Labor Market Selection and the Dynamics of a Recovery

By Lukas Mann*

Draft: August 30, 2023

This paper proposes selection in the labor market as a solution to the puzzle of slow and near-linear recoveries. A simple twist in the matching technology of an otherwise standard matching model delivers such selection and generates the same recovery unemployment dynamics as in the data. Early in the recovery, composition effects and separations depress job creation incentives and therefore job finding rates. As observed empirically, this effect becomes much stronger for less productive, unemployed workers who under slack markets often get outranked by their more productive, employed peers. As these workers struggle to find jobs, negative composition effects create a feedback loop that slows down the adjustment of unemployment back to steady state. The model is able to match the last 6 recovery processes in the US economy closely.
JEL: E24, J64, J82
Keywords: Unemployment, recoveries, selection, heterogeneity, composition effects

Over the past 50 years, recessions in the US labor market have been characterized by a recurrent pattern: Unemployment rises within a short span of time, only to follow a near-linear downward recovery path. As pointed out recently by Hall and Kudlyak (2022), the shape of these recovery paths present a puzzle to the literature studying unemployment fluctuations. This is because, under plausible calibrations, the pure convergence dynamics of a system with stable transition rates predict an almost instantaneous and exponentially shaped recovery of unemployment back to its steady state level once the initial shock has subsided. Standard matching models typically fail to generate realistic recovery-dynamics, because an increase in the number of unemployed workers substantially improves matching prospects for firms, which increases the number of vacancies too much relative to what is observed empirically.

In this paper, I develop an equilibrium model of labor market selection that is able to generate realistic unemployment dynamics during a recovery. In the

* Department of Economics, Princeton University. Email: lmann@priceton.edu. I thank Gianluca Violante for his constructive comments, support and encouragement. I also thank John Grigsby, Gregor Jarosch, Richard Rogerson, Kieran Larkin, María Alexandra Castellanos, and various seminar participants at Princeton University, the University of Essex, Royal Holloway University of London, the University of Bristol, QMUL, the Econometric Society Summer Meetings, and Yale University for their constructive feedback.

model, the relatively slow adjustment of unemployment is the outcome of a feedback mechanism. Two assumptions deliver this mechanism: First, workers are ex-ante heterogeneous. Second, they compete with each other for the same jobs. The correct recovery dynamics then follow naturally: Early on, labor markets are slack. Under slack markets, more applicants compete for the same vacancy and low-productivity workers are less likely to match successfully. This skews the pool of job searchers towards lower-quality workers, in turn reducing the incentive to post vacancies and keeping markets slack. As a result, the recovery drags on and the unemployment rate converges to steady state very slowly.

In order to model the selection process, this paper introduces a new matching mechanism based on ranking different workers who compete for the same job. It extends the famous matching framework of Diamond, Mortensen and Pissarides (DMP) to allow for more than one match per vacancy and per searcher. In this matching framework, an additional congestion force is at play: In the DMP model, after a negative shock, a decrease in the value of a job is compensated by higher matching rates per vacancy. In the model presented here, during a recession, firms face a higher number of encounters per vacancy. These additional encounters do not substantially increase the number of matches but do compensate the firm by allowing it to be more selective about its applicant pool, so that in equilibrium the benefit of a vacancy continues to equal its cost. Since the rate of successful matches per vacancy does not increase substantially at the beginning of the recovery, the recovery is slowed relative to a recovery generated by traditional matching models.

When taken to the data, the model captures the empirically observed slow speed of many recoveries. Once transition rates into and between non-participation and unemployment are taken as given, the shape of the recovery path of unemployment is almost perfectly replicated by the model. Applying the model to the recoveries from the six most recent recessions, I find that the model captures the most salient features of post-recession transition rates into employment and therefore predicts an unemployment recovery path that is very close to the near-linear pattern observed in the data.

The work presented here relates to several different strands of literature. The question of sluggish post-recession unemployment adjustment has received attention by several papers, such as Hall and Kudlyak (2022) and Gregory, Menzio and Wiczer (2021). Like me, Gregory, Menzio and Wiczer (2021) employ a model with heterogeneous workers to predict the behavior of the unemployment rate. While their focus lies on an endogenous separation margin, I tie separations directly to the data and focus on the job finding margin during multiple recoveries, highlighting the general equilibrium feedback loop created by selection. Hall and Kudlyak (2022) discuss a wide variety of mechanisms that have the potential to explain slow and near-linear recoveries. Among them, composition-based effects stand out as one of the leading approaches to achieve model outcomes that match the data. This idea is somewhat controversial in the literature and there is an active debate on the cylicality of the composition of the unemployed (see Barnichon and Figura (2015) and Mueller (2017) for two papers on opposite sides of this issue). The model in this paper predicts that recoveries are primarily slowed through composition effects, but also shows that under selection, not much cyclicality in the average quality of the unemployment pool is needed to obtain strong composition effects.

One of the key challenges in creating realistic recovery outcomes is to make the unemployment rate sufficiently detached from its long-term steady state. Therefore, this paper also relates more broadly to the large literature studying amplifiers of unemployment fluctuations, initiated by Shimer (2005) and Hagedorn and Manovskii (2008) and more recently revisited by Eeckhout and Lindenlaub (2019), Mercan, Schoefer and Sedláček (2021), and Hall (2017), among others. Pries (2008) investigates the potential of worker heterogeneity to generate persistent unemployment fluctuations and finds that even under worker heterogeneity quantitatively much remains to be explained. A contribution of this paper is to show that worker heterogeneity can fully account for persistent unemployment dynamics once selection is also taken into account. In addition, my model matches not only the persistence but also the shape of recoveries. The closest paper to this idea is perhaps Ferraro (2018), which studies the skewness of the long-run unemployment distribution in addition to its variance, although the focus is on matching long-run aggregate statistics rather than individual recoveries. The model in Ferraro (2018) also features a form of selection - some workers simply become unemployable in some states. In contrast, my way of modeling selection offers an arguably more realistic description and micro-foundation of individual job finding rates as a function of worker quality and the aggregate state.

The ranking model introduced here is similar in spirit to Blanchard and Diamond (1994), who like me study a model in which vacancies meet with several workers, choosing the one with the lowest recorded unemployment duration. Their paper, however, emphasizes the role of ranking for wages, not for the dynamics of unemployment. In fact, the ranking assumption in their model has no effect on the behavior of aggregate unemployment and only carries implications for wages and individual unemployment duration. In contrast, ranking as conceptualized in this paper has major allocative consequences: Because of selection, hiring shifts away from the unemployed and towards the employed. Firms facing lower job values are compensated in match quality rather than matching probability, further slowing recovery. In allowing for multiple encounters per vacancy and per worker, my paper also contributes to attempts at modelling a more realistic application process in job search, such as Wolthoff (2018) and Birinci, See and Wee (2020).

The idea that the unemployed have on average more difficulty finding a job than the employed also relates to Engbom (2021) who studies a model in which the expected value of a match falls with a rise in the number of unemployed since they apply for less suitable jobs. The model presented here highlights that less success in job finding by the unemployed can be due to a different reason:

The unemployed simply tend to be less productive workers who consequently will be the top candidate for a position less frequently. In matching this feature of the data, the model is therefore able to generate endogenously the finding in Faberman et al. (2017) that the employed are typically more successful in their search than the unemployed.[1] In addition, in the model, the wage premium of the typical hire from employment (relative to a hire from unemployment) due to observable differences is close the value of 17 log points identified in their paper.

The rest of the paper proceeds as follows: Section I illustrates the puzzle of rapid and near-linear recoveries, section III outlines the model, section IV discusses the calibration and section V presents the results of the simulations, with a special emphasis on the fit of the unemployment rate after a recession. Section VI concludes.

I. The recovery puzzle

To see why a standard DMP model produces unrealistic recovery dynamics, consider the following equation describing continuous time unemployment dynamics in a broad class of labor market models with only two states, employment and unemployment:

$$\dot{u}_t = -\lambda u_t + \delta(1 - u_t) \tag{1}$$

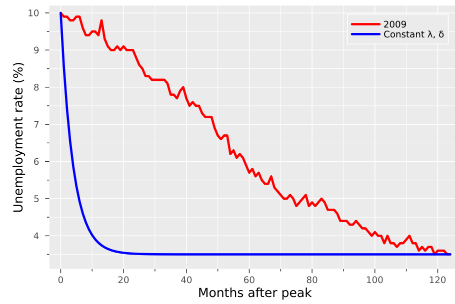
Here, λ is the rate at which the unemployed find a job, δ is the separation rate into unemployment, and u is the unemployment rate. In an environment where λ and δ have reverted back to steady state, such as a Diamond-Mortensen-Pissarides (DMP) model, at steady state productivity, this differential equation has the solution

$$u_t = e^{-(\lambda+\delta)t}\left(u_0 - \frac{\delta}{\lambda+\delta}\right) + \frac{\delta}{\lambda+\delta} \tag{2}$$

The half life of excess unemployment arising from this equation, a convenient expression for the speed at which excess unemployment dissipates, is $\frac{\log 2}{\delta+\lambda}$. Using monthly values of $\delta = 0.015$ and $\lambda = 0.25$, which are motivated by typical ranges found in CPS data, this half life is 2.6 months. Such a recovery is extremely short and incompatible with observed dynamics in the data. Figure 1 shows this. On the left hand panel, it displays the unemployment series for the 5 most recent pre-Covid recoveries in the data. The right hand panel displays a counterfactual recovery according to equation (2), in which separation and job finding rates are constant over the course of the recovery. As is evident from the figure, such a recovery is far too rapid to fit the data well.

[1]That is, as in their paper, the employed in receive more offers per meeting which can be interpreted as receiving more offers per application.

(a) Unemployment, 5 pre-Covid recoveries



(b) Dynamics of equation 2 (2009 recovery)

Figure 1.

The reason for the excess recovery speed in the standard model is the following: Holding productivity constant, in times of high unemployment, vacancies are likely to encounter many applicants, driving up the value of a vacancy. Therefore, vacancies are posted up to the point where the vacancy filling rate returns to steady state. This happens to also be the point at which the job finding rate is at steady state. Absent other factors, the recovery then follows equation 2.

Fixing this shortcoming of the workhorse matching model requires generating a mechanism by which the value of posting a vacancy is depressed until long into the recession, lowering job finding rates. I show in this paper that selection can play an amplifying role in translating generally slack markets into environments with a low exit rate from unemployment. This mechanism crucially hinges on the idea that less productive workers are comparatively more affected by slack markets. The next section provides empirical evidence that this is indeed the case.

II. Sensitivity of UE rates to the cycle

To motivate the premise that selection can play an important role in generating unemployment dynamics, I use data from the National Longitudinal Survey of Youth (NLSY) to show the following empirical fact: Workers who over their lifetime have a lower job finding probability, are much more exposed to the business cycle than those who have a high job finding probability. The NLSY follows a cohort of workers over their entire working life. I use the full representative sample of workers aged between 14 and 21 on December 31, 1978. This leaves me with 12.7 million weekly observations of 6075 workers between 1979 and 2018. I then bin workers into 5 quantiles q of life time job finding rates and calculate the average monthly job finding rate in every year for each of those groups. I run the

Table 1—: Sensitivity of job finding rates across UE prob. quantiles

UE Prob. Quantile (q)	1st	2nd	3rd	4th	5th
Coefficient (β_1)	-0.62	-0.43	-0.09	0.06	0.006
	(0.20)	(0.15)	(0.13)	(0.12)	(0.08)

Robust standard errors in parentheses.

following simple regression for each group:

$$\log \mathrm{UE}_t^q = \beta_0 + \beta_1 \log \mathrm{UR}_t + \gamma_1 t + \gamma_2 t^2 + \varepsilon_t^q$$

where UR_t is the average monthly unemployment rate (from the BLS) during a given year and I fit a second order polynomial to absorb age effects on the job finding rate.

Table 1 collects the coefficient β_1 for each quantile q. It is obvious from the table that the lower quantiles are much more exposed to the business cycle - these workers' job finding rates drop much more during times when unemployment is high. Simply put, workers who systematically have trouble finding jobs suffer much more from recessions than workers who find jobs easily. This finding motivates the matching mechanism of the model which has the same implication - when unemployment rises, job finding rates decline a lot for workers who already have more difficulty in gaining employment. In the model, this happens because low productivity workers are more likely to be outranked by high productivity competitors. Crucially, this effect becomes much stronger under slack markets, since slack market environments involve more competition of workers for jobs. This feeds back into adverse composition effects during the recovery phase that dampen job creation.

III. Model

The model is a three-state-model - a somewhat less common choice in the literature, but useful for two reasons. First, omitting non-participation as a state inhibits a realistic treatment of composition effects. Second, the data features a non-negligible number of transitions to and from non-participation. Such abstraction necessarily comes at a cost when trying to model realistically the value of a job to the firm. Any model abstracting from non-participation (i.e. any model with two employment states) can feature one of two calibration choices: One possibility is to estimate the separation rate in a range that matches the EU transition probability in the data. Such models would realistically capture unemployment rate inflows but overstate the duration of employment compared to the data and therefore make the job value more sensitive to changes in the effective discount factor (which is almost entirely determined by separation rates). Alternatively, the model could feature realistic discount factors and employment durations by choosing the transition rate into unemployment and non-participation as the benchmark separation rate. Those models, however, would strongly overstate the inflows into unemployment and therefore the turnover rate of the unemployed population, implying a higher job finding rate than in the data. For these reasons, I opt for a three-state-model in which workers can be employed, unemployed or non-participating. I set this model in discrete time in order to facilitate comparison with monthly data.

The effects of worker heterogeneity lie at the core of my mechanism so I consider three dimensions of worker heterogeneity - productivity, and separation probability by successive status (EU and EN). The model abstracts from the effects of firm heterogeneity which I leave to future research. The following section sets up this model.

A. Environment

Workers: The economy is populated by heterogeneous workers indexed by their type $i \in \{1, \cdots, I\}$ and distributed across types with probability measure μ_i. At the end of any period, a worker is in one of three states, non-participating, unemployed or employed. Every worker type is characterized by a tuple (y_i, r_i, d_i^u, d_i^n), where y_i is their productivity, $r_i = i$ is their type rank and d_i^u, d_i^n are their relative (to the aggregate) transition rates into unemployment and non-participation respectively. I estimate y_i from wage data under a Nash bargaining assumption, and d_i^u, d_i^n from transition rates. In the estimation procedure, explained in more detail in section IV, worker types are grouped on observables. The type rank of a worker determines whether the order in which they are chosen when they compete for a vacancy. I assume that this ranking is exogenous and corresponds to the steady state wage of a worker type and motivate this assumption in section IV.

Timing: Time is discrete and infinite. Each period consists of three stages. First, some random selection of employed and unemployed workers receives a non-participation shock, which will push them out of the labor force at the beginning of the next period. The probability of this shock for a given worker depends on the worker type as specified below. Second, among workers who do not receive this shock, a random selection receives an unemployment shock which will make them unemployed at the beginning of the next period. In the third stage, those workers who are not hit by any of these shocks participate in a matching mechanism. I refer to the third stage as the hiring or matching stage. After it has passed, all transitions conclude and the number of employed, non-participating and unemployed workers is measured. Figure 2 illustrates this timing assumption.

During the first stage, each employed worker of type i transitions into non-participation with probability $\delta_t^{en,i} = d_i^n \delta_t^{en}$. Each unemployed worker transitions into non-participation with probability δ_t^{un}. During stage two, non-participating and employed workers transition into unemployment with probabilities δ_t^{nu} and $\delta_t^{eu,i} = d_i^u \delta_t^{eu} / (1 - d_i^n \delta_t^{en})$ respectively (this makes the unconditional probability of

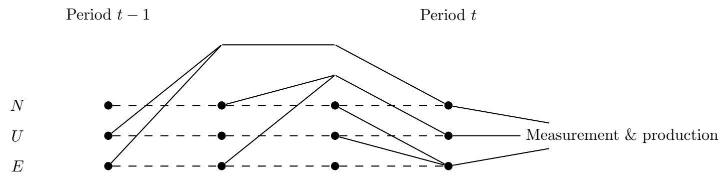


Figure 2. : Timing assumptions

a flow into unemployment for an employed worker at the beginning of the period equal to $d_i^u \delta_t^{eu}$). Finally, during the matching stage, conditional on the employment state, workers become searchers with probability 1 (if unemployed prior to the matching stage), s_e (if employed) or s_n (if non-participating). As a searcher, they partake in the matching mechanism described in section III.B. If matched, they transition into employment, otherwise they retain their employment status.

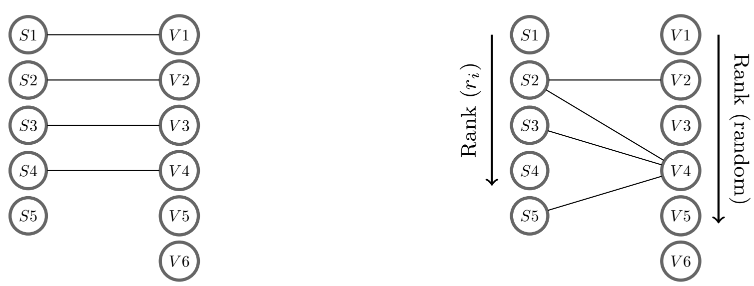
Firms: The firm side of the model is simple and follows the DMP tradition, in particular by imposing a free entry assumption. There is a measure of atomistic and homogeneous firms. In each period, during the hiring stage, firms can decide to post an arbitrary number of vacancies at cost κ. The vacancy is then subject to the matching process described in section III.B that determines whether and with what kind of worker it is filled. If successfully filled, production begins immediately after the hiring stage and continues for every period after the hiring stage until the worker stochastically separates into unemployment or non-participation.

Bargaining: Once matched, wages are set according to Nash bargaining. I assume that the outside option for any worker is always the same, irrespective of the matched searchers previous employment status. I assume that for any worker, their outside option is to follow the transition probabilities into non-participation or unemployment conditional on not remaining in employment. Importantly, I assume this for all matches, including those from employment. This assumption side-steps the problem of non-convex bargaining sets for models with Nash bargaining and on-the-job search raised in Shimer (2006).

B. Matching and selection

The matching stage is the heart and main innovation of the model. It features a ranked selection mechanism which favors the unemployment exit rate for workers of a high rank: During times of slack markets, hires tend to come from a larger pool of applicants, which favors better workers. This propagates a skewed composition of searchers (low quality workers do not get absorbed into employment) and therefore endogenously slows the recovery by discouraging vacancy creation. It also leads to hires out of employment over hires out of unemployment, particularly in times of slack markets, further strengthening this effect. The mechanism is based on a two-sided multiplicity of encounters - a searcher can encounter many different vacancies and a vacancy many different searchers. I outline the technical details of this mechanism below.

To set the stage for the matching mechanism, consider a matching environment with M matches, V open vacancies and $L = U + s_e \cdot E + s_n \cdot N$ job searchers where U, E and N are the number of workers in the unemployment, employment and non-participating states respectively and s_e (s_n) is the probability of search for an employed (non-participating) worker. Given some sufficiently small $\varepsilon > 0$, M, V and L correspond to natural numbers $n_M \in \mathbb{N}$, $n_V \in \mathbb{N}$ and $n_L \in \mathbb{N}$ in the following way: Let n_M, n_V and n_L be the smallest natural numbers satisfying $\varepsilon \cdot n_M \in B_\varepsilon(M), \varepsilon \cdot n_V \in B_\varepsilon(V), \varepsilon \cdot n_L \in B_\varepsilon(L)$. That is, the real numbers are a re-scaled stand-in for large natural numbers which approximately have the same ratio among each other as the original real numbers.[2] We are ultimately interested in the limit $\varepsilon \to 0$. However, this discrete representation allows us to think about matches, searchers and vacancies in a discrete environment. There are two natural assumptions one could make about the relationship between matches, searchers and vacancies. One could assume that matches are evenly assigned to searchers and vacancies, each vacancy and each searcher receiving at most one match. In this instance, if every match results in a job, each searcher has a job finding probability of $\frac{n_M}{n_L} \to \frac{M}{L}$ which corresponds to the matching mechanism from the baseline DMP model. In the model proposed here, I introduce a different assumption: Instead of being evenly distributed across searchers and vacancies, I assume that meetings[3] are *randomly* assigned to a vacancy and a searcher. Figure 3 illustrates the difference for the case $n_M = 4, n_V = 6, n_L = 5$.



(a) Standard matching (even assignment)

(b) Many-to-many (random assignment)

Figure 3. : Illustration of the matching mechanism with $n_M = 4, n_V = 6, n_L = 5$

In particular, under this assumption, any vacancy and any worker can end up

[2]Strictly speaking, n_M, n_V and n_L are functions of ε and M, V and L respectively, but for ease of notation this dependence is suppressed.

[3]In this environment, not all meetings can lead to successful matches. To streamline nomenclature, I shall refer to an encounter between worker and vacancy as a "meeting" or "encounter", and to a meeting that leads to a job as a "(successful) match".

with any number of meetings, at most one of which can result in a successful match. I therefore call this matching process many-to-many matching, which throughout the paper I contrast with the one-to-one matching assumption that is standard in the search literature. For any vacancy j, denote by m_j the (random) number of meetings assigned to that vacancy and denote $q_\varepsilon = \frac{n_M}{n_V} \to q = \frac{M}{V}$. For vacancy j, the probability of getting $m_j = k$ meetings is

$$\begin{aligned}
P(m_j = k) &= \left(1 - \frac{1}{n_V}\right)^{n_M - k} \left(\frac{1}{n_V}\right)^k \binom{n_M}{k} \\
&= \left(1 - \frac{1}{n_V}\right)^{n_M - k} \left(\frac{1}{n_V}\right)^k \frac{n_M!}{k!(n_M - k)!} \\
&= \left(\left(1 - \frac{1}{n_V}\right)^{n_V}\right)^{q_\varepsilon} \left(1 - \frac{1}{n_V}\right)^{-k} q_\varepsilon^k \left(\prod_{j=0}^{k} \left(1 - \frac{j}{n_M}\right)\right) \frac{1}{k!} \\
&\xrightarrow{\varepsilon \to 0} e^{-q} q^k \frac{1}{k!}
\end{aligned}$$

Hence, the limiting distribution of meeting frequency per vacancy is Poisson with parameter q. Employing an analogous argument, on the searchers' side, the number of meetings for a given searcher is Poisson-distributed with parameter λ. Assuming that each searcher can only take one job and that vacancies can be filled with at most one worker, the potential multiplicity of meetings per searcher and per vacancy forces us to commit to a selection mechanism by which only a subset of meetings become successful matches. I will now describe this selection mechanism.

There exists a common ranking of worker types, r_i, which is constant over time. Suppose that within type, workers are randomly ranked each period. For a given period this yields a ranking for every worker in the pool of searchers. Further suppose, that each period posted vacancies are randomly ranked[4]. Thus, every searcher can be characterized by their rank $p_L \in [0, 1]$ within the distribution of searchers. Likewise, every vacancy can be characterized by their (random) rank $p_V \in [0, 1]$. Now, assume that at the start of a matching stage all vacancies make offers in order of their rank p_V. Each vacancy starts out by making an offer to their encountered searcher with the highest rank. If this searcher has not previously received (and therefore accepted) an offer from a higher ranked vacancy, they accept the offer, otherwise they reject. In case of rejection the vacancy makes an offer to whoever has the second highest rank among their encounters, in case of rejection they make an offer to the third-ranked encounter and so on. In some cases, a vacancy might receive a positive number of meetings but not receive any successful match because the searchers they meet have all received better ranked offers.

[4]Since firms are assumed to be identical, a random ranking is natural in this setting

To describe this process by means of example, consider figure 3. Here, the highest-ranked vacancy (V1) has not encountered a searcher and therefore stays unmatched (the same happens to vacancies 3, 5 and 6). V2 then makes an offer to S2 and S2 accepts. Next, V4 makes an offer to S2, but S2 is already matched and therefore rejects, so V4 makes an offer to S3 which is accepted. S5 ends the matching phase without a match even though it has encountered V4, because S5 is outranked by S3. All remaining searchers who have not encountered a vacancy also remain unmatched.

It is easy to show that this successive way of resolving encounters is conditionally Pareto optimal, even if workers are assumed to be indifferent between vacancies. Intuitively, every encountered worker with a rank higher than the one the firm matches with has been matched with another firm. To find a Pareto improvement, at least one firm has to be matched with a higher ranked worker than in the original allocation. This means that another firm of higher rank will lose the match with this worker. To compensate this firm, a worker of even higher quality has to switch out of their old match into this firm and so on. At some point, it will be impossible to compensate the firm since they do not have a higher-ranked encounter.

We can now determine the probability of a match given the rank of a particular worker. For this, define

$$f(p_L, p_V) = P(p_L \text{ receives an offer from a vacancy of rank } p_V \text{ or higher})$$
$$g(p_L, p_V) = P(p_V \text{ receives no acceptance from a searcher of rank } p_L \text{ or higher})$$

For a given searcher, the number of meetings with firms greater than p_V is Poisson-distributed with parameter $\lambda(1 - p_V)$. Similarly, the number of meetings of the firm that feature a worker of rank greater than p_L is Poisson-distributed with parameter $q(1 - p_L)$. Let $\mu \sim \text{Pois}(\lambda(1 - p_V))$ be the number of meetings with vacancies ranked above p_V for a given searcher and $m \sim \text{Pois}(q(1 - p_L))$ be the number of meetings with searchers ranked above p_L for a given vacancy. The probability that a searcher p_L receives an offer from a vacancy with rank p_V or higher is equal to the probability that at least one of the searcher's meetings higher than p_V gets rejected by all their meetings greater than p_L. The probability that a vacancy p_V receives no acceptance from a searcher of rank p_L or higher is equal to the probability that all the vacancy's meetings of rank above p_L get rejected by the worker (because that worker received an offer higher than p_V). Thus,

$$f(p_L, p_V) = \mathbb{E}\left[1 - \left(1 - (1-p_V)^{-1}\int_{p_V}^{1} g(p_L, \tilde{p}_V)\, d\tilde{p}_V\right)^{\mu}\right]$$

$$g(p_L, p_V) = \mathbb{E}\left[\left((1-p_L)^{-1}\int_{p_L}^{1} f(\tilde{p}_L, p_V)\, d\tilde{p}_L\right)^{m}\right]$$

and therefore, using the fact that $\mathbb{E}\left[p^m\right] = \exp(\lambda(p-1))$ for any number p and Poisson-distributed random variable $m \sim \mathrm{Pois}(\lambda)$, we arrive at the following differential equation for f:

$$1 - f(p_L, p_V) = \exp\left(-\lambda \int_{p_V}^{1} \exp\left(-q\int_{p_L}^{1}\left[1 - f(\tilde{p}_L, \tilde{p}_V)\right]\, d\tilde{p}_L\right)\, d\tilde{p}_V\right) \tag{3}$$

It is straightforward to verify that this equation reduces to the edge case

$$f(1, s) = 1 - \exp\big(\lambda(p_V - 1)\big)$$

for $p_L = 1$. This is intuitive - a searcher who will be hired first by any firm will get a job offer better than p_V with exactly the probability of a meeting higher than p_V.

Differential equation 3 can be solved numerically, which yields solutions for f and g. An interesting object is $f(p_L, 0)$, which describes the job finding probability for a given worker p_L. Assuming a Cobb-Douglas production function for meetings $M = aL^{\omega}V^{1-\omega}$, q and λ are related through the meeting efficiency parameter a by the formula

$$\lambda_t = a^{\frac{1}{\omega}} q_t^{\frac{\omega-1}{\omega}} \tag{4}$$

This implies that for a given meeting efficiency a, it is possible to determine the job finding probability as a function of the rank within the distribution of searchers using a single scalar, the number of meetings per searcher. The meeting efficiency a plays an important role for the relationship between a searcher's job finding rate and their rank. For larger values of a the slope of the job finding curve will be steeper. Holding the average job finding probability constant, figure 4 shows how the the job finding probability as a function of the searcher rank varies with the chosen value of the meeting efficiency a. Intuitively, as a goes to zero, there are far more vacancies than meetings (recall that we are holding the average job finding rate constant). Therefore, meetings with a given vacancy are virtually guaranteed to be the vacancy's only contact, meaning that the proba-

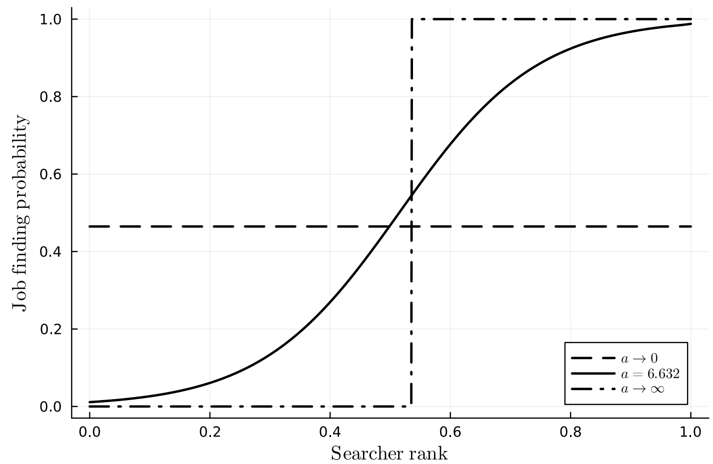


Figure 4. : 2009 steady state job finding probability by searcher rank for different a

bility of a match approximately equals the probability of a meeting, which is the same for all searchers. As a goes to infinity, all vacancies are connected to all searchers, meaning that vacancies can always find a match until there are either no vacancies or no searchers left. Holding the average job finding rate constant, this corresponds to the situation where all vacancies end up with a match, but only the highest ranked searchers do. The value for a used in the baseline calibration introduces a relationship between rank and job finding probability between these two extreme cases.

Figure 5 illustrates the relationship between λ and the job finding rate for a given searcher rank for the calibration of meeting efficiency used in my baseline calibration. As can be seen from the figure, a higher meeting efficiency means an effective shutdown of access to the job market for the lower percentiles of the searcher distribution. This is because in times of low job finding rates, holding the meeting efficiency constant, vacancies will on average get many applicants. This allows them to be more selective about who to hire which effectively means that the lower end of the job searcher distribution will be crowded out by higher ranked candidates, even if they find one or several meetings.

An alternative way to look at the dynamics of this matching framework is to view the offer and response mechanism as a process during which some of the meetings created by the meeting function get destroyed while others turn into successful matches. Let $\tilde{\lambda}(p_L) = f(p_L, 0)$ denote the job finding rate, i.e. the

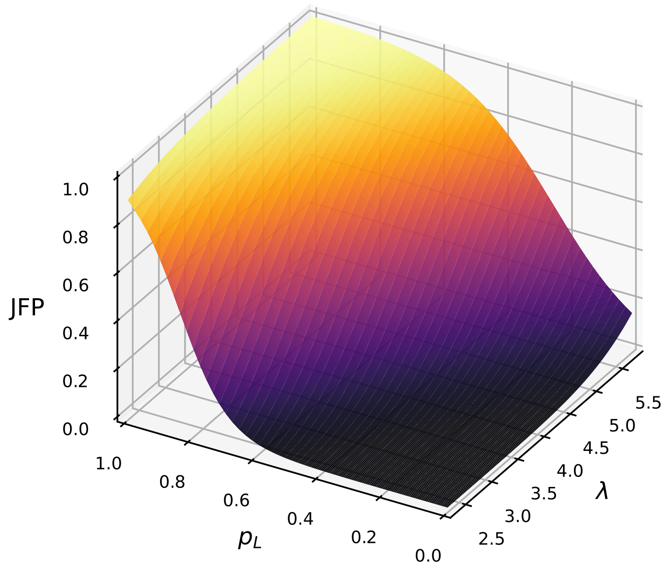


Figure 5. : Job finding probability as a function of λ and p_L for $a = 5.56$

number of successful matches per searcher at rank p_L. Now, define

$$\sigma(p_L) = \frac{\tilde{\lambda}(p_L)}{\lambda} = \frac{f(p_L, 0)}{\lambda} \tag{5}$$

The ratio of successful matches to total meetings at rank p_L is given by $\sigma(p_L)$ and is fully determined by λ (or, equivalently, q). Note that the model nests standard one-to-one matching - one can simply set $\sigma(p_L) \equiv 1$ to deliver the standard equations in this case. Denote by $J(p_L)$ the benefit of a match with a worker of rank p_L. Denote by $\bar{J}$ the average value of a meeting to the firm. For a firm with a random vacancy rank, the ex-ante density of successful match probability by rank is given by $\sigma(p_L)$ and the expected average value of a meeting in period t is given by

$$\bar{J}_t = \int_0^1 \sigma_t(p_L) J_t(p_L)\, dp_L \tag{6}$$

As in the DMP model, the cost of posting a vacancy has to equate the benefit. As the expected number of meetings per vacancy is q and the vacancy posting cost is κ, this leads to the well-known equilibrium condition

$$\kappa = q_t \bar{J}_t \tag{7}$$

However, one feature of this equation sets it apart from the canonical DMP condition: Here, as opposed to the DMP model, $\bar{J}$ *directly* depends on q through $\sigma(.)$. For higher values of q, σ becomes more skewed towards higher values as firms receive more applications per vacancy and therefore become more selective. Note however that, conditional on $J(p_L)$, there is exactly one value of λ (or, equivalently, market tightness or q) that solves this equation. In this sense, there is still a direct mapping between the value of a meeting and the job finding rate, as in the DMP model. This simplifies immensely the computation of equilibrium because all transition probabilities into employment can be computed from the initial distribution and the scalar λ_t. Hence the computation of equilibrium is achievable with a simple fixed-point search: One simply needs to find the sequence $\{\lambda_\tau\}_{\tau=0}^{\infty}$ that produces a sequence of distributions and transition probabilities that map into a series of meeting values $\bar{J}_t$ that through equations 7 and 4 in turn correspond to $\{\lambda_\tau\}_{\tau=0}^{\infty}$.

C. Laws of motion and value functions

With this matching mechanism specified, the model is almost fully defined. We can write down transition probabilities from period to period by defining the following transition matrix:

$$\Theta_t^i = \begin{pmatrix} (1-\delta_t^{nu})(1-s_n\sigma_t(p_L^t(i))\cdot\lambda_t) & \delta_t^{un} & \delta_t^{en,i} \\ \delta_t^{nu} & (1-\delta_t^{un})(1-\sigma_t(p_L^t(i))\cdot\lambda_t) & (1-\delta_t^{en,i})\delta_t^{eu,i} \\ (1-\delta_t^{nu})s_n\sigma_t(p_L^t(i))\cdot\lambda_t & (1-\delta_t^{un})\sigma_t(p_L^t(i)\cdot\lambda_t & (1-\delta_t^{en,i})(1-\delta_t^{eu,i}) \end{pmatrix}$$

This matrix enables us to conveniently write the value functions of a worker of type i in a non-participating, unemployed or employed state, denoted $V_t^{N,i}, V_t^{U,i}$ and $V_t^{E,i}$ respectively. Define $\mathcal{V}_t^i = (V_t^{N,i}, V_t^{U,i}, V_t^{E,i})'$. Then,

$$\mathcal{V}_t^i = (b, b, w_t^i)' + \frac{1}{1+r}\left(\Theta_{t+1}^i\right)'\mathcal{V}_{t+1}^i \tag{8}$$

Likewise, for the firm, denote by J_t^i the value of a successful match with worker i. This value is determined by the following Bellman equation:

(9)
$$J_t^i = y_i - w_t^i + \frac{1}{1+r}\left[(1-\delta_{t+1}^{en,i})(1-\delta_{t+1}^{eu,i})(1-s_e\cdot\sigma_{t+1}(p_L^{t+1}(i))\cdot\lambda_{t+1})\right]J_{t+1}^i$$

where $p_L^t(i)$ is the average rank of type i in period t.[5] Finally, because of Nash bargaining, the wage w_t^i is set in every period such that

$$J_t^i = (1-\beta)(J_t^i + V_t^{E,i} - \gamma_{t+1}^i V_t^{N,i} - (1-\gamma_{t+1}^i)V_t^{U,i}) \tag{10}$$

where $\gamma_t = \delta_t^{en,i}/(1-(1-\delta_t^{en,i})(1-\delta_t^{eu,i}))$ denotes the probability of separating into non-employment in period $t+1$ conditional on separating in period $t+1$.

Next, define $\mathcal{E}_t^i = (N_t(i), U_t(i), E_t(i))'$. The law of motion of the type-state distribution is

$$\mathcal{E}_t^i = \Theta_t^i \mathcal{E}_{t-1}^i \tag{11}$$

To determine the expected value of a meeting during a particular period, it is useful to change the integration measure of equation 6:

$$\begin{aligned}\bar{J}_t &= \left(\int_0^1 \sigma_t(\tilde{p}_L)\, d\tilde{p}_L\right) J_t \\ &= \underbrace{\left(\int_0^1 \sigma_t(\tilde{p}_L)\, d\tilde{p}_L\right)}_{(1)} \int_0^1 \underbrace{\left(\frac{U_t^-(i) + s_n N_t^-(i) + s_e E_t^-(i)}{\int U_t^-(i) + s_n N_t^-(i) + s_e E_t^-(i)\, d\tilde{\mu}_i}\right)}_{(2)} \underbrace{\frac{\sigma_t(p_L^t(i))}{\int_0^1 \sigma_t(\tilde{p}_L)\, d\tilde{p}_L}}_{(3)} \underbrace{J_t^i}_{(4)}\, d\mu_i\end{aligned} \tag{12}$$

where J_t denotes the average value of a match, $\bar{J}_t$ denotes the average value of a meeting, and $N_t^-(i) = (1-\delta_t^{nu})N_t(i)$, $U_t^-(i) = (1-\delta_t^{un})U_t(i)$ and $E_t^-(i) = (1-\delta_t^{en,i})(1-\delta_t^{eu,i})E_t(i)$ denote the number of non-participating, unemployed and employed workers of type i after the separation stage respectively. Conditional on terms (2) and (4) in this equation, there is exactly one tuple $(\lambda_t, q_t, \sigma_t)$ (jointly determined by the matching function and equation 5) that solves the free entry condition (equation 7). That is, conditional on the composition of unemployed, employed and non-participating searchers by type, and conditional on the value each type generates to the firm, the model endogenously generates the job finding rate conditional on type and a corresponding tightness. It is now possible to define an equilibrium.

Equilibrium: *Given an initial state $\{N_0(i), U_0(i), E_0(i)\}_{i\in I}$ and a path of separation and non-employment transition probabilities $\{\delta_t^{eu}, \delta_t^{en}, \delta_t^{un}, \delta_t^{un}\}_{t=1}^{\infty}$, an equilibrium is a sequence of type-state distributions $\{\{N_t(i), U_t(i), E_t(i)\}_{i\in I}\}_{t=1}^{\infty}$, worker meeting rates $\{\lambda_t\}_{t=1}^{\infty}$, firm meeting rates $\{q_t\}_{t=1}^{\infty}$, meeting success rate*

[5]This equation is an approximation - σ is evaluated at the median worker within each type. The precise version would feature an integral over workers within type but since the mass of each type is small and σ is a smooth function, the approximation is very good.

functions $\{\sigma_t\}_{t=1}^{\infty}$*, worker values* $\{V_t^{N,i}, V_t^{U,i}, V_t^{E,i}\}_{t=1}^{\infty}$*, meeting values* $\{\bar{J}_t\}_{t=1}^{\infty}$ *and wages* $\{w_t^i\}_{t=1}^{\infty}$ *such that equations 4, 5, 7, 8, 9, 10, 11, and 12 are satisfied and* $\sigma_t()$ *corresponds to the function* $f()$ *that solves equation 3 given* (λ_t, q_t)*.*

Equation 12 provides a useful decomposition of $\bar{J}_t$. First, a meeting can only create value for the firm when it turns into a match. Therefore, term (1) describes the average probability of a meeting becoming a successful match. The remaining terms (2,3,4) decompose the value of a successful match. Term (2) captures a composition effect: As the probability of search varies by employment state and different worker types have a different distribution over employment states, not all worker types are equally likely to be encountered by the firm, even accounting for their distribution in the population (μ_i). The value of a match therefore has to be adjusted for the likelihood of encountering particular worker types. Next, when there are several meetings per vacancy, firms will be able to choose the highest quality worker who accepts their offer. This skews the likelihood that a meeting will become a successful match towards higher searcher quantiles. This also has to be accounted for when calculating the value of a match which is done by including term (3), the "selection effect". Term (4) captures the value of a match conditional on worker type. I call fluctuations in this term the "direct effect".

In the model, the composition effect is at the heart of the persistently low UE rate during the early recovery. As unemployed workers are more likely to search than their non-participating or employed peers, the quality of the pool of searchers is affected by the share of each worker type that is in each employment state. As the recovery progresses, low job finding rates during the recession and in the beginning of the recovery lead to a composition effect that makes worse workers more likely to search, which makes encounters less lucrative for firms. This effect depresses the number of vacancies, which depresses the job finding rate and raises the expected number of encounters per match, which, through selection, in turn propagates the composition effect.

The selection margin also plays an important role in generating differential recovery dynamics for the employed and the unemployed. In equilibrium, when the value of a job decreases, this term endogenously rises to offset this decrease because more selective firms will be able to hire workers from higher up the quality distribution. In the beginning of a recovery, firms will cut back vacancies until the selection effect makes the benefit of posting a vacancy equal to its cost. This leads to more selective hiring, skewing the odds of successful job search more towards high quality workers.

Finally, over the course of the recovery, the direct value of creating a job conditional on matching with a particular worker type also changes due to changes in separation and poaching/job finding rates over the course of the recovery.

IV. Calibration

I calibrate the model on wage and transition probability data from the basic monthly files of the Current Population Survey (CPS) and compute the transition probabilities exactly as described in Shimer (2012). An exception to this is the series on job-to-job flows, which I take from Fujita, Moscarini and Postel-Vinay (2020) and adjust for seasonality as I do with the series from Shimer (2012).

There are 5 aggregate parameters that I hold constant across recoveries: The meeting function coefficient on job seekers, ω, the bargaining weight of the worker, β, the flow value of unemployment and non-participation (which I assume to be identical and homogeneous across workers), b, the interest rate r, and the meeting function shifter a.

Consider ω first. In the many-to-many matching model, not all meetings translate into matches. Furthermore, the average match probability conditional on a meeting depends on market tightness. Therefore, traditional estimates of the elasticity of the matching function with respect to searchers (often assumed to be the unemployed only) estimate an object that is not identical to ω. However, it turns out that the choice of ω is not critical to the results. This is because, under high values of a, such as the one I calibrate, the number of successful matches is approximately equal to the minimum of searchers and vacancies. Intuitively, a very large number of meetings leads to all vacancies and searchers being connected, so that meetings turn into matches until there are either no remaining vacancies or no remaining searchers. Mathematically, the effective matching function becomes $\tilde{M} = \min\{L, V\}$ for $a \to \infty$, regardless of ω. This has two important consequences.

First, in any calibration with a high meeting efficiency a and realistic job finding rates, matches track vacancies closely, whereas vacancies reported in JOLTS tend to be more volatile than hires over the cycle. However, JOLTS vacancies are measured only once at the end of the month, which generates a a disconnect between vacancies posted over the course of the month when their average duration fluctuates. Dividing the number of vacancies in JOLTS by their average length (from Davis, Faberman and Haltiwanger (2013)) yields estimates of the number of vacancies posted within a month that co-move closely with the number of matches. Moreover, some authors have emphasized the disconnect between recruiting effort and vacancy posting (see Gavazza, Mongey and Violante, 2018), which suggests that the concept of a vacancy in this model does not necessarily correspond perfectly to vacancies as observed in JOLTS.

The second consequence is that, in my calibration, the choice of ω has a very limited influence on the results. Regardless, I choose to use estimates of the matching function to inform ω. In particular, I follow the mean estimate in the survey of Petrongolo and Pissarides (2001) in setting $\omega = 0.4$. All results continue to hold for alternative choices of $\omega = 0.2$ or $\omega = 0.6$.

Targets for the worker bargaining parameter are somewhat elusive in the literature. Since estimates of this parameter, as far as they exist, are based on standard one-to-one matching functions, these estimates are not precisely transferable to the setting many-to-many matching. Nonetheless, to keep things simple, I follow the common practice to choose a value corresponding to the meeting elasticity, i.e. $\beta = \omega = 0.4$[6]. Next, I set b to be just below the minimum of the lowest observed steady state wage of a worker type, a value corresponding to about 59% of the average wage in the 2009 steady state. I set an interest rate of 1% per annum. While this choice is somewhat ad hoc, it turns out that the interest rate also does not have a large effect in this model, as the discount factors are mostly determined by separation rates which are orders of magnitude larger than any reasonable candidate values for the interest rate.[7]

Finally, the calibration value for the meeting efficiency a is chosen to match the ratio of the job finding probability in the third quintile relative to the first. This is a natural choice of calibration target, since a governs the relationship between searcher rank and job finding probability, as illustrated in figure 4. Figure 6 shows a comparison between all quintiles of individual average job finding rates in the NLSY[8] and compares it to the quintiles in the model for the 2009 steady state.

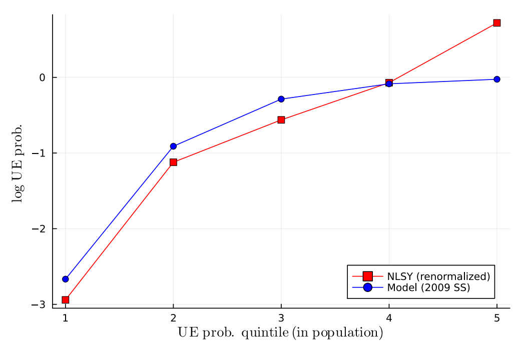


Figure 6. : Model steady state versus NLSY data

For every recovery, I re-calibrate the relative search probabilities s_e, s_n as well as steady state transition parameters $\delta^{eu}_{ss}, \delta^{en}_{ss}, \delta^{un}_{ss}, \delta^{nu}_{ss}$ to match respective aggregate period-to-period transition probabilities in steady state for every recovery, as detailed in table 2. Target steady state period-to-period transition rates are defined as the average transition rate during the last twelve months of an observed recovery. One exception to this is the EE transition probability for the first three

[6]This practice is commonly motivated with the Hosios condition. The Hosios condition is derived in a model environment with one-to-one matching and therefore does not apply in the present setting.

[7]This is largely a consequence of constant payoffs over the course of a job in this model. Back-loading of firm payoffs increases the dependency of job creation incentives on the interest rate. See e.g. Hall (2017).

[8]Since job dynamics in the NLSY turn out to feature slightly less turnover, I re-normalize the empirical rates to match the population mean of the individual job finding rate in the model's 2009 recovery steady state.

Table 2—: Internally calibrated values

Target value							
Parameter	Target	1975	1983	1992	2003	2009	2020
κ	ur_{ss}	0.06	0.055	0.039	0.047	0.035	0.035
s_e	EE	0.0283	0.0283	0.0283	0.0241	0.0233	0.024
s_n	NE	0.0495	0.0489	0.0498	0.0473	0.0438	0.044
δ_{ss}^{eu}	EU	0.0146	0.0145	0.0113	0.0116	0.0088	0.0093
δ_{ss}^{en}	EN	0.0337	0.0285	0.0286	0.0294	0.031	0.0325
δ_{ss}^{nu}	NU	0.0244	0.0228	0.0209	0.0208	0.0155	0.0149
δ_{ss}^{un}	UN	0.229	0.213	0.253	0.245	0.258	0.264
Parameter value							
Parameter	Target	1975	1983	1992	2003	2009	2020
κ	ur_{ss}	53.22	52.47	59.39	57.17	68.18	68.46
s_e	EE	0.0386	0.0383	0.0364	0.0326	0.0305	0.0312
s_n	NE	0.145	0.145	0.117	0.141	0.11	0.107
δ_{ss}^{eu}	EU	0.0198	0.0194	0.0146	0.0156	0.0118	0.0125
δ_{ss}^{en}	EN	0.0569	0.0471	0.0454	0.0492	0.0518	0.0545
δ_{ss}^{nu}	NU	0.0244	0.0228	0.0209	0.0208	0.0155	0.0149
δ_{ss}^{un}	UN	0.229	0.213	0.253	0.245	0.258	0.264

Table 3—: Externally calibrated values

Single parameters		
Parameter	Value	Explanation
ω	0.4	Petrongolo and Pissarides (2001)
β	0.4	$= \omega$
r	0.01 p.a.	
b	4.053	minimum value of steady state wage distribution

Distributional parameters	
Parameter	Target
d_i^u	relative EU prob. by worker type (2009m10-2020m2)
d_i^n	relative EN prob. by worker type (2009m10-2020m2)
y_i	w_{ss}^i (average wage by worker type, 2019)
r_i	rank of average wage by worker type (2019)

recoveries in the sample, as the series from Fujita, Moscarini and Postel-Vinay (2020) only extends back to September 1995. I therefore match an EE transition probability of 2.83% which is the average transition probability over the earliest 12 months observed in the EE series. I further calibrate κ so that the resulting steady state meeting rate λ_{ss} generates a steady state unemployment rate equal to the last unemployment rate observed during the recovery phase. I now turn to calibrating the variables that are specific to the worker type. The calibration of each worker type is based on CPS data from the recovery after the Great Recession (Oct 2009 - Feb 2020). I define a worker as any person between the age of 16 and 69. To separate them into types, I first group workers on the following observables: I use a worker's age, their sex, their ethnicity and finally their education, grouping each worker into bins corresponding to these categories. Except for the lowest age bin, every age bin comprises 5 years, so I group workers aged 16-19, 20-24, 25-29, and so on up to 65-69. I divide workers based on education into workers with no degree, high school degree, vocational degree, bachelor's degree, master's degree and doctorate degree. There are four race bins: White, Black, Asian and one bin for any worker outside these first three categories. Finally, I bin based on sex as recorded in the CPS. A worker type i is then defined as a group of workers that belong to the same age, education, race and sex bin. I exclude types with less than 100 recorded periods of employment or less than 100 recorded periods of unemployment. Then I calculate average transition rates from employment into unemployment and non-participation within each bin. These (re-normalized) transition rates serve as estimates of d_i^u and d_i^n respectively. To calibrate the steady state wage, I record the average real wage (deflated by the CPI) in 2019 for each type in the CPS data. I use these wages as calibration targets to calibrate productivity y_i based on the Nash bargaining assumption as follows: Given b and transition probabilities into employment as well as a surplus sharing rule β, I set y_i for every worker type such that it delivers exactly a steady state wage of w_{ss}^i. I also record the number/weight of workers in the CPS and assign those weights accordingly when I simulate the model. It turns out that this method of assigning worker types leads to a very granular distribution: No worker type comprises more than 2.8% of workers and only three types comprise more than 2% of workers.

Finally, I assume that the rank of a worker within the population is also determined by its steady-state wage. This is in some sense a shortcut: In an ideal version of the model, workers would be ranked in order of the value they generate for the firm, J_t^i. In this ideal version of the model, the rank would therefore be endogenous. However, when solving the model, different rankings of workers would lead to different rankings of J_t^i because changing the ranking of workers affects the probability of poaching, wages, and the steady state distribution of worker types over employment states. To calibrate this model, it would thus be necessary to find a fixed point over these rankings, the existence of which is not guaranteed. To avoid these complications, the assumption that firms rank workers by their steady-state wage is a good approximation. As figure 7 shows, the steady state wage of a worker correlates strongly (0.976 for the 2009 calibration) with the firm value of a match with that worker. Therefore, in a world where

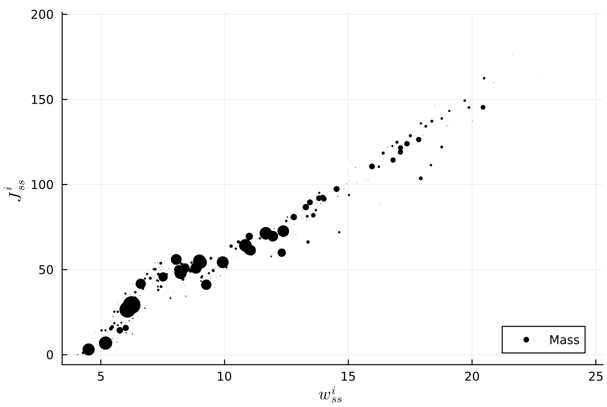


Figure 7. : Scatter plot of w^i_{ss} and J^i_{ss} (2009 calibration)

firms cannot observe worker characteristics, can observe the worker's past wage and incorrectly believe this estimate to be somewhat noisy, the ranking corresponds to the optimal ranking from a firm perspective in steady state.[9] Overall, the ranking rule of ranking by steady state wage comes very close to the optimal ranking rule from the view of firms, providing some justification for this auxiliary assumption.

V. Results

The calibrated model can now be put to the test. I apply the model to the most recent six labor market recoveries, defined as the phase during which the unemployment rate falls from its peak during a given recession to a stable trough (just before the subsequent recession). Table 4 summarizes these start and end points for the recoveries studied in this paper. Each recovery is identified by their start year, i.e. the year of peak unemployment. I initialize a recovery as follows. Starting from steady state, 5 years prior to the peak of the unemployment rate, I choose a sequence of $\{\lambda_t, \delta^{eu}_t, \delta^{en}_t, \delta^{un}_t, \delta^{nu}_t, s^n_t\}$ so that the UE, EU, EN, UN, NU and NE transition probabilities from the data are exactly replicated. Note that during the run-up to the recovery (i.e. during the recession itself) I allow the

[9]Alternatively, instead of falsely believing in a noisy estimate, one could "smooth out" the distribution of worker types in three-dimensional space sufficiently for this ranking to become optimal and then approximate the simulation under the discrete distribution here. Small amounts of smoothing would be sufficient to achieve this.

Table 4—: Start and end points of recoveries

Recovery	Start point	Unemployment rate	End point	Unemployment rate
2020	2020m4	14.7%	2023m3	3.5%
2009	2009m10	10.0%	2020m2	3.5%
2003	2003m6	6.3%	2007m11	4.7%
1992	1992m6	7.8%	2000m12	3.9%
1982	1982m12	10.8%	1990m7	5.5%
1975	1975m3	9.0%	1979m12	6.0%

probability of search for a non-participating worker to vary in order to enable precise matching of all transition rates simultaneously. This ensures that the stocks of non-participating, unemployed and employed workers track the empirically observed stocks closely. In period 0, before the first simulation period, I multiply δ_t^{eu} and δ_t^{en} by a factor that guarantees a perfect match of the unemployment rate in this period. I then start the simulation, setting s_n to its calibrated value for the remainder of the simulation. This means that from that point forward, the value of a job and all distributions move endogenously according to the rules of the model. In particular, conditional on the initial distribution and the aggregate separation paths into unemployment and non-participation, there is no further input from the data into the model. Any similarity of the recovery between model and data is therefore generated endogenously, only taking separations and transitions among the two non-employment states as given.

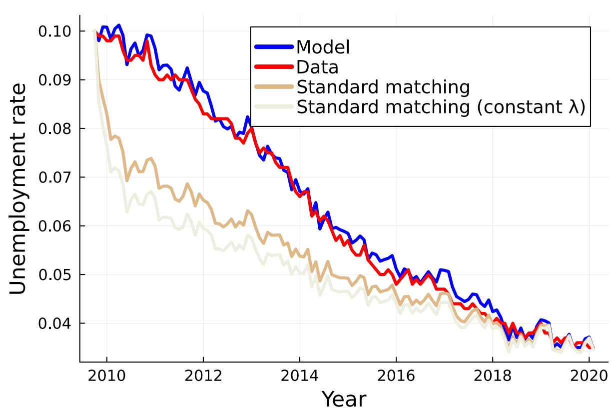


Figure 8. : Unemployment: Data vs. baseline vs. standard matching (2009 rec.)

A. Unemployment and transition rates

Figure 8 shows the modelled recovery path along with the path observed in the data for the recovery following the Great Recession, which is the longest and most recent in the sample. As a reference, the figure also shows two other recovery paths generated by modifications of the model. The first modification is a (re-calibrated) version with the same initial state distribution and standard DMP matching (i.e. $\sigma(p_L) \equiv 1$). The second is a version with DMP matching where the job finding rate λ is held constant. The first serves as a benchmark to show that a DMP framework cannot solve the recovery puzzle. The second is an illustration as to why. As is clear from the figure, the DMP model does not generate a drop in the job finding rate that comes close to matching the observed recovery. Unemployment remains at a level above its steady state for a while because exogenous separations increase the inflow into unemployment. While increased separations and the initial composition effect also cause the job finding rate to drop slightly in the DMP model (visible as a slightly higher unemployment rate in the case where λ is allowed to vary), the drop is small and the new unemployment rate quickly falls to a value well below its starting point. Because the dynamics of a system with homogeneous job finding rates are fast, the unemployment rate quickly converges to the lower level, from which it recovers gradually.

In contrast, the matching mechanism in the many-to-many matching model generates an unemployment series that follows the series observed in the data extremely closely. As figure 8 suggests, the reason the many-to-many model can deliver the correct recovery dynamics is because it delivers the correct average job finding rate for unemployed workers. Figure 9 confirms this: The model generates a post-recession drop in the UE transition probability that only slowly adjusts upward as the recovery progresses. The same figure also highlights another important feature of the model: As in the data, while the UE transition probability is well below steady state level at the beginning of the recovery, the EE transition probability barely drops at all. This feature is generated endogenously through the ranking assumption of the model: The employed are typically higher-ranked workers who still tend to end up on top of their encountered vacancies' appli-

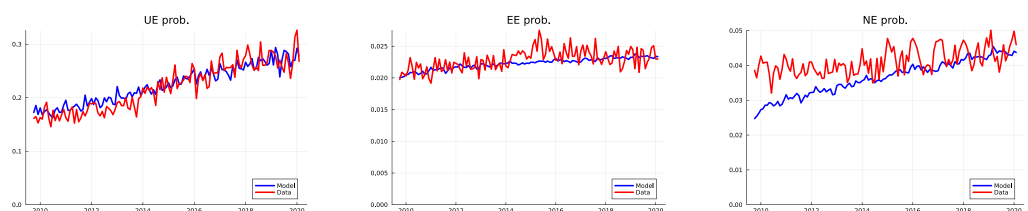


Figure 9. : Transition probabilities in model and data (2009 recovery)

cant pools even if market tightness is low and every vacancy receives many more matches than in normal times. The unemployed however tend to be lower-ranked workers. If they encounter a vacancy, they are likely to be outranked by another worker. The post-recession environment amplifies this effect: As market tightness drops, there are more applicants for a given position, making it even more likely for such workers to be outranked. This slows down the recovery of unemployment, because the outflow from the stock of unemployed workers decreases. Finally, figure 9 also shows that the model somewhat underestimates the NE transition rate during the early recovery.[10]

B. The role of selection

To highlight the role that selection plays in this process, figure 10 shows the counterfactual selection probabilities if every worker would transition into employment with the average aggregate probability (i.e. if workers were selected randomly but holding vacancy creation and all other variables constant). It is

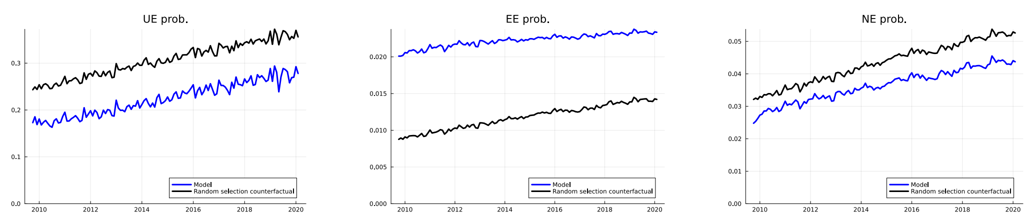


Figure 10. : Actual and no-selection transition probabilities in the baseline model

immediate that selection depresses the job finding rate for the unemployed and elevates it for the employed. This is particularly true for the beginning of the recovery, where the (percentage) drop in the UE transition probability would be significantly stronger without selection. It is also evident from the figure that employed searchers are more than twice as successful looking for a job as they would be if they did not enjoy a ranking advantage.

An alternative way to assess the role of selection is to recalibrate a model version in which selection is eliminated by ranking all workers randomly each period.[11] Importantly, this is not identical to the standard matching mechanism, as the

[10]There are at least two potential reasons for this: The first is that, mechanically, the model does not assume the UN or NU transition probabilities to be type-dependent, generating similar compositions for both pools. In reality there might be a composition of worker types in the pool of non-participating that tends to make this pool more successful. Secondly, the NE-margin might in part be driven by a matching technology that lies outside the traditional random-search framework, such as recalls into old jobs after taking time off. However, it is clear that this mismatch between model and data does not adversely affect the model fit for the unemployment rate.

[11]To facilitate comparison, a is held constant.

model retains multiplicity of encounters and therefore also the properties of the many-to-many meeting function, aside from selection. Again, we can clearly see

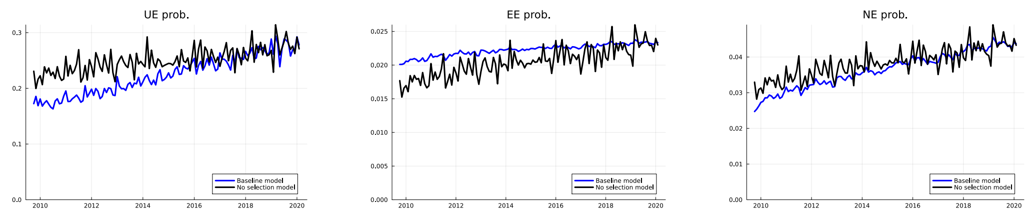


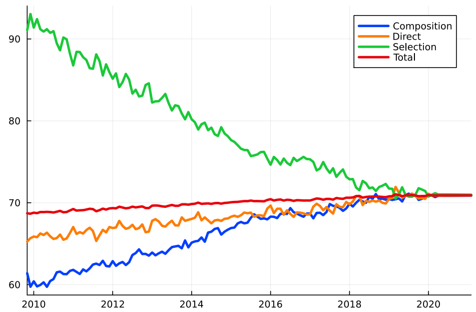
Figure 11. : Transition probabilities in the baseline and no-selection models

that the UE rate suffers a far less severe drop early during the recovery in the no selection model compared to the baseline model and the data. The EE rate drops much more than both the baseline model and the data. The NE rate, likewise, is slightly more stable than in the baseline model, although the difference is minor.

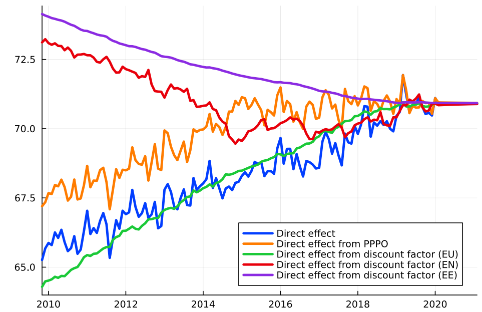
Overall, these figures show that selection plays an important role in translating slack markets into environments with a low exit rate out of unemployment. However, selection works to amplify the effect of slack markets on the unemployment exit rate. To fully understand the recovery dynamics resulting from the model, it therefore remains to answer why market tightness is low in the aftermath of a recession and does not quickly revert back to steady state after the shock has disappeared, as it does in the baseline DMP model.

C. Forces depressing job creation

To analyze the forces responsible for this, recall that one can map J and λ into each other per equations 4 and 7, which jointly give $\lambda_t = a^{\frac{1}{\omega}} \left(\frac{J_t}{\kappa}\right)^{\frac{1-\omega}{\omega}}$, establishing that the number of meetings per searcher is a monotone transformation of the job creation incentive. Therefore, a decomposition of J is informative about the fundamentals determining the state of the labor market, captured by λ. We can therefore decompose the average value of a match J_t into changes in the three terms (2), (3) and (4). Note that the combined change has to move one-to-one with the expected number of meetings per vacancy. Concretely, I decompose J_t by holding two out of the two terms at their steady state level while letting the remaining one vary as it does in the model simulation. Figure 12a shows the result of this decomposition for the Great Recession. As is clear from the figure, both the direct effect and the composition effect play a role in depressing the value of a match. Quantitatively, the composition effect plays the larger role in my calibration. This means that in the model, the initial composition of searchers is skewed towards worse workers (generated by low market tightness during the



(a) Decomposition of J_t (2009 rec.)



(b) Direct effect decomp. (2009 rec.)

Figure 12. : Decompositions of J_t

recession). This composition effect then gets propagated as high quality workers continue to quickly transition into employment while lower quality workers find it near impossible to find jobs under the slack labor market conditions present during the early phase of the recovery. Only over time, this effect becomes less important and market conditions relax to the point where the dynamics reach steady state.

This finding warrants further discussion. There is a long and active debate in the literature on whether composition-based explanations for persistent unemployment fluctuations are consistent with the data. A long line of papers argues that the pool of unemployed workers becomes worse in the immediate aftermath of a recession. This is built into many models of unemployment fluctuations with worker heterogeneity, such as Pries (2008), Ravenna and Walsh (2012) or Ferraro (2018). Barnichon and Figura (2015) provide evidence that in the aftermath of the Great Recession, the pool of unemployed shifted towards workers with lower job finding rates, consistent with the finding in this paper. For the same episode, Gregory, Menzio and Wiczer (2021) argue that the pool of unemployed shifted towards workers with high separation rates. On the other hand, Mueller (2017) shows that the composition of the pool of unemployed is counter-cyclical and shifts to higher-wage workers in recessions.

My setting differs in some aspects from these papers. In addition to the unemployed, both the employed and the non-participating search in my setting. Nonetheless, the qualitative decline of the unemployment pool is an important driving force of the mechanism. Importantly though, the mean worker in unemployment is not what governs firm decisions in the model. Instead, the importance of each worker in unemployment depends on their likelihood of selection, given by $\frac{\sigma_t(p_L^t(i))}{\int_0^1 \sigma_t(\tilde{p}_L)\, d\tilde{p}_L}$. This weighting does much to amplify the cyclicality of the composition of the pool of unemployed. Figure 13 demonstrates this for the recovery from the Great Recession: The black line traces out the expected value of a job for workers in the unemployed population, given by $\int_0^1 \frac{U_t^-(i)}{\int U_t^-(i)d\tilde{\mu}_i} J_t^i d\mu_i$. The green line traces the same value, but weighted by the selection probability, i.e. $\int_0^1 \frac{U_t^-(i)}{\int U_t^-(i)d\tilde{\mu}_i} \frac{\sigma_t(p_L^t(i))}{\int_0^1 \sigma_t(\tilde{p}_L)\, d\tilde{p}_L} J_t^i d\mu_i$. As can be seen in the figure, the green line exhibits much stronger cyclicality than the black line.

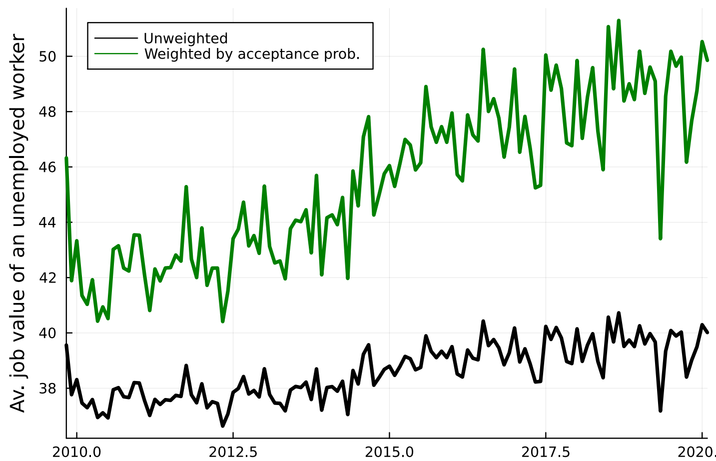


Figure 13. : Expected job value J_t^i among unemployed

Thus, while the model does imply a pro-cyclical composition of the pool of unemployed, the average quality of the worker is nearly acyclical. Figure A7 in the appendix shows that this pattern is even more pronounced for some of the other recoveries in my sample.

Next, I briefly turn to the direct effect. It can be further decomposed into the effects coming from the different components of the discount factor as well as the effect that comes from the per-period payoff (PPPO) for the firm through the interplay of market conditions and wages. To do this, I perform a decomposition of equation 9 by the same principle: I plot J_t while holding all distributions constant at their steady state, and holding constant all but one of $\delta_t^{en,i}$, $\delta_t^{eu,i}$, $\sigma_t(p_L(i))\lambda_t$ or w_t^i. Figure 12b plots this decomposition of the direct effect. It is clear that the main effect depressing the direct effect is the transition probability from employment into unemployment which remains above steady state for a long time into the recovery. The transition probability from employment into non-participation is depressed during the recovery, increasing the value of a match. The transition rate into employment is also depressed as the aggregate job finding probability falls. This, too, works towards increasing the value of a job. Finally, wages have a moderate downward effect on the job value. This comes from the Nash bargaining assumption in connection with a deteriorating value of on the job search under reduced poaching rates which must be compensated by rising wages.

We can now turn to other recoveries, applying the same model calibration to them (but re-calibrating steady state value of all δs and κ to match new steady states). Figure A6 shows the unemployment fit for these other recoveries. The

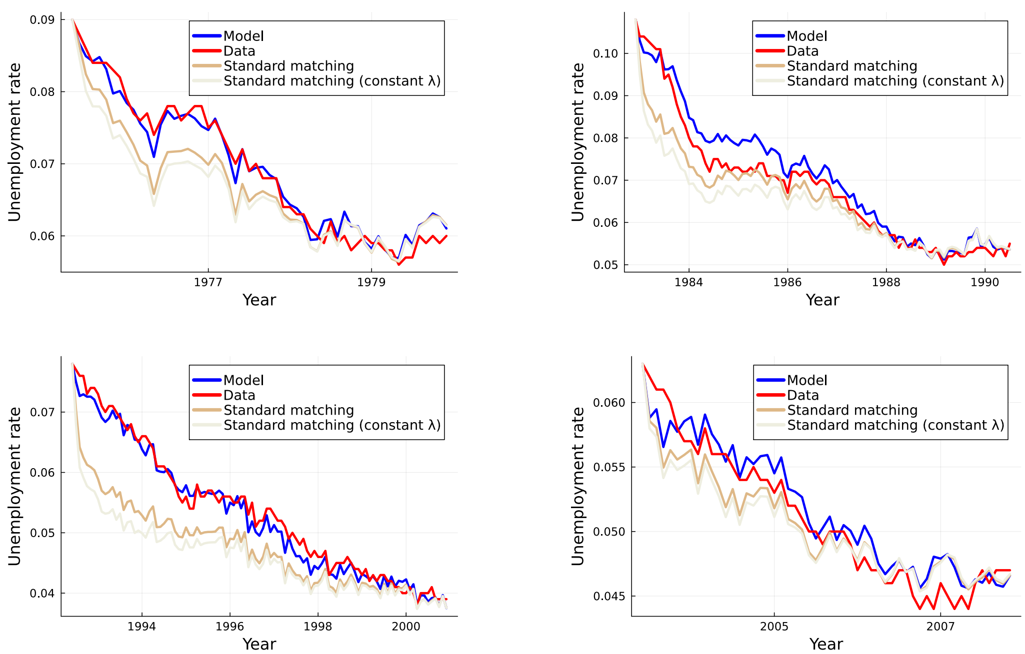


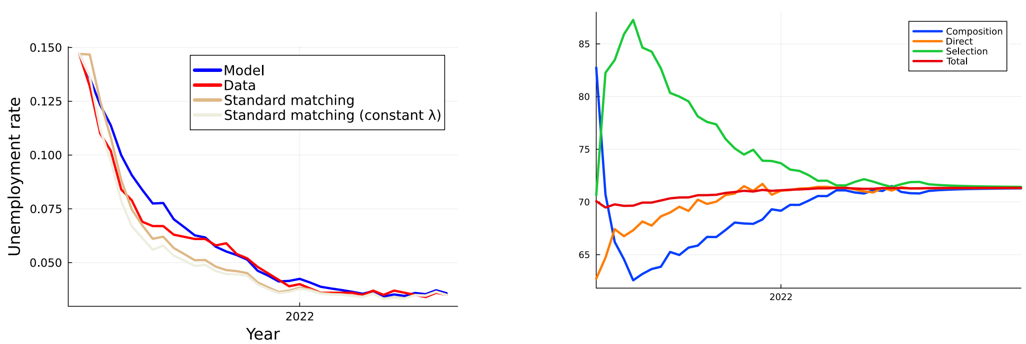
Figure 14. : True and simulated unemployment series for other recoveries

fit of the simulated unemployment series compared with the data is excellent for three out of the four recoveries (1975, 1992 and 2003) and good for the fourth (1982). The simulation of the 1982 recovery is somewhat of an outlier in that the unemployment rate plateaus at a level that is slightly too high towards the first half of the recovery. Aside from this, the many-to-many matching model is able to endogenously generate the exact shape of recovery observed after every of the last 5 major recessions in the US. The figure also shows that the model replicates the series both for recoveries with a large direct contribution from the separation margin (such as 1975 and 2003) as well as for recoveries in which this direct effect is not nearly sufficient to explain the elevated unemployment rate (such as 1992 and 2009).

The model is also successful in capturing the quality difference between the pool of employed and unemployed searchers. To demonstrate this, I compute the wage premium for hires out of employment relative to hires out of unemployment. Faberman et al. (2017) estimate this premium to be 36 log points, 17 log points can be explained with observable characteristics. In the baseline model, the wage premium of a hire out of employment in the 2009 recovery steady state is 19 log points.

D. The Covid Recovery

The model is able to reproduce the surprising regularity in the behavior of unemployment rates during the 1975, 1982, 1992, 2003 and 2009 recoveries. However, in more recent history, the US has experienced another recovery that falls a little outside the pattern established by its predecessors: The labor market readjustment after the Covid pandemic of 2020. In many respects, the Covid recovery is unusual. Transition dynamics in and out of unemployment were confounded by a large number of re-calls into old jobs as demand in the US economy began to recover and supply restrictions were loosened. The unemployment rate jumped up to 14.7% but then adjusted much faster than seen in previous recoveries. This arguably poses an interesting challenge to the model. Can the selection-based model predict the dynamics that arose from the 2020 recovery?



(a) 2020 recovery

(b) Decomposition of J_t

Figure 15. : Model predictions for the Covid recovery

Figure 15 shows the model-implied dynamics. Although standard matching and many-to-many matching perform equally in the first phase of the recovery, the baseline model gains the upper hand in the latter phase of the recovery. Notably, the two models do not differ much in their implications for the unemployment rate, in contrast to their predictions for other recoveries.

VI. Conclusion

This paper makes two contributions. First, it develops a ranked matching capable of modeling selection in a tractable way. As in the data, this matching mechanism predicts that more productive workers return to the workforce more quickly than their less productive and but also explains the higher business cycle sensitivity of less productive workers. Second, I show that in a heterogeneous agent economy with selection, composition and separation shocks are enough to explain the shape of recovery under this matching mechanism. Thus, the changing composition of the pool of job seekers and selection can be considered a primary source of sluggish adjustment during past recoveries. The model is able to closely replicate the unemployment series of all of the last 6 recoveries in the US economy.

The ranked matching mechanism opens up more avenues for future research. In particular it links the efficiency of the meeting technology to inequality in labor market outcomes. While in this paper I have abstracted from changes in the meeting efficiency, a model along these lines could assess the impact of better matching technologies such as the rise of online application systems on inequality. I leave this question for future research.

In this paper I take separation rates as given and therefore side-step the perhaps more fundamental question of why they stay elevated for a long time after the initial recession shock has subsided. This is another question future research should address.

REFERENCES

Barnichon, Regis, and Andrew Figura. 2015. "Labor market heterogeneity and the aggregate matching function." *American Economic Journal: Macroeconomics*, 7(4): 222–249.

Birinci, Serdar, Kurt See, and Shu Lin Wee. 2020. "Job applications and labor market flows." *FRB St. Louis Working Paper*, , (2020-23).

Blanchard, Olivier Jean, and Peter Diamond. 1994. "Ranking, unemployment duration, and wages." *The Review of Economic Studies*, 61(3): 417–434.

Davis, Steven J, R Jason Faberman, and John C Haltiwanger. 2013. "The establishment-level behavior of vacancies and hiring." *The Quarterly Journal of Economics*, 128(2): 581–622.

Eeckhout, Jan, and Ilse Lindenlaub. 2019. "Unemployment cycles." *American Economic Journal: Macroeconomics*, 11(4): 175–234.

Engbom, Niklas. 2021. "Contagious Unemployment." *NBER Working Paper.*

Faberman, R Jason, Andreas I Mueller, Ayşegül Şahin, and Giorgio Topa. 2017. "Job search behavior among the employed and non-employed." *NBER Working Paper.*

Ferraro, Domenico. 2018. "The asymmetric cyclical behavior of the US labor market." *Review of Economic Dynamics*, 30: 145–162.

Fujita, Shigeru, Giuseppe Moscarini, and Fabien Postel-Vinay. 2020. "Measuring Employer-to-Employer Reallocation." *NBER Working Paper*.

Gavazza, Alessandro, Simon Mongey, and Giovanni L Violante. 2018. "Aggregate recruiting intensity." *American Economic Review*, 108(8): 2088–2127.

Gregory, Victoria, Guido Menzio, and David G Wiczer. 2021. "The alpha beta gamma of the labor market." *NBER Working Paper*.

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

Hall, Robert E. 2017. "High discounts and high unemployment." *American Economic Review*, 107(2): 305–30.

Hall, Robert E, and Marianna Kudlyak. 2022. "Why Has the US Economy Recovered So Consistently from Every Recession in the Past 70 Years?" *NBER Macroeconomics Annual*.

Mercan, Yusuf, Benjamin Schoefer, and Petr Sedláček. 2021. "A Congestion Theory of Unemployment Fluctuations." *NBER Working Paper*.

Mueller, Andreas I. 2017. "Separations, sorting, and cyclical unemployment." *American Economic Review*, 107(7): 2081–2107.

Petrongolo, Barbara, and Christopher A Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic Literature*, 39(2): 390–431.

Pries, Michael J. 2008. "Worker heterogeneity and labor market volatility in matching models." *Review of Economic Dynamics*, 11(3): 664–678.

Ravenna, Federico, and Carl E Walsh. 2012. "Screening and labor market flows in a model with heterogeneous workers." *Journal of Money, Credit and Banking*, 44: 31–71.

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

Shimer, Robert. 2006. "On-the-job search and strategic bargaining." *European Economic Review*, 50(4): 811–830.

Shimer, Robert. 2012. "Reassessing the ins and outs of unemployment." *Review of Economic Dynamics*, 15(2): 127–148.

Wolthoff, Ronald. 2018. "Applications and interviews: Firms' recruiting decisions in a frictional labour market." *The Review of Economic Studies*, 85(2): 1314–1351.

*

Appendix

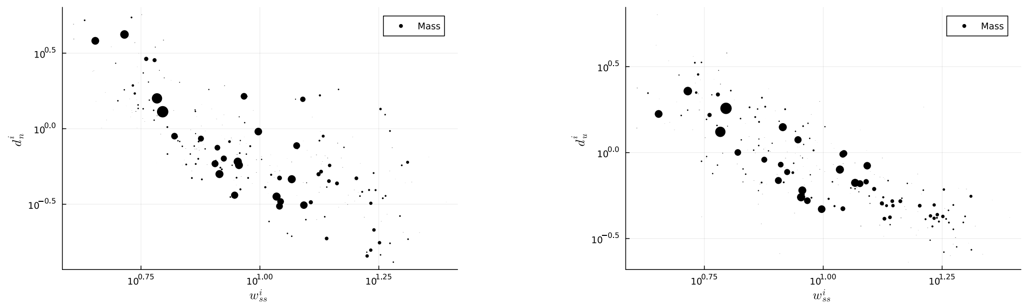


Figure A1. : Calibrated wages and relative transition probabilities by type

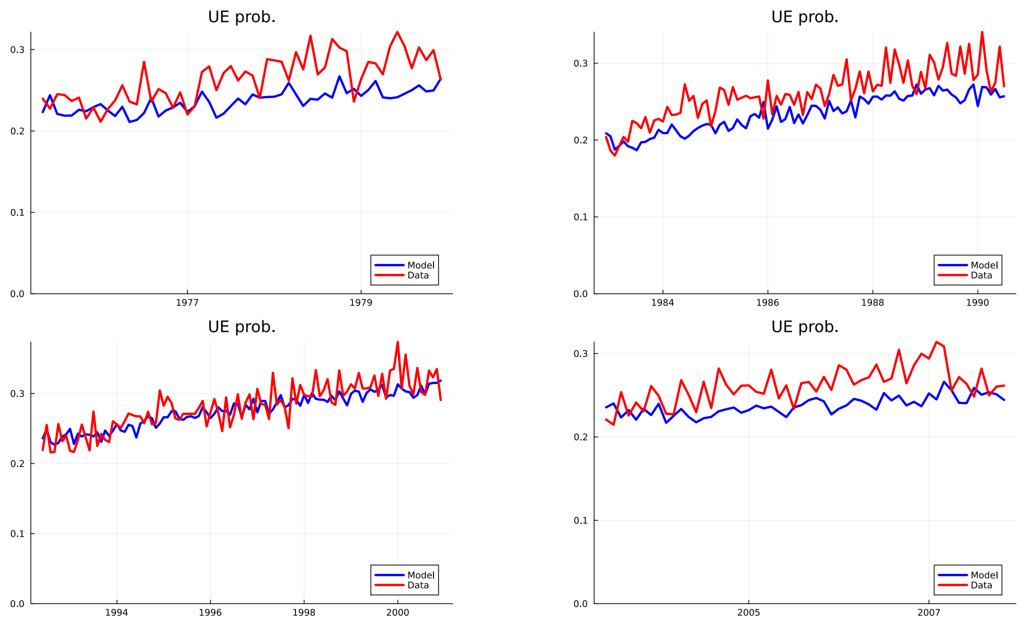


Figure A2. : UE transition probability for 1975-2003 recoveries

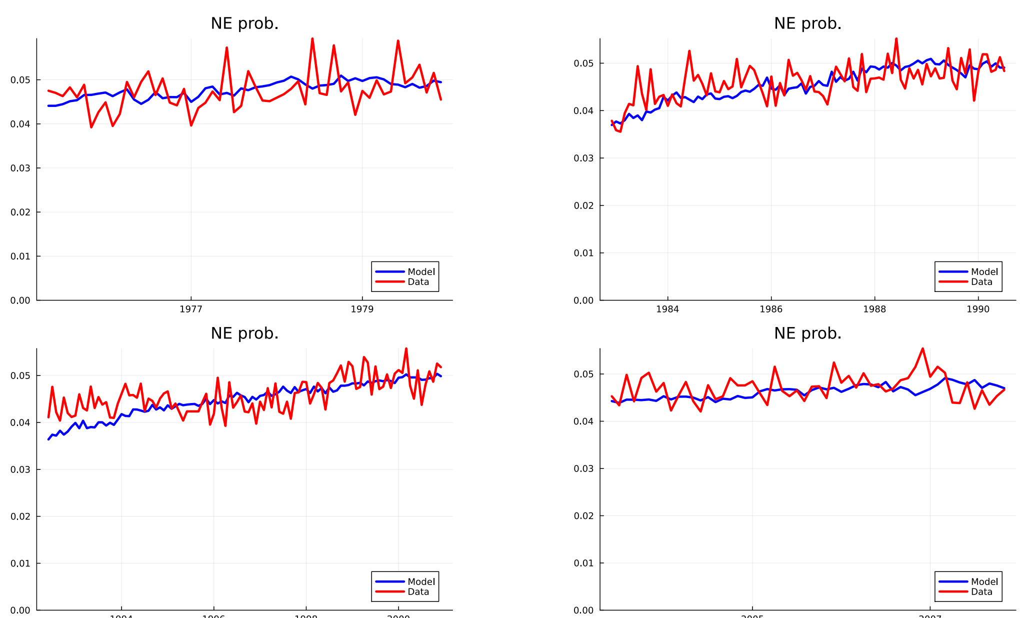


Figure A3. : NE transition probability for 1975-2003 recoveries

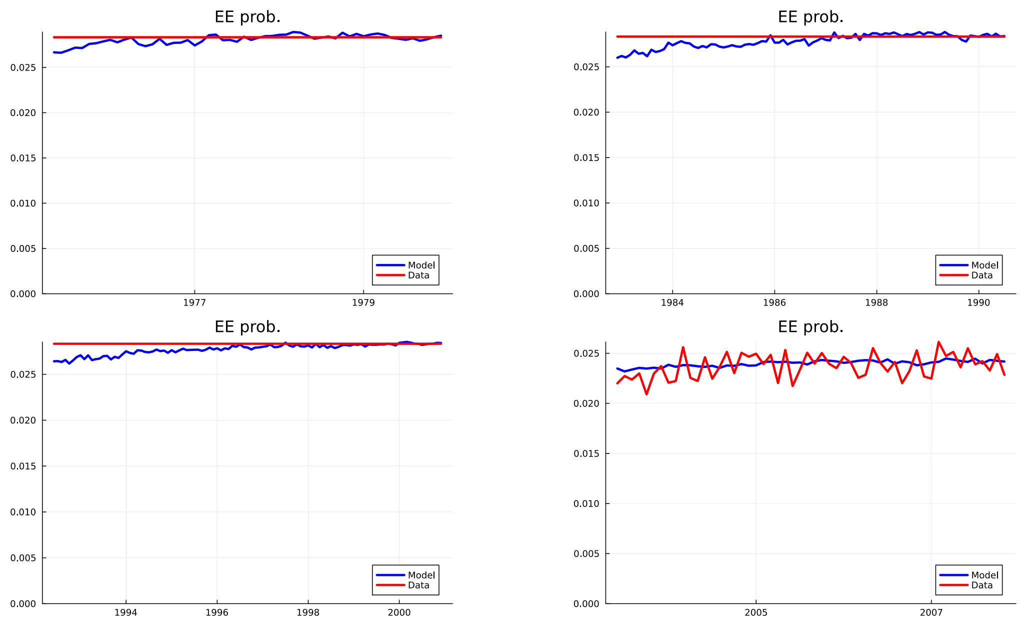


Figure A4. : EE transition probability for 1975-2003 recoveries

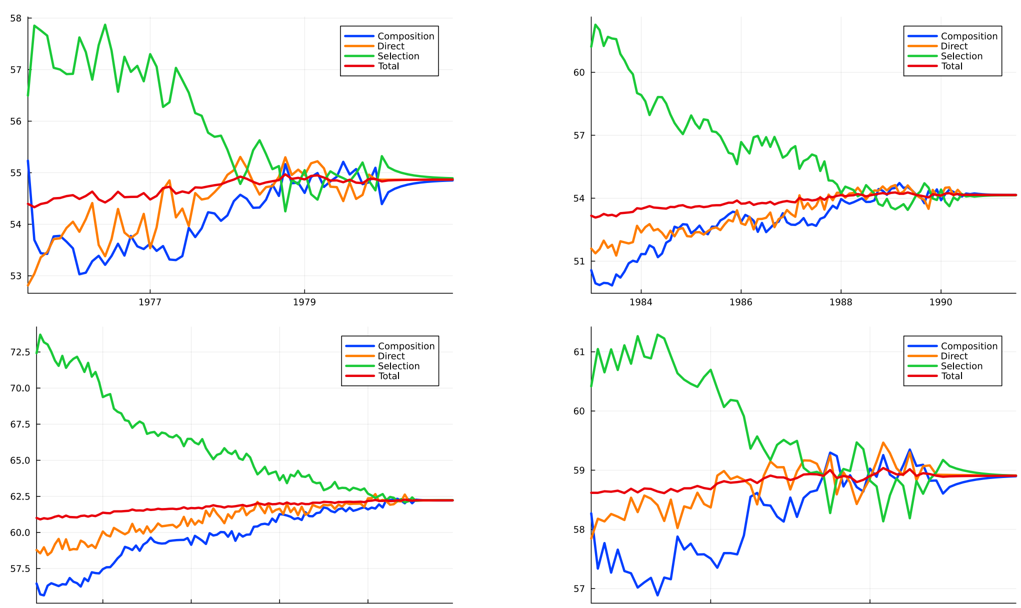


Figure A5. : Decomposition of J_t for 1975-2003 recoveries

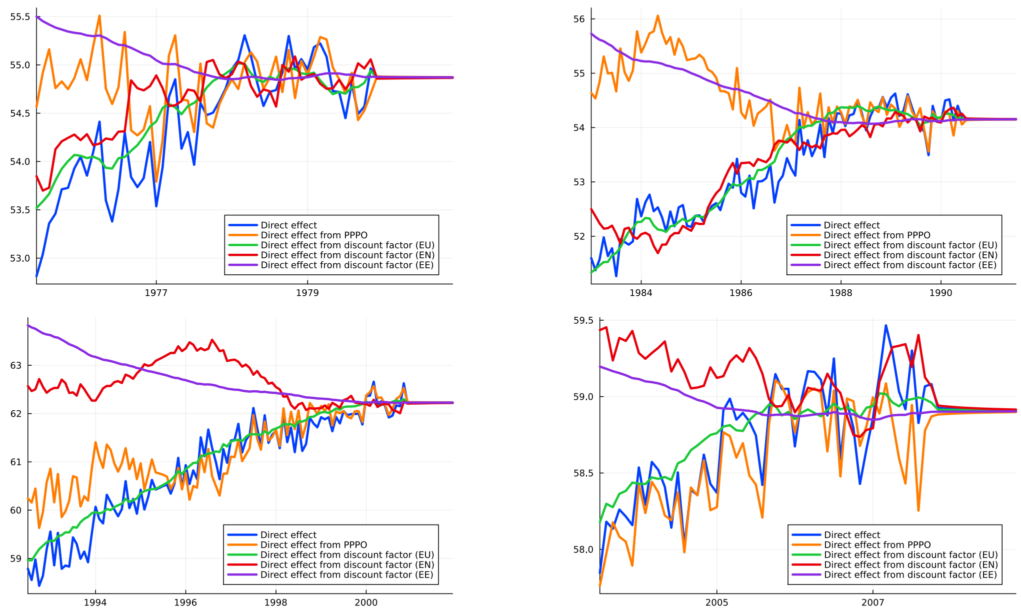


Figure A6. : Decomposition direct effect of J_t for 1975-2003 recoveries

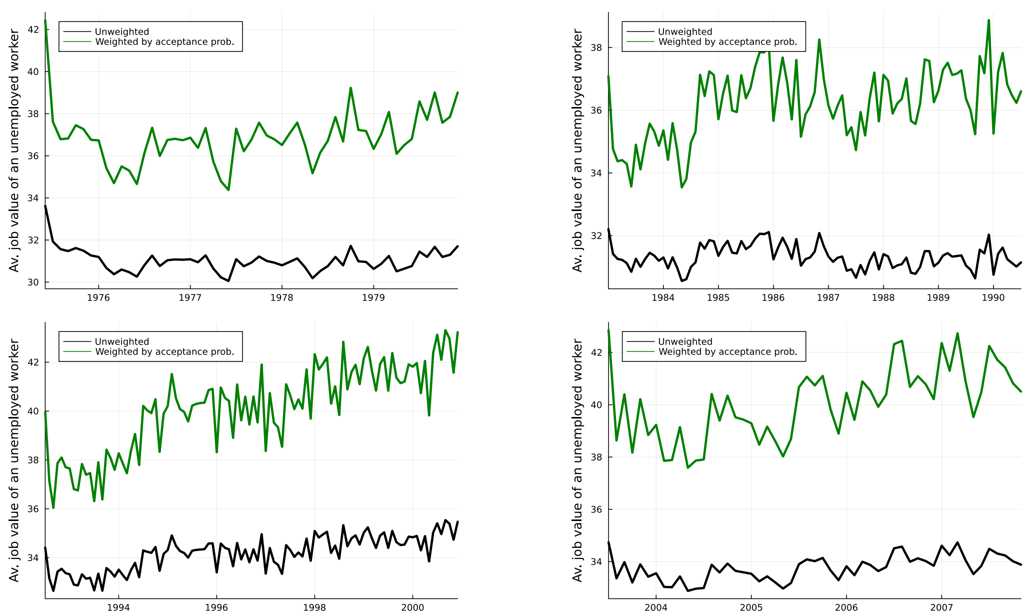


Figure A7. : Composition of the unemployed for 1975-2003 recoveries