal{S})}{v_0(\gamma, \mathcal{S})}\right) q(z|\gamma, \mathcal{S}, k)\right).$$

   7. Given $q(z; \gamma, \mathcal{S})$, the matching rates of workers can be computed, as in equations (8) and

(9).

$$\begin{aligned}
m(z;\gamma,\mathcal{S}) =& q(z;\gamma,\mathcal{S})v_0(\gamma,\mathcal{S}) \\
m(z,\varnothing;\gamma,\mathcal{S}) =& m(z;\gamma,\mathcal{S}).\frac{ug_u(z)}{ug_u(z)+\int^z e_+(z';\gamma,\mathcal{S})dz'g_e(z)} \\
m(z,z';\gamma,\mathcal{S}) =& m(z;\gamma,\mathcal{S}).\frac{e_+(z';\gamma,\mathcal{S})g_e(z)}{ug_u(z)+\int^z e_+(z';\gamma,\mathcal{S})dz'g_e(z)} \quad \text{for, } z'<z
\end{aligned}$$

8. Given the matching rates, next period's unemployment rate and distribution of match quality amongst the employed can be computed using equation (10) and (11).

$$\begin{aligned}
u'(\gamma,\mathcal{S}) =& u - \int m(z;\gamma,\mathcal{S})dz + \gamma\int^{z^\star}\int \tilde{e}(z';\mathcal{S},z^\star)g(z|z')dz'dz \\
e'(z;\gamma,\mathcal{S}) =& e_+(z;\gamma,\mathcal{S}) + \int m(z,z';\gamma,\mathcal{S})dz' - \int m(z',z;\gamma,\mathcal{S})dz'
\end{aligned}$$

9. Given next period's unemployment $u'(\gamma,\mathcal{S})$ and the distribution of the employed $e'(z;\gamma,\mathcal{S})$ a transition matrix $\pi_S(\mathcal{S}';\gamma,\mathcal{S})$ can be defined. This function tells us the state of the economy tomorrow (u',m_1',m_2'), given the state today (γ,u,m_1,m_2).
10. Compute the job finding rate of an employed worker in match quality z as

$$f(z;\gamma,\mathcal{S}) = \frac{\int m(z',z;\gamma,\mathcal{S})dz'}{\tilde{e}(z;\mathcal{S},z^\star)}.$$

11. Guess the value to a firm, of a match of quality z given the aggregate state variables, $\Pi_0(z,\gamma,\mathcal{S})$.
    12. This guess is updated, as in equation (4), as

$$\Pi_1(z,\gamma,\mathcal{S}) = (1-\beta)z - \omega_0 + \frac{1}{1+r}\int\Bigg[(1-\gamma)\left(1-f(z;\gamma,\mathcal{S})\right)\pi(\gamma',\gamma)\Pi_0(z,\gamma',\mathcal{S}')d\gamma' + \gamma\int\pi(\gamma',\gamma)\int_{z^\star}^{z} g(z'|z)\Pi_0(z',\gamma',\mathcal{S}')dz'd\gamma'\Bigg]$$

    13. A tolerance is computed as

$$\text{tol}_\Pi(z,\gamma,\mathcal{S}) = \text{abs}\left(\frac{\Pi_0(z,\gamma,\mathcal{S})-\Pi_1(z,\gamma,\mathcal{S})}{\Pi_0(z,\gamma,\mathcal{S})}\right)$$

    14. If the maximum of $\text{tol}_\Pi(z,\gamma,\mathcal{S})$, an element of dimension $50\times 20\times 10^3 = 1,000,000$, exceeds 0.01, set $\Pi_0(z,\gamma,\mathcal{S}) = \Pi_1(z,\gamma,\mathcal{S})$ and return to step 12. Otherwise, proceed to step 15.
15. Compute the value of posting a vacancy as defined in (7) as

$$V(v_0,\gamma,\mathcal{S}) = -\kappa + \int_{z^\star} q(z;\gamma,\mathcal{S})\left(\Pi_1(z,\gamma,\mathcal{S})-\sigma\right)dz.$$

16. Update $v_0(\gamma, \mathcal{S})$ by method of bisection and return to step 5, increasing $v(\cdot)$ if $V(\cdot) > 0$ and decreasing if $V(\cdot) < 0$.
17. Stop when find two vacancy levels $v_0(\gamma, \mathcal{S})$ and $v_1(\gamma, \mathcal{S})$ for which $V(v_0, \gamma, \mathcal{S}) > 0$, $V(v_1, \gamma, \mathcal{S}) < 0$ and for all values of γ, u, m_1 and m_2

$$\text{tol}_v(\gamma, \mathcal{S}) = \left(\frac{v_1(\gamma, \mathcal{S}) - v_0(\gamma, \mathcal{S})}{v_0(\gamma, \mathcal{S})}\right) < 0.01$$

A.3.1 Defining a grid for $\mathcal{S}$

Before the model is solved, two steady-state versions of the model are solved in which γ is fixed at $\gamma = \frac{\gamma_0}{1-\rho} \pm 3\sqrt{\frac{\sigma_\gamma^2}{1-\rho^2}}$. In this setting there is a unique unemployment rate and distribution of match quality in the economy. The solution algorithm is the same with two exceptions. Since, in this scenario $\mathcal{S}$ is stable there is no transition matrix to be computed and there is an outer loop where (u, m_1, m_2) are iterated over. Once solved, the two extreme cases for (u, m_1, m_2) are taken and the grid is equispaced between the two extremes.

A.4 Sampling Distribution

Recall, conditional on actively looking for a job, every period a worker of state $s \in \{u, e\}$ draws a new match productivity from the cdf

$$\tilde{G}_s(z) = \exp(-\lambda_s \exp(-\lambda_z z)) \quad \text{where } \lambda_s > 0 \text{ , } \lambda_z > 0 \text{ and } z \geq 0.$$

However, embedded in this distribution is the probability a worker receives no offers, which occurs with probability $\tilde{G}_s(0) = \exp(-\lambda_s)$. Thus, conditional on receiving an offer, the best productivity job follows the cdf

$$\tilde{G}_s(z|z > 0) = \frac{\exp(-\lambda_s \exp(-\lambda_z z)) - \exp(-\lambda_s)}{1 - \exp(-\lambda_s)} \qquad \text{where } \lambda_s > 0, \quad \lambda_z > 0 \quad \text{and} \quad z > 0.$$

As is pointed out in footnote 4 in the paper, this is a truncated Gumbel distribution with location parameter $\mu_s := \log(\lambda_s)/\lambda_z$ and scale parameter $\beta := 1/\lambda_z$.

$$\tilde{G}_s(z|z > 0) = \frac{\exp(-\exp(-(z - \mu_s)/\beta)) - \exp(-\exp(\mu_s/\beta))}{1 - \exp(-\exp(\mu_s/\beta))} \qquad \text{where } \mu_s \in \mathbb{R}, \quad \beta > 0 \quad \text{and} \quad z > 0.$$

The associated density function, on $x = \frac{z-\mu_s}{\beta}$, with s index dropped for brevity, is

$$\tilde{g}\left(x\right) = \frac{1}{\beta\left(1 - e^{-e^{\mu/\beta}}\right)}.e^{-\left(x+e^{-x}\right)} \qquad \text{where } \mu \in \mathbb{R}, \quad \beta > 0 \quad \text{and} \quad x > -\frac{\mu}{\beta}.$$

The associated moment generating function is given by

$$\begin{aligned} M_z(t) = \mathbb{E}(e^{tz}) &= \frac{1}{1 - e^{-e^{\mu/\beta}}} \int_0^\infty e^{tz} \frac{1}{\beta} e^{-\left(x+e^{-x}\right)} dz \\ &= \frac{1}{1 - e^{-e^{\mu/\beta}}} \int_{-\mu/\beta}^\infty e^{t(\beta x+\mu)} e^{-x} e^{-e^{-x}} dx. \end{aligned}$$

After a change of variables, letting $y = e^{-x}$, and hence, $x = -\log(y)$ and $dy = -e^{-x}dx$.

$$\begin{aligned} M_z(t) &= \frac{1}{1 - e^{-e^{\mu/\beta}}} \int_{e^{\mu/\beta}}^0 -e^{t\mu} e^{-t\beta \log(y)} e^{-y} dy \\ &= \frac{e^{t\mu}}{1 - e^{-e^{\mu/\beta}}} \underbrace{\int_0^{e^{\mu/\beta}} y^{-t\beta} e^{-y} dy}_{\text{the lower incomplete gamma function}} \\ &= \frac{e^{t\mu}}{1 - e^{-e^{\mu/\beta}}} \gamma(1 - t\beta, e^{\mu/\beta}) \end{aligned}$$

Where the lower incomplete gamma function given by

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

The first moment can be computed as

$$\begin{aligned} \frac{\partial M_z(t)}{\partial t}\Big|_{t=0} &= \left[\frac{e^{t\mu}}{1 - e^{-e^{\mu/\beta}}} \left(\mu\gamma(1 - t\beta, e^{\mu/\beta}) - \frac{\partial \gamma(1 - t\beta, e^{\mu/\beta})}{\partial t} \right) \right]_{t=0} \\ &= \mu - \frac{\beta}{1 - e^{-e^{\mu/\beta}}} \frac{\partial \gamma(\alpha, e^{\mu/\beta})}{\partial \alpha}\Big|_{\alpha=1}. \end{aligned}$$

Substituting back the expression for μ and β and returning the s subscript yields,

$$\begin{aligned} \overline{z}_s &= \frac{1}{\lambda_z} \left(\log(\lambda_s) - \frac{1}{1 - e^{-\lambda_s}} \frac{\partial \gamma(\alpha, \lambda_s)}{\partial \alpha}\Big|_{\alpha=1} \right) \\ &= \frac{1}{\lambda_z} \left(\log(\lambda_s) - \frac{\partial \log \gamma(\alpha, \lambda_s)}{\partial \alpha}\Big|_{\alpha=1} \right) \end{aligned}$$

Looking at the log derivative,

$$\frac{\partial \log \gamma(\alpha, \lambda_s)}{\partial \alpha}\Big|_{\alpha=1} = \left(\frac{1}{\gamma(\alpha, \lambda_s)} \frac{\partial \gamma(\alpha, \lambda_s)}{\partial \alpha} \right)_{\alpha=1}$$

The lower incomplete gamma function ($\gamma(\alpha, x)$) can be written as the ordinary gamma function ($\Gamma(\alpha)$) minus the upper incomplete gamma function ($\Gamma(\alpha, x)$), where

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt \quad , \quad \Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1}e^{-t}dt \quad \text{and} \quad \gamma(\alpha, x) = \int_0^x t^{\alpha-1}e^{-t}dt$$

Hence,

$$\frac{\partial \log \gamma(\alpha, \lambda_s)}{\partial \alpha}\Big|_{\alpha=1} = \left(\frac{1}{\gamma(\alpha, \lambda_s)} \left(\frac{\partial \Gamma(\alpha)}{\partial \alpha} - \frac{\partial \Gamma(\alpha, \lambda_s)}{\partial \alpha} \right) \right)_{\alpha=1}$$

Going through term by term, at $\alpha = 1$,

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

$$\left(\frac{\partial \Gamma(\alpha)}{\partial \alpha} \right)_{\alpha=1} = -\gamma \quad \text{where } \gamma \text{ is the the Euler-Mascheroni constant}$$

$$\left(\frac{\partial \Gamma(\alpha, \lambda_s)}{\partial \alpha} \right)_{\alpha=1} = E_1(\lambda_s) + e^{-\lambda_s}\log(\lambda_s) \quad \text{where } E_1(\lambda_s) \text{ is the exponential integral and defined}$$

in footnote 8. The identity can be found on page 262 of Abramowitz and Stegun (1970).

Substituting into the expression for the mean productivity sampled in employment state s yields

$$\overline{z}_s = \frac{1}{\lambda_z(1 - e^{-\lambda_s})} \left(\log(\lambda_s) + E_1(\lambda_s) + \gamma \right).$$

A.5 Fit of Model — absent screening

Table 9: Fit of the Model

Empirical Moments

	s	f_0	f_1	u	v	θ
Mean	0.014	0.219	0.019	0.066	0.027	0.447
Standard dev.	0.002	0.048	0.003	0.019	0.004	0.17
Correlations						
f_0	-0.485					
f_1	-0.322	0.641				
u	-0.73	-0.668	0.86			
v	-0.913	0.651	0.876	-0.78		
θ	0.48	0.698	-0.745	-0.921	0.886	
Autocorrelation	0.666	0.893	0.824	0.993	0.929	0.979
Labor Share	Mean = 0.59					

Simulated Moments

	s	f_0	f_1	u	v	θ
Mean	0.017	0.211	0.017	0.081	0.027	0.335
Standard dev.	0.002	0.003	0×10^{-3}	0.01	0.001	0.039
Correlations						
f_0	-0.707					
f_1	-0.072	0.668				
u	-0.049	-0.714	-0.217			
v	-0.91	0.054	0.981	0.331		
θ	0.479	-0.204	0.337	-0.947	-0.063	
Autocorrelation	0.843	0.882	0.937	0.981	0.355	0.939
Labor Share	Mean = 0.574					

Data covers the U.S economy from 2001 to 2014, inclusive. Employment dynamics and the unemployment rate are computed from the CPS and vacancy rates from the JOLTS. All are computed at a monthly frequency. The labor share comes from the BLS and is reported at a quarterly frequency, it is logged and HP filtered as described in the main body.

A.6 Accuracy of Aggregation

To assess the validity of the restricted information set imposed on firms, a straightforward comparison is made. How close are the simulated labor dynamics in the model to ones that firms would forecast? A transition rate is defined as $x \in \{s, f_0, f_1\}$ and can represent the probability an employed worker transitions to unemployment (s) and vice-versa (f_0), and the probability an employed worker switches employer (f_1). In addition to the unemployment rate, these will also depend on the state variable $e_t(z)$. Hence, to evaluate the accuracy of the aggregation synthetic transition rates are computed, defined as $\tilde{x}$. These are computed assuming that the firm's approximation of the distribution of match quality amongst the employed is correct. In other words, that $e_t(z) = \tilde{e}(z; u, m_1, m_2, z^\star)$, where $\tilde{e}(\cdot)$ is defined by equation (12).

To evaluate the proximity of the objects x and $\tilde{x}$ a linear regression on simulated data is run, taking the form

$$\log(x_t) = \beta_x \log(\tilde{x}_t) + \epsilon_t.$$

Results are displayed in Table 10. For the approximation to be working well there are two metrics to be evaluated. Firstly, the coefficient β_x should be approximately equal to one. This indicates that there is little bias in the firms forecasting. A $\beta_x > 1$ for example, would imply that the transition rate in question is systematically higher than a firm's expectation. Secondly, the dispersion of the error component ϵ_t. This is a measure of the precision of the aggregation. Consequently, included in Table 10 is both the estimated coefficient and R^2, standard errors are omitted as they are rounded to zero to any meaningful significance level.

Table 10: Accuracy of the Aggregation

	s	f_0	f_1
Coefficient	0.999	1.000	1.006
R^2	1.000	1.000	0.940
MAE	0.006	0.000	0.029

MAE is the mean absolute error (in logs) and defined as $\frac{1}{T}\sum_{t=1}^{T} |\log(x_t) - \log(\tilde{x}_t)|$.

Overall, the aggregation appears to work well. A firm's approximation of the job finding rate of the unemployed is indistinguishable from the model's simulation. This is unsurprising since, other than through the level of vacancies, the distribution plays no direct role in determining this rate. The other two rates see small differences in their forecast and realization. For the separation rate, differences in the mass of workers at the infimum of the distribution will lead to errors in the forecast. For job-to-job transition differences across the whole support matter and hence why this prediction is the poorest. The mean absolute error provided in Table 10 can be interpreted as approximately the mean proportional error between the actual and predicted rate. Hence, on average, when forecasting next period's job-to-job transition rate a firm will be incorrect in the order of 3%.