Unemployment Cycles*

Jan Eeckhout[†] and Ilse Lindenlaub[‡]

October 2018

Abstract

The labor market by itself can create cyclical outcomes, even in the absence of exogenous shocks. We propose a theory in which the search behavior of the employed has profound aggregate implications for the unemployed. There is a strategic complementarity between active on-the-job search and vacancy posting by firms which leads to multiple equilibria: in the presence of sorting, active on-the-job search improves the quality of the pool of searchers. This encourages vacancy posting, which in turn makes costly on-the-job search more attractive – a self-fulfilling equilibrium. The model provides a rationale for the Jobless Recovery, the outward shift of the Beveridge Curve during the boom and for pro-cyclical frictional wage dispersion. Central to the model's mechanism is the fact that the employed *crowd out* the unemployed when on-the-job search picks up. We illustrate this mechanism in a stylized calibration exercise.

Keywords. On-The-Job Search. Strategic Complementarity. Multiplicity. Unemployment Cycles. Sorting. Mismatch. Job-To-Job Flows. Jobless Recovery. Beveridge Curve Shift.

*We are grateful to Jess Benhabib, Katarina Borovickova, Hector Chade, Mike Golosov, Boyan Jovanovic, Guido Menzio, Fabien Postel-Vinay, Edouard Schaal and Randy Wright for helpful discussion and comments. We also benefited from the feedback of many seminar audiences. Eeckhout gratefully acknowledges support from the ERC, Grant 339186, and from RecerCaixa. We benefitted from excellent research assistance of Julia Faltermeier, Adrià Morrón-Salmerón and Shubhdeep Deb.

†UPF (Barcelona GSE and ICREA) and University College London, `j.eeckhout@ucl.ac.uk`

‡Yale University, `ilse.lindenlaub@yale.edu`

1 Introduction

Business cycles have a wide variety of origins, ranging from financial crises, over oil price shocks, to productivity spurts and slowdowns. Often, of all economic agents, workers are those affected most dramatically, mainly through unemployment. For long, researchers – most notably Diamond (1982) – have asked whether frictional markets can generate cyclical outcomes, even in the absence of any exogenous shocks or changes in fundamentals. But so far there has been no compelling mechanism where the *labor market by itself* can generate cycles and that fits the facts. In this paper, we propose a simple theory that generates self-fulfilling unemployment fluctuations and that can account for the key labor market facts: Our model provides a simple rationale for large fluctuations in unemployment, vacancies and job-to-job flows, for the Jobless Recovery (the fact that it takes a long time for unemployment to drop even after vacancies and productivity have recovered), for the outward shift of the Beveridge Curve during recovery and for the evolution of frictional wage dispersion over the business cycle. These important business cycle aspects of the labor market cannot be rationalized in the standard random search model of the labor market.

The main contribution of this paper is to develop a theoretical mechanism that can explain these phenomena, where the main driving force is the *search behavior of the employed.* Singling out the *employed* to explain unemployment may seem counterintuitive. But with a share of over ninety percent of the labor force, any minor change in the behavior of the employed, who vie for job openings side by side with the unemployed in the same labor market, has profound aggregate implications for unemployment. Even if they search much less intensively than the unemployed, simply because of their size, *on average* almost half of the new jobs are filled by workers who were employed already. Most importantly, we document that there is a large *cyclical variation* in the composition of searchers, ranging from 32% of employed workers in the recession to 48% in the boom, mostly due to the change in the search behavior of the employed over the business cycle. This variation in the composition of searching workers is a novel empirical finding; our work suggests that it is important for the cyclical dynamics of unemployment.

We contribute to the literature by spelling out a model that features a *strategic complementarity* between on-the-job search (OJS) and vacancy creation, giving rise to multiple equilibria. In their search decision, workers trade off the *matching rate* against the *cost of searching.* In turn, in their vacancy posting decision, firms take into consideration both the *expected quality* (or productivity) as well as *duration* of the job. When workers actively search on-the-job, there are two opposing effects on the firm's value of a job. First, in the presence of sorting, searchers tend to move to jobs with higher match quality, and with active search the relative number of on-the-job searchers compared to unemployed searchers is higher (which we refer to as a *composition externality* in the pool of searchers), pushing the value of a job up. But at the same time, under active search, the expected duration of a job is shorter, pushing the value of a job down. It is precisely the interplay between the composition externality and the different duration of a job that is at the root of the multiplicity.

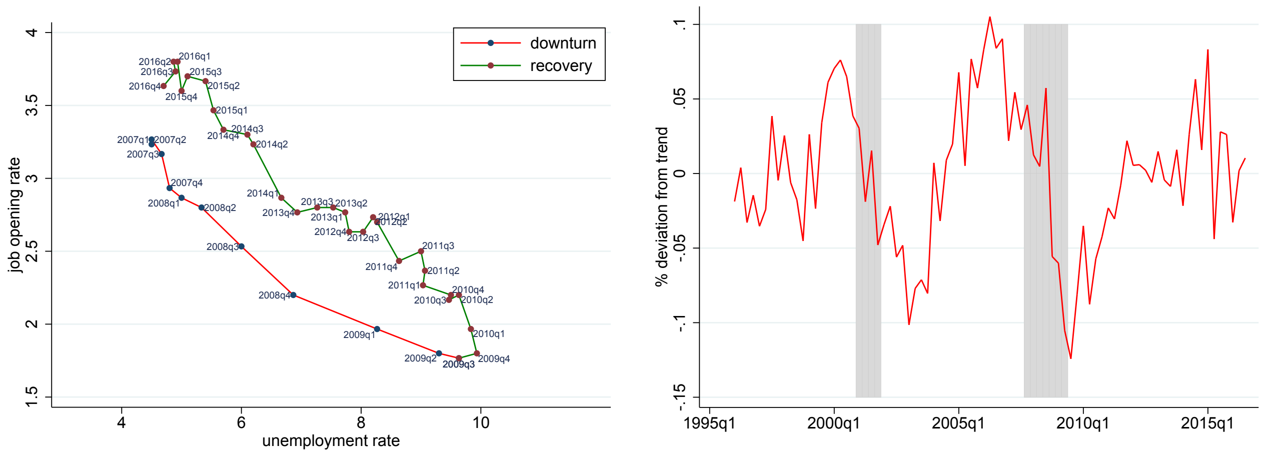
With active OJS, the favorable change in the composition of searchers (and thus higher expected match quality) dominates the shorter job duration, which creates incentives for vacancy posting. More vacancies in turn create incentives for workers to actively search on-the-job since it is easier to find one. Active job search has become self-fulfilling. This high churning outcome corresponds to an economic *boom* with active on-the-job search, high employment-to-employment (EE) transitions, little mismatch, abundant job creation, low unemployment and high aggregate output. But there is also another equilibrium, the *recession*, where workers do not actively search on-the-job, where the pool of searchers has relatively few on-the-job searchers and the expected productivity of a job is low. For firms, the shorter duration of jobs formed with on-the-job searchers here dominates the impact of the composition externality. As a result, firms have little incentives to post vacancies. This generates a low matching rate for workers that does not compensate the cost of search. Again, this low search intensity is self-fulfilling. It leads to low worker turnover and high mismatch, low aggregate output and high unemployment. In the recession, workers experience grim job prospects and hang on to their existing jobs, even if mismatched. Firms take solace in the long duration of jobs, even if they are of low productivity.

This purely endogenous labor market mechanism is new in the literature. The underlying strategic complementarity between vacancy posting and OJS intensity builds on three features: 1. endogenous OJS, 2. sorting (job ladder) with mismatch, and 3. endogenous vacancy creation. These ingredients give rise to two closely related composition shifts that generate the necessary feedback effects for self-fulfilling equilibria: the composition of employed workers across the job ladder and the composition of searchers. First, under active OJS, the job ladder is more 'fluid', so more workers transit from mismatched jobs at the bottom to better jobs at the top of the job ladder. This means that – counterintuitively at first sight – the share of *employed workers* who search on-the-job is smaller during the boom, simply because less workers are stuck at the bottom of the ladder. Second, however, the likelihood for a vacant job to draw an on-the-job searcher rather than an unemployed searcher is higher under active OJS, because the fewer employed workers who search do so more intensely. Endogenous OJS intensity thus affects the efficiency units of searchers. Based on efficiency units, the *pool of searchers* shifts towards on-the-job searchers during a boom (above referred to as the composition externality). Both composition shifts are absent in any standard random search models, which is why they do not generate endogenous fluctuations. We will provide direct empirical evidence for these cyclical composition changes below.

While we derive most of our analytical results focussing on steady state equilibria, we also find parameter regions where *multiple dynamic equilibria* exist. In these situations, based on the local behavior of the model's dynamic system, both steady states are saddle foci. For the global dynamics, this implies that both the active and passive job search equilibrium are stable manifolds on which, for a given belief about aggregate OJS behavior, the economy converges on oscillating paths to the corresponding steady state. We find initial conditions for which there are multiple dynamic equilibrium paths converging to two different steady states. To which steady state the economy will converge depends on the agents' beliefs about whether aggregate OJS is active or not.

While the contribution of this paper is predominantly theoretical in that we identify a new mechanism behind fluctuations that originates exclusively in the labor market, we also numerically illustrate this mechanism: We calibrate the model to the US economy during the Great Recession and show that it is quantitatively consistent with large cyclical fluctuations in labor market outcomes, as well as with the jobless recovery, the shift in the Beveridge Curve during the recovery, and pro-cyclical frictional wage dispersion. We now discuss each in turn.

First, even in the absence of any exogenous shocks (for instance to productivity and to financial markets), our model can be made consistent with large *cyclical variations* in unemployment and vacancies (Figure 1a) as well as EE transitions (Figure 1b) in the data. In our model, these fluctuations stem from the labor market itself but are more difficult to generate in the standard random search model with shocks to fundamentals (see Shimer (2005)).



(a) Beveridge Curve During and After Great Recession.

(b) Quarterly Employment-to-Employment Rate (detrended using HP filter, see Data Appendix for details)

Figure 1: Employment-to-Employment Flows and the Beveridge Curve.

The second labor market feature the model can account for is the *Jobless Recovery*. Even after productivity has picked up following the recession, unemployment has remained sluggishly high. It took until 2016, seven years after the end of the Great Recession, for unemployment to be back at 5%. Here we identify a new underlying channel where the employed searchers are *crowding out* the unemployed ones during recovery. At the end of a recession, as beliefs switch to an active OJS regime and firms add many vacancies, the composition of the pool of searchers changes. Overall, job creation picks up, but jobs go disproportionately to the on-the-job searchers (who are abundant after the recession), and not to the unemployed. All the renewed activity thus initially translates in better jobs for the employed, but not in improved prospects for the unemployed.

Third, the model is consistent with the observed cyclical variation in frictional wage dispersion. We use our model's implied *Mean-min wage ratio* (originally developed in Hornstein, Krusell, and Violante (2011)) to assess frictional wage dispersion in the data. First, we show that this measure of wage dispersion is highly pro-cyclical – a finding we believe is new in the literature. We then show that our model matches these observed patterns. Contrary to the phenomenon of jobless recovery, the primary force here is not the fluctuation in the composition of *searchers* but the change in the composition of *employed workers* across the job ladder. In the boom with active OJS, the job ladder is much more 'fluid' with many workers moving into the upper rungs, fueling wage dispersion. In contrast, in the recession with low search intensity of employed workers, many of them are stuck in bad jobs – the job ladder fails – and frictional wage dispersion is low. This composition shift of employed workers across the job ladder is more important for pro-cyclical wage dispersion than the movements in wages themselves.

Finally, while these findings mainly rely on steady state comparisons of our calibrated economy, we exploit the *equilibrium dynamics*, where we assess the response to a positive (unexpected and permanent) expectations shock that pushes the economy out of the recession. Following this expectations shock while the economy is in the recession, we study the transition path to the boom steady state, and show that our model can trigger an outward shift of the *Beveridge Curve*. Recently there has been a renewed interest in the Beveridge curve because of a sizable outward shift following the Great Recession (see Figure 1a).[1] An outward shift is often interpreted as a decrease in matching efficiency: to maintain a given level of unemployment, a larger number of vacancies needs to be posted. A deterioration of labor market efficiency during the recovery is puzzling. One would expect to the contrary that part of the recovery is due to improved matching. We argue that the shift of the Beveridge Curve is not due to matching efficiency but rather due to a shift in the *effective* market tightness. While the Beveridge Curve relates to the ordinary market tightness, i.e. the ratio of vacancies to unemployed, the *effective* market tightness is given by the ratio of vacancies to *all searchers*, including in the denominator not only the unemployed but also on-the-job searchers. The fact that OJS picks up during recovery leads, for a given number of vacancies, to a decrease in the *effective* labor market tightness. Job offers start going disproportionally to employed searchers, crowding out the unemployed workers and resulting in lower job finding probabilities for them. For a given vacancy rate, there are more unemployed workers, the Beveridge Curve shifts. This phenomenon is closely related to the Jobless Recovery we discussed

[1]The outward shift is substantial. For example, with 1.9% of vacancy creation, the unemployment rate at the end of 2008 was 7.5% while with the same vacancy creation, the unemployment rate at the beginning of 2010 was 9.5%. Most people (including the NBER) would argue that 2010 was a solid recovery, yet unemployment was higher.

above. Responsible for the shifting Beveridge Curve is therefore a large difference across equilibria in the *endogenous argument* of the matching function – i.e., the effective market tightness – and not the exogenous matching technology.

To validate our mechanism further, we end by providing direct evidence for two features that underlie the model's endogenous fluctuations and that are responsible for generating our main results: (i) procyclical search intensity of on-the-job searchers; (ii) cyclical composition shifts in both the pool of searchers and the pool of employed workers across the job ladder.

Related Literature. We are intellectually indebted to several earlier contributions and ideas that have shaped our thinking on this topic. A pioneer of self-fulfilling employment fluctuations is Diamond (1982).[2] His model features multiplicity due to a thick market externality from increasing returns to scale in the matching technology: the more people search, the higher the probability of trading. While our source of multiplicity is similar since it also stems from endogenous behavior affecting the matching function, we do not rely on increasing returns to matching, a counterfactual feature of the matching technology (Pissarides and Petrongolo (2001)). Like Diamond (1982), our model has Keynesian elements in the sense that beliefs can generate business cycles. In contrast, most literature on the cyclical implications of labor search theory is neoclassical (fluctuations are driven entirely by productivity shocks).

The source of multiplicity in our model is also related to Burdett and Coles (1997). Their driving force is not Diamond (1982)'s market size externality, but rather a *selection externality* that affects the steady state distribution of heterogenous types. If high types believe other high types are not selective and also accept matches with low types, the equilibrium distribution of searchers will have a low fraction of high types and hence it pays off to be non-selective oneself. While the composition effects of active OJS in our model are somewhat similar to this selection externality, that model has quite different predictions: In Burdett and Coles (1997) it is difficult to map the two equilibria into boom and recession. In their selective equilibrium, mismatch is low and output is high (as in a boom), but unemployment is *high* as well, whereas in the non-selective equilibrium, mismatch is high and output is low (as in the recession), but unemployment is *low*.

Kaplan and Menzio (2016) ask the archetypical Keynesian question whether externalities in the goods market can affect employment in the labor market. In their model, if unemployment is believed to be high, then demand for goods will be low and more workers search for low prices, both leading to less production and thereby to high unemployment.[3] In an interesting approach that also features a demand externality, Schaal and Taschereau-Dumouchel (2015) have a model with multiple equilibria (without search) but focus on equilibrium selection using global games. While this guarantees a unique

[2] Diamond and Fudenberg (1989) further analyze the non-stationary rational expectations equilibrium of this model.

[3] Howitt and McAfee (1992) address a similar question. See also Shleifer (1986) for a model with multiplicity and output fluctuations through the timing of bringing innovation to the market.

equilibrium, it maintains the amplification through multiple steady states. Schaal and Taschereau-Dumouchel (2016) embed this mechanism into a random search model and show that it can quantitatively account for the volatility and persistence of labor market variables in the US.

The difference between this literature and our model is that, rather than exogenous demand externalities in the goods market, our feedback mechanism originates in labor market itself: it is based on a strategic complementarity between OJS and vacancy posting and consistent with the well documented pro-cyclicality of EE flows in the data.[4] Not only do we find that pro-cyclical on-the-job search can account for large labor market fluctuations, our mechanism also rationalizes the jobless recovery, the shift in the Beveridge Curve during recovery and pro-cyclical fluctuations in frictional wage dispersion.

Self-fulfilling multiple equilibria in search models have been used beyond the labor market: Burdett, Imai, and Wright (2004) build a marriage market model where multiplicity stems from the strategic interaction of partners' on-the-match search *within a match*. Moen, Nenov, and Sniekers (2015) have a model of the housing market where multiplicity arises because homeowners who switch houses coordinate whether to sell their current house before or after they buy the new house. These are interesting approaches but differ from our work in terms of mechanism and objectives.

Last, in his seminal paper, Shimer (2005) argues that in the standard Diamond-Mortensen-Pissarides (DMP) model of unemployment, productivity fluctuations cannot account for the fluctuations in unemployment and vacancies observed in the data.[5] Hall (2005) (wage stickiness) and Hagedorn and Manovskii (2008) (the high value of unemployment) have offered explanations to counter Shimer's finding, and can indeed create labor market volatility from small productivity shocks. We do not see our contribution in adding to this debate. But we note that our model generates considerable amplification by relying on *endogenous reallocation* of workers to jobs with higher match-specific productivity in the boom and without alluding to any shock in the fundamentals (like Total Factor Productivity).

The paper is organized as follows. Section 2 introduces the model. Section 3 analyzes multiple steady state equilibria and Section 4 multiple dynamic equilibria. Section 5 contains a stylized quantitative exercise and provides additional empirical evidence for the model mechanism. Section 6 concludes.

2 Environment

We build a model of random search where workers search both when unemployed and employed. It features a stylized two-step job ladder: we assume that all jobs found out of unemployment have low

[4] The only paper that also obtains fluctuations driven exclusively by the labor market is Golosov and Menzio (2015). The model features moral hazard where it is most efficient to provide incentives through firing during recessions. Interestingly, this requires decreasing, not increasing returns to matching. Both their model and questions are different from ours.

[5] In a theory of rest and search unemployment, a variation of the DMP search model, Jovanovic (1987) shows that productivity fluctuations also generate pro-cyclical search behavior (in addition to pro-cyclical productivity and counter-cyclical unemployment) as here, but without the amplification from equilibrium multiplicity that we highlight.

match productivity, and all jobs found out of an existing job have high match productivity. This stylized set-up aims to capture the main forces of search models with OJS and sorting in a tractable way: a worker who already has a job will only move to a new job if the new job is more productive. Therefore, the types of job matches out of unemployment are on average less productive than those that form when moving from an existing job and, as a consequence, firms prefer hiring employed workers.[6]

Agents and Technology. Time is continuous, $t \in [0, \infty)$. There is a measure one of risk neutral workers in the economy. A worker is unemployed and searching for a job, or employed, in which case she can choose to search *actively* or *passively* on-the-job. We assume that OJS only takes place in low productivity jobs (see Online Appendix I.2 for a model that relaxes this assumption but preserves the key mechanism).[7] The flow utility from being unemployed is b and the flow utility of employment is equal to the wage, w_t. The search cost when unemployed or under *passive* OJS is normalized to zero and the search cost for *active* search when employed is k, so costs of OJS increase in search intensity. Workers maximize the value of employment: they search actively if the gain from active search exceeds the cost. Otherwise, they search passively at no cost.

There is a large measure of potential jobs (or firms). Firms can open a job/vacancy by paying a flow cost c. If they stay inactive their payoff is zero. Firms are risk neutral and maximize the discounted sum of profits. Denote the measure of job openings by v_t. All jobs are ex ante identical, but ex post heterogeneous in their job productivity y. We assume the technology is given by $f(y) = py$, where p is aggregate and $y \in \{\underline{y}, \overline{y}\}$ is match-specific productivity.[8] When a job is filled by an unemployed worker, the productivity is $\underline{y}$ and when it is filled by a formerly employed worker the productivity is $\overline{y}$, with $\underline{y} < \overline{y}$.[9] This captures in a stylized way the economy's job ladder: Workers tend to be better matched after they switch jobs. This is reflected in the data by substantial wage gains after EE transitions, even controlling for the earnings growth experienced by similar workers (Moscarini and Postel-Vinay (2017)). We model this job ladder as improvements in the match-specific component of a worker-firm pair.

Denote the measure of the unemployed by u_t; the measure of the employed in a low productivity job $\underline{y}$ by γ_t; and the measure of the employed in high productivity jobs $\overline{y}$ by ξ_t. Since the measure of

[6]We discuss this assumption that firms prefer hiring employed over unemployed workers in more detail in the *Remarks on the Assumptions* below, but already want to note that there are many plausible micro foundations for it: For example, productivity of employed workers is higher due to human capital accumulation on the job/learning by doing; or unemployed workers are less productive due to skill loss during unemployment; or employed workers are better able to direct their search and thus sort more effectively; or the differential search intensity by workers at lower vs. higher rungs of the job ladder leads to differential match duration, making firms prefer employed over unemployed workers.

[7]The implicit assumption is that search costs in high productivity jobs are too high relative to the gains, so that no more search occurs to increase the wage further after one round of OJS.

[8]Below we assume in addition that cost k and unemployment benefits b are proportional to p (consistent with Chodorow-Reich and Karabarbounis (2016) that the value of unemployment is pro-cyclical.

[9]If the surplus of a low type match is positive (an assumption we make), it is optimal for the firm to accept this match even if that surplus is lower than the surplus of a high type match.

workers is equal to one, feasibility requires that $u_t + \gamma_t + \xi_t = 1$.

Market Frictions and Search. Meetings between jobs and workers are stochastic, and are modeled by means of a standard matching function $m(v_t, s_t)$, where m is strictly increasing and concave in both arguments, and has constant returns to scale. Matching function m takes as inputs the measure of vacancies, v_t, and the measure of searchers, denoted by s_t (including employed and unemployed). Therefore the matching rate for a worker is $m(\theta_t)$, where $\theta_t = \frac{v_t}{s_t}$, and that of a firm is $q(\theta_t) = \frac{m(\theta_t)}{\theta_t}$. Job separation is exogenous and constant over time, occurring at Poisson rate $\delta_t = \delta$.

Employed workers always engage in passive search at no cost (some job opportunities arrive without search effort) which leads to a match at rate $\lambda_0 m(\theta_t)$, where $\lambda_0 > 0$ is the search intensity of passive on-the-job searchers relative to the search intensity during unemployment, which is normalized to one. They can also engage in active search at intensity $\lambda_0 + \lambda_1$ (with $\lambda_1 > 0$), by incurring the search cost k.[10] In return they get a higher chance of a match, $(\lambda_0 + \lambda_1)m(\theta_t)$. The total number of searchers is expressed in "efficiency units" weighted by search intensity so the effective measure of workers searching for a job, s_t, is given by $u_t + \lambda_0\gamma_t$ if all employed workers in $\underline{y}$-jobs search passively and by $u_t + (\lambda_0 + \lambda_1)\gamma_t$ if all employed workers in $\underline{y}$-jobs search actively. Thus, λ_0 and λ_1 reflect the efficiency of the matching technology on-the-job. The resulting market tightness is a function of the total measure of searchers: When all workers actively search on the job, the market tightness is given by $\theta_t = \frac{v_t}{u_t + (\lambda_0 + \lambda_1)\gamma_t}$ and when they only search passively on-the-job, it is given by $\theta_t = \frac{v_t}{u_t + \lambda_0\gamma_t}$. Notice that we distinguish the effective market tightness $\theta_t = \frac{v_t}{s_t}$ that takes into account *all* effective job searchers from the conventional market tightness, here denoted by $\Theta_t = \frac{v_t}{u_t}$, which only takes into account the unemployed searchers.

Individual Decision Problems and Bellman Equations. We denote the value of a vacant job by V_t, of a filled job by J_t, of an unemployed worker by U_t, and of an employed worker by E_t. When we denote the value of a job for an employed worker, we use the notation $\underline{E}_t$ ($\overline{E}_t$) to indicate that she is employed in a low (high) productivity job. Likewise, $\underline{J}_t$ ($\overline{J}_t$) denotes the value of a low (high) productivity job that is filled with a worker coming out of unemployment (out of a low productivity job). Denote by $\omega_t \in [0, 1]$ the decision by the individual worker whether to actively search on-the-job and by $\mathbf{\Omega}_t \in [\mathbf{0}, \mathbf{1}]$ the behavior of all workers in a symmetric strategy equilibrium (bold-face indicates from now on the behavior of the *aggregate* economy). Even though we also show under which conditions a mixed strategy equilibrium with interior ω_t and $\mathbf{\Omega}_t$ exists, we focus for the most part on pure strategies. That is, all agents in low productivity jobs either do search actively or they do not, hence $\mathbf{\Omega}_t \in \{\mathbf{0}, \mathbf{1}\}$ and $\omega_t \in \{0, 1\}$, where 1 (or $\mathbf{1}$) indicates active and 0 (or $\mathbf{0}$) passive search. Throughout we assume

[10] An alternative way of modeling this would be through *continuous* search intensity, where workers choose an interior non-zero search intensity under a convex cost. This could also give rise to multiple equilibria with a high and a low intensity of OJS. Unfortunately, we cannot solve that case analytically. Observe that our cost is a step function and hence convex.

that individual search behavior ω_t is private information, so a firm cannot make the wage contingent on search effort. For the remainder, we also use the notation $\lambda(\omega_t) = \lambda_0 + \omega_t\lambda_1$ for the individual search intensity and $\lambda(\mathbf{\Omega}_t) = \lambda_0 + \mathbf{\Omega}_t\lambda_1$ for the aggregate search intensity. Moreover, we make explicit that tightness is a function of search behavior as $\theta_t(\mathbf{\Omega}_t) = \frac{v_t}{s_t(\mathbf{\Omega}_t)} = \frac{v_t}{u_t+\lambda(\mathbf{\Omega}_t)\gamma_t}$, and similarly for searchers $s_t(\mathbf{\Omega}_t)$ and wages $w_t(\mathbf{\Omega}_t)$, which we discuss in detail below. Note that all values and other endogenous variables are functions of search behavior as well, but for the most part we suppress this dependence.

Workers: We can write the values of a worker as follows.

$$rU_t = pb + m(\theta_t(\mathbf{\Omega}_t))(\underline{E}_t - U_t) + \dot{U}_t \tag{1}$$

$$r\underline{E}_t = \underline{w}_t(\mathbf{\Omega}_t) - \omega_t pk + \lambda(\omega_t)m(\theta_t(\mathbf{\Omega}_t))(\overline{E}_t - \underline{E}_t) - \delta(\underline{E}_t - U_t) + \dot{\underline{E}}_t \tag{2}$$

$$r\overline{E}_t = \overline{w}_t(\mathbf{\Omega}_t) - \delta(\overline{E}_t - U_t) + \dot{\overline{E}}_t \tag{3}$$

where $\dot{U}_t$ is the time derivative of U_t (and similarly for $\dot{\underline{E}}_t$ and $\dot{\overline{E}}_t$).

Importantly, *individual* search decisions, ω_t, affect only the value of the employed in low productivity jobs, $\underline{E}_t$, namely through the cost of job search k and the increased rate of finding a job by λ_1. *Aggregate* search behavior from the population at large, $\mathbf{\Omega}_t$, enters the values through two channels: It affects the job finding probabilities of workers through market tightness, $\theta_t(\mathbf{\Omega}_t)$, and thereby the value of the employed in a low productivity job as well as the value of the unemployed. It also affects wages, $w_t(\mathbf{\Omega}_t)$ which depend on the belief whether workers search actively on-the-job or not.

Firms: The value of a vacancy to a firm is given by,

$$rV_t = -c + q(\theta_t(\mathbf{\Omega}_t))\left[\frac{u_t}{s_t(\mathbf{\Omega}_t)}\underline{J}_t + \frac{\lambda(\mathbf{\Omega}_t)\gamma_t}{s_t(\mathbf{\Omega}_t)}\overline{J}_t - V_t\right] + \dot{V}_t. \tag{4}$$

reflecting the expected value of filling a vacancy (either with an unemployed worker which occurs at rate qu/s or with an employed searcher which occurs at rate $q\lambda\gamma/s$), net of vacancy posting cost c. Because we assume free entry and a large measure of potential entrants, the value of a vacancy V_t will be driven to zero in equilibrium. Observe that the measure of vacancies adjusts instantaneously: Whenever V_t is positive, vacancies are created frictionlessly to set the expected profits back to zero.

The values of a filled low and high productivity job to the firm are given by:

$$r\underline{J}_t = p\underline{y} - \underline{w}_t(\mathbf{\Omega}_t) - [\lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)) + \delta](\underline{J}_t - V_t) + \dot{\underline{J}}_t \tag{5}$$

$$r\overline{J}_t = p\overline{y} - \overline{w}_t(\mathbf{\Omega}_t) - \delta(\overline{J}_t - V_t) + \dot{\overline{J}}_t. \tag{6}$$

The flow value of a high type job in (6) is output net of wages. Once it is filled, the job lasts until there is exogenous separation at rate δ. The low type job in (5) similarly generates a flow value of $p\underline{y} - \underline{w}$ and separates exogenously at rate δ, but in addition faces separation from OJS, which happens at rate λm.

Wage Setting. Wages are determined as in the sequential auction framework by Postel-Vinay and Robin (2002) (see also Dey and Flinn (2005)). Employment contracts stipulate a fixed wage over time that the employer commits to and that can be renegotiated only if both parties agree. Firms can fire workers and workers are free to quit at will. As a result, when workers receive outside offers, wages may be renegotiated: Current and outside employers Bertrand-compete for the worker. The worker goes to the match that generates a higher total match value and receives a wage such that her continuation value equals the match value with the least productive of the two competing firms (i.e. the match value of her outside option). If no outside offer arrives, wages remain unchanged. If the worker is hired out of unemployment, wages are such that she receives the option value of unemployment.

A firm hiring an unemployed worker will thus offer a wage $\underline{w}_t$ that makes her indifferent between accepting the job and remaining unemployed, $\underline{E}_t = U_t$. Likewise, the firm offers a wage $\overline{w}_t$ to an employed worker such that she is indifferent between joining the new firm with high productivity job $\overline{y}$ and staying at the old firm in low productivity job $\underline{y}$. Hence, the new firm pays the worker the highest wage that the previous firm could have paid, pinned down by $\underline{J}_t = V_t$ in the previous firm. This matching of outside offers yields the following wages in low and high productivity jobs:

$$\begin{aligned} \underline{w}_t(\mathbf{\Omega}_t) &= p\left[b\left(\frac{r+\lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))+\delta}{r+\delta}\right) - \frac{\lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))}{r+\delta}\underline{y} + \mathbf{\Omega}_t k\right] && (7) \\ \overline{w}_t(\mathbf{\Omega}_t) &= p\underline{y}. && (8) \end{aligned}$$

Observe that wages reflect the population-wide behavior of on-the-job searchers $\mathbf{\Omega}_t$ and not the individual level search behavior ω_t. That is, the wage reflects the firm's belief about the workers' search behavior but cannot condition on the actual (unobserved) search behavior of the particular worker that is hired. Note that $\overline{w}_t(\mathbf{\Omega}_t)$ is a constant and thus time-invariant even out of steady state.

Labor Market Dynamics. At any point in time, the laws of motion for unemployment and employment across the job ladder satisfy:

$$\begin{aligned} 1 &= u_t + \gamma_t + \xi_t && (9) \\ \dot{\gamma}_t &= u_t m(\theta_t(\mathbf{\Omega}_t)) - \gamma_t[\delta + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)] && (10) \\ \dot{\xi}_t &= \gamma_t\lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)) - \xi_t\delta && (11) \end{aligned}$$

Equation (9) ensures feasibility: the measure of workers consists of unemployed u_t, on-the-job searchers who work in low productivity jobs γ_t, and workers who obtained their high productivity job through OJS ξ_t, and is equal to the measure of the entire worker population. In equation (10), the change in the stock of on-the-job searchers equals the difference between the flow into low productivity jobs from unemployment and the flow out of low productivity jobs, which consists of separations at rate δ and the outflow due to OJS at rate λm. Finally, the change in the stock of workers in high productivity jobs equals the difference between in- and outflow from high productivity jobs, given by equation (11).

Definition of equilibrium. We can now define equilibrium.

Definition 1. *For a given sequence $\{\mathbf{\Omega}_t\}_{t\geq 0}$, a Perfect Foresight Equilibrium is a path $\{U_t, \underline{E}_t, \overline{E}_t, V_t, \underline{J}_t, \overline{J}_t, \theta_t, u_t, \gamma_t, \xi_t, \underline{w}_t, \overline{w}_t, \omega_t, \}_{t\geq 0}$ such that for all $t \in [0, \infty)$:*

1. *$U_t, \underline{E}_t, \overline{E}_t, V_t, \underline{J}_t, \overline{J}_t$ satisfy the Bellman equations (1)-(6);*
2. *Given $\{\mathbf{\Omega}_t\}_{t\geq 0}$, $\omega_t = \mathbf{\Omega}_t$ maximizes $\underline{E}_t$;*
3. *There is free entry of firms: $V_t = 0$;*
4. *Wages: $\underline{w}_t$ is such that $\underline{E}_t = U_t$ and given by (7); $\overline{w}_t$ is such that $\underline{J}_t = V_t$ and given by (8);*
5. *u_t, γ_t, ξ_t satisfy the laws of motion (9)-(11);*
6. *$\lim_{t\to\infty} \theta_t$ is finite and initial conditions u_0, γ_0, ξ_0 are given.*

Note that in this equilibrium definition, we assume that the sequence $\{\mathbf{\Omega}_t\}_{t\geq 0}$ is deterministic and that the agents never anticipate a deviation from this deterministic path of aggregate OJS, i.e. the agents have perfect foresight.

3 Steady State Equilibrium

We first focus on steady state equilibrium where we assume that beliefs about the profitability of OJS and thus search behavior is constant over time, that is $\mathbf{\Omega}_t = \mathbf{\Omega}$ for all $t \in [0, \infty)$. We solve the system of equilibrium equations, where we set time derivatives to zero and drop time subscripts. First we solve for wages, $\underline{w}$ in (7) and $\overline{w}$ in (8), then we pin down θ from free entry where firms take wages and search behavior as given. Finally, for given θ, we determine the stocks u, γ and ξ from the steady-state flow-balance equations. This guarantees that all but requirement 2. from Definition 1 are satisfied. In what follows, we therefore analyze conditions under which requirement 2. is also satisfied, i.e. under which there is no profitable deviation in *individual* on-the-job search behavior ω from the *aggregate* search behavior $\mathbf{\Omega}$. At the same time, we establish under which conditions *both* steady states coexist.

3.1 Multiplicity

We construct two candidate steady state equilibria in which either no employed worker in a low productivity job searches actively, $\mathbf{\Omega} = \mathbf{0}$, or all workers search actively in such jobs, $\mathbf{\Omega} = \mathbf{1}$. For a steady-state to exist, it is sufficient to check that *one-shot deviations* by a firm or a worker are not profitable. To exclude the firms' one-shot deviations is straightforward since firms only have a participation decision to make and, in the presence of free entry, this yields zero profits (if they do not participate they also make zero profits). Note that, in our current setup, we have restricted the contract space to constant wages until the arrival of an outside offer, as is customary in this literature. In Online Appendix II we show that even if firms can offer a wage contract with backloading, they nonetheless do not want to deviate from constant wages under natural parameter restrictions and there continue to be multiple steady state equilibria.

On the worker side, it is sufficient to check one-shot deviations from the worker's strategy in the low-productivity job. The value of unemployment is pinned down by the exogenous flow benefits b. As a result, the value of unemployment U is independent of the worker's search intensity ω. Likewise, the worker's value of being employed in a high productivity job $\overline{E}$ is independent of search behavior since there is no search decision at the top rung of the job ladder. This implies that U and $\overline{E}$ are independent of the search decision ω, and we can directly check the deviations of those who are employed in low productivity jobs and who obtain $\underline{E}$.

To evaluate deviations by an individual worker, we introduce the following notation. If a worker in a low productivity job deviates from the search behavior of all others for an instant dt and then reverts to the equilibrium behavior $\mathbf{\Omega}$, we denote his value by $\underline{E}(\omega|\mathbf{\Omega})$ with $\omega \neq \mathbf{\Omega}$. This captures the notion of the one-shot deviation principle, or equivalently Bellman optimality. In turn, the value of a worker who does not deviate is $\underline{E}(\omega|\mathbf{\Omega})$ where $\omega = \mathbf{\Omega}$.

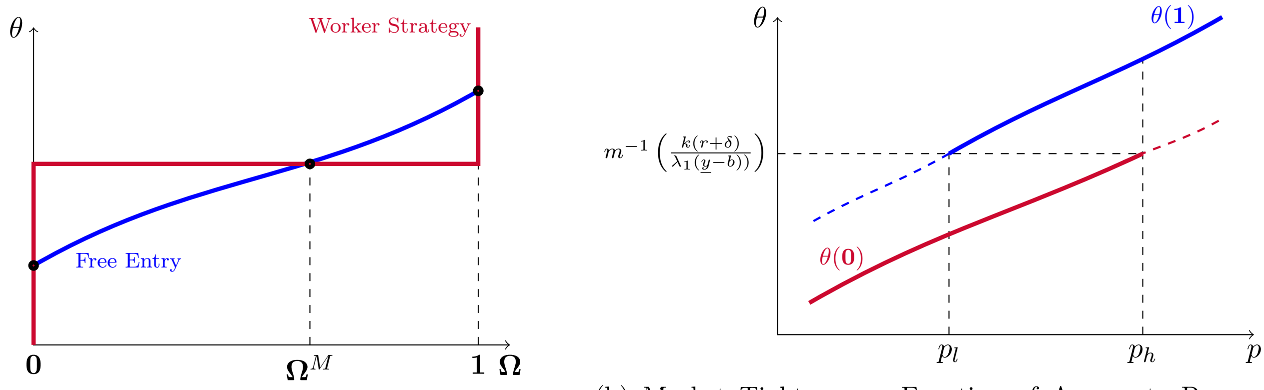
We now check two possible deviations and derive conditions under which those deviations are not individually rational: (i) when all workers in low productivity jobs are actively searching on-the-job, there is no deviation by a single agent to stop active search if:

$$\underline{E}(1|\mathbf{1}) \geq \underline{E}(0|\mathbf{1}) \quad \Longleftrightarrow \quad m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right) \leq \theta(\mathbf{1}).$$

(ii) when no worker is actively searching on-the-job, there is no deviation of a single agent to start active search if:

$$\underline{E}(0|\mathbf{0}) \geq \underline{E}(1|\mathbf{0}) \quad \Longleftrightarrow \quad \theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right).$$

These two no-deviation conditions give rise to the following result.



(a) Steady State Equilibria: No Active OJS $\mathbf{\Omega} = \mathbf{0}$; Active OJS $\mathbf{\Omega} = \mathbf{1}$; Mixed Strategy $\mathbf{\Omega}^M \in (\mathbf{0}, \mathbf{1})$.

(b) Market Tightness as Function of Aggregate Productivity: Equilibria (Solid Segments) and Multiplicity Range $p \in [p_l, p_h]$.

Figure 2: Multiple Steady State Equilibria.

Lemma 1. *There exist multiple steady state equilibria if and only if*

$$\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right) \leq \theta(\mathbf{1}).$$

All proofs are in Appendix A. Under the condition that the market tightness is not too extreme, there exist two pure strategy steady state equilibria, one where all workers in low productivity jobs search actively and one where no one searches actively. We show in the Appendix that whenever the two pure strategy equilibria exist, there is also a mixed strategy equilibrium where every agent searches actively on-the-job with probability $\omega = \mathbf{\Omega} \in (\mathbf{0}, \mathbf{1})$, i.e. in every interval of time dt workers randomize between the choice of search effort (see Appendix A.3 for the formal statement).[11] This is illustrated in Figure 2a, where we plot the mutual best-responses of workers' search effort and firms' vacancy posting (reflected by labor market tightness). The workers' best response to tightness is an increasing step function and the firms' best response of tightness to workers' search effort is an increasing function as well, indicating the strategic complementarity between search effort and vacancy posting. The intersections at $\mathbf{\Omega} = \mathbf{0}$ and $\mathbf{\Omega} = \mathbf{1}$ mark the pure strategy steady state equilibria while the interior intersection indicates the mixed strategy steady state. In what follows, we focus attention on the two pure strategy steady states.

Of course, tightness $\theta(\mathbf{\Omega})$ is an endogenous object. We now provide a necessary and sufficient condition for multiplicity in terms of the primitives of the model.

Proposition 1 (Aggregate Productivity and Multiplicity)**.** *There are multiple steady state equilibria if and only if aggregate productivity p is such that $p \in [p_l, p_h]$. The interval $[p_l, p_h]$ is non-empty for an open set of parameters $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$.*

[11] Given a continuum of agents, this mixed strategy equilibrium is equivalent to purification where a fraction chooses the pure strategy $\omega = 1$ and the remainder chooses 0.

See Appendix A for the exact expressions for the productivity bounds $[p_l, p_h]$, which are complicated functions of the model's remaining parameters. This result rewrites the necessary and sufficient condition for multiple steady state equilibria from Lemma 1 as a condition on the exogenous productivity parameter p that needs to lie in a certain interval (which we show is not empty). The intuition is straightforward: If aggregate productivity is too high, $p > p_h$, then all workers in low productivity jobs want to search actively to take advantage of jobs with high match-specific component whose productivity $\overline{y}$ is now augmented by high aggregate productivity p. The passive search equilibrium breaks down and the active search equilibrium is unique. The opposite occurs if productivity is too low, $p < p_l$. In Figure 2b we illustrate the multiplicity region, by plotting equilibrium tightness θ as a function of aggregate productivity p. Market tightness is always increasing in productivity, both with and without active OJS, but for any given value of p, θ is higher with active OJS. The solid θ-segment indicates the productivity range for which a certain steady state exists. If $p \in [p_l, p_h]$, then both steady states exist, corresponding to the condition in Lemma 1 that $\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right) \leq \theta(\mathbf{1})$ (see y-axis, Figure 2b).

This condition for multiplicity can also be expressed in terms of any of the exogenous variables other than p. We want to highlight one more of these conditions, which shows that the existence of multiple equilibria is closely related to the gains from sorting, i.e. to the difference $\overline{y} - \underline{y}$.

Proposition 2. *(Gains from Sorting and Multiplicity). There are multiple steady state equilibria if and only if $\overline{y} \in [\overline{y}_l(\underline{y}), \overline{y}_h(\underline{y})]$ for each $\underline{y}$. The interval $[\overline{y}_l(\underline{y}), \overline{y}_h(\underline{y})]$ is non-empty for an open set of parameters $(\lambda_0, \lambda_1, \underline{y}, p, k, c, b, r, \delta)$.*

For sufficiently low productivity gains from sorting, or equivalently low gains from OJS (measured by $\overline{y} - \underline{y}$), there is a unique equilibrium with no active OJS. In this case, the costs of OJS given by the direct search cost k incurred by the worker and the indirect search cost incurred by the firm due to shorter expected duration of a job outweigh the productivity gains from OJS. Hence, firms post few vacancies and the dominant strategy is not to search actively.

At the other extreme, when productivity gains from OJS are sufficiently large, then the gains from OJS swamp its costs. Then the dominant strategy of employed workers is to search actively and firms' vacancy posting is high.

Our results in this section illustrate the mechanism that gives rise to the strategic complementarity between worker and firm behavior and hence to multiplicity when aggregate productivity and the gains from sorting are not too extreme. Firms trade off the expected quality or productivity of a job (which can vary endogenously due to the composition externality) against job duration. And workers trade off the matching rate against the cost of searching. With active OJS, there is more sorting, the composition of the pool of searchers improves, and the value of a job to any firm is higher, which creates incentives for vacancy posting. More vacancies in turn create incentives for workers to actively search on-the-job since it is easier to find a job. This gives rise to a steady state equilibrium with active search. There is also an equilibrium where workers do not actively search on-the-job, where the pool of searchers has relatively few on-the-job searchers and is of relatively low match quality. For firms, the shorter duration of jobs filled under OJS then dominates the impact of the composition externality, and as a result, they post few vacancies. This indeed leads workers to not search actively.

Remarks on the Assumptions. Before analyzing the equilibrium properties, we discuss some of our main assumptions and their role for the multiplicity of steady state equilibria.

First, for tractability we assume in our baseline model that there is a two-step job ladder, where on-the-job searchers receive a *deterministic* match-specific productivity upgrade. This assumption is consistent with the evidence that unemployed workers accept lower-quality jobs offers (Faberman, Mueller, Sahin, and Topa (2017)) and that there is substantial wage growth as workers climb up the job ladder (see for example Faberman (2015), Haltiwanger, Hyatt, and McEntarfer (2015), and Gertler, Huckfeldt, and Trigari (2016)), which is pro-cyclical (Haltiwanger, Hyatt, Kahn, and McEntarfer (2018)). This is supporting evidence that our simplified job ladder is not a poor approximation. It could be rationalized by several micro-foundations (e.g. human capital accumulation/learning by doing; adverse selection; differential sorting/directed search of employed and unemployed workers) but exploring them in depth would require an entirely different model and go beyond the scope of this paper. Instead, we analyze in the Online Appendix the most natural generalization of our model, which is to dispense with the reduced-form job ladder and introduce *stochastic* productivity upgrades, where both unemployed and employed workers receive them with the same probability and also can search for an unrestricted number of rounds (Online Appendix I.1).[12]

What transpires from this exercise is that a similar strategic complementarity between OJS and vacancy posting generates multiple steady state equilibria in more general environments. This shows that the multiplicity is *not only* due to the specific job ladder we assume but roots more deeply in the interplay between composition externality and job duration. Even though overall job duration is lower under active search, the potential productivity gain and the fact that a match is of longer duration when formed with an employed compared to an unemployed searcher, make on-the-job searchers for firms attractive. Like in our baseline model, this triggers a strategic complementarity between search intensity and vacancy posting.Thus our simple model inherits all the important features of the more

[12] This extension is similar to Postel-Vinay and Robin (2002) (but with two job types) where heterogenous job offers randomly arrive to both employed and unemployed workers (see also models of sorting with OJS and a continuum of types that are computationally solved by Lise and Robin (2017) and Lamadon, Lise, Meghir, and Robin (2013)). In this model, there are many more candidate equilibria (namely $2^4 = 16$), depending on various choices of search intensity at different parts of the job ladder. In principle, one needs to check no-deviation conditions for each of these candidate equilibria. We pick two specific equilibria out of the 16 candidate ones and show that they can co-exist for certain parameter values. We also analyze a second extension, which is even closer to the baseline model, where we allow for more than one round of OJS while keeping our deterministic productivity upgrades (Online Appendix I.2).

general setup but has the benefit of being considerably more tractable (in the generalized model, there are $2^4 = 16$ potential equilibria, depending on various choices of search intensity at different parts of the job ladder). Below, we make the deliberate choice to bring our baseline model to the data instead of the generalized one, since it would be very difficult to estimate that model while ensuring both the existence of two particular equilibria (eight no-deviation conditions have to be satisfied, instead of two in our baseline model) while excluding the possibility that any of the other fourteen equilibria co-exists.

Second, we assume in our baseline model that productivity is match-specific. Online Appendix I.3 shows that multiplicity of equilibria does not hinge on this assumption either. There we introduce permanent *ex-ante* productivity differences of firms, i.e. firms can either open a low or a high productive vacancy. Employed and unemployed workers meet both types of vacancies with the same probabilities.

Third, we assume that unemployed workers (as opposed to employed workers) do not choose their search intensity endogenously, for three reasons: 1. The empirical studies on how search intensity of the unemployed varies over the cycle are inconclusive. There is evidence on both (slightly) counter-cyclical search intensity (Mukoyama, Patterson, and Şahin (2018)) and pro-cyclical search intensity (Schwartz (2014)). 2. Note that our results would go through if we interpreted the OJS intensity *relative* to the unemployed search intensity, and this is in fact the interpretation we should adopt based on our empirical evidence below.[13] 3. In our model, even if we introduced endogenous search intensity of the unemployed, they would always (independent of the business cycle) choose a *unique* level of search intensity. This is due to the sequential auction wage setting without worker bargaining power where the value of unemployment is constant and so are the gains from search during unemployment.

Fourth, we do not include the flows of those Not in the Labor Force (NiLF) because the counter-cyclicality of their search effort is unlikely to be a confounding factor to our mechanism.[14]

Fifth, we assume that both employed and unemployed workers randomly search for jobs in a single labor market. For our mechanism to work, there cannot be completely segregated labor markets since in that case the discussed composition externality would be shut down (but completely segregated/directed search markets would be a questionable assumption as well).

Sixth, our mechanism that generates multiplicity does not hinge on the contractual setting with fixed wages. It is known that in the presence of endogenous OJS, commitment to a fixed wage can be improved upon with a time varying contract (see for example Lentz (2014)). We show in Online Appendix II that even if firms can deviate from a fixed wage to a simple contract with backloading, there is multiplicity.

[13]If the unemployed searched with intensity $\lambda_u \neq 1$, then the flow from unemployment to employment would read $UE = \lambda_u um$. Therefore, the search intensity of the employed (implied by the UE and EE flows) would read: $\lambda = \frac{EE}{UE}\frac{u}{\gamma}\lambda_u$. Since below we use $\frac{EE}{UE}\frac{u}{\gamma}$ to compute a measure of search intensity of the employed, what we effectively obtain is a measure of search intensity of the employed λ *relative* to the search intensity of the unemployed λ_u, and we will show in Section 5.4 that this relative measure is higher in booms than in recessions.

[14]While these flows are sizable, the cyclical properties of the search intensity of those NiLF is the opposite of that of the employed: counter-cyclical. In Appendix C.1, Figure 13b, we report a measure of their search intensity.

Last, we assume that separations are constant in the model. This is clearly not borne out in the data, see Fujita and Ramey (2009). However, our focus is on how the interaction between search intensity of the employed and vacancy creation by firms can generate multiplicity, which is why we make the simplifying assumption of exogenous and constant separations.

3.2 Properties

In the standard random search model without OJS (e.g. Pissarides (2000)), the steady state allocation is at the intersection of the Beveridge Curve (BC), i.e. the flow-balance condition of unemployment,

$$u = \frac{\delta}{\delta + m(\theta(\mathbf{\Omega}))} \tag{BC}$$

and the free entry condition in the (u, v)-space. Here, however, since matching probabilities are a function of the *effective* market tightness $\theta = \frac{v}{u+\lambda\gamma} = \frac{v}{s}$, the Beveridge Curve also depends on the stock of on-the-job searchers γ, given by flow-balance condition (10), which we label the γ-Curve or (γC):

$$\gamma = \frac{\delta m(\theta(\mathbf{\Omega}))}{[\delta + m(\theta(\mathbf{\Omega}))][\delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))]}. \tag{γC}$$

To find the steady states in our model, we therefore plot the equilibrium system of three equations, given by free entry (4) with $V = 0$, (BC) *and* (γC), in terms of u, v and γ, where v is a transformation of θ: $v = \theta(u+\lambda\gamma)$. Both (BC) *and* ($\gamma$C) give vacancies v as a function of (u, γ). For the sake of clarity, we combine these equations in Figure 3 and plot their intersection BC $\cap$ γC, which gives v as a function of searchers $s = u + \lambda\gamma$ and thus is the *effective Beveridge Curve* that takes all searchers into account:

$$s = u + \lambda(\mathbf{\Omega})\gamma = \frac{\delta(\delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))) + \delta\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))}{(\delta + m(\theta(\mathbf{\Omega})))(\delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))} \tag{BC^s}$$

Note that (BC^s) has similar properties as (BC) (it is downward sloping and convex).

The intersection between effective Beveridge Curve (BC^s) and free entry condition (4) (with $V = \dot{V} = 0$) marks the steady state for a given $\mathbf{\Omega}$. In Figure 3, we plot both the steady state equilibrium for active ($\mathbf{\Omega} = \mathbf{1}$, blue circle) and for passive OJS ($\mathbf{\Omega} = \mathbf{0}$, red circle) but omit the one in mixed strategies for clarity. Our next result compares the properties of the multiple steady states whenever they coexist.

Proposition 3 (Properties of Steady States)**.** *Let there be multiple steady state equilibria. Then:*

1. *conventional market tightness is higher with active OJS:* $\Theta(\mathbf{1}) \geq \Theta(\mathbf{0})$*;*

2. *unemployment is lower with active OJS:* $u(\mathbf{1}) \leq u(\mathbf{0})$*;*

3. *the measure of vacancies is higher with active OJS:* $v(\mathbf{1}) \geq v(\mathbf{0})$*;*

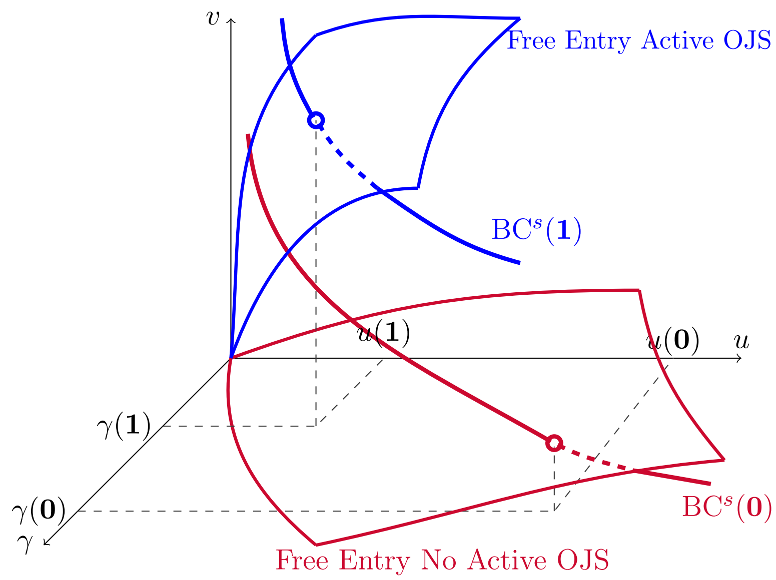


Figure 3: Pure Strategy Steady States at the Intersection of Free Entry Plane and Beveridge Curve (BC^s).

4. *EE flows, defined as* $EE = \lambda\gamma m$, *are higher with active OJS:* $EE(\mathbf{1}) \geq EE(\mathbf{0})$;

5. *the share of on-the-job searchers in all searchers increases with active OJS:* $\frac{\lambda(\mathbf{1})\gamma(\mathbf{1})}{s(\mathbf{1})} \geq \frac{\lambda(\mathbf{0})\gamma(\mathbf{0})}{s(\mathbf{0})}$;

6. *the share of on-the-job searchers in employed workers decreases with active OJS:* $\frac{\gamma(\mathbf{1})}{1-u(\mathbf{1})} \leq \frac{\gamma(\mathbf{0})}{1-u(\mathbf{0})}$;

7. *the Mean-min wage ratio,* $Mm = \frac{\frac{\gamma}{1-u}\underline{w}+\frac{\xi}{1-u}\overline{w}}{\underline{w}}$, *increases with active OJS (for* k *sufficiently small);*

8. $BC(\mathbf{1})$ *is shifted outward relative to* $BC(\mathbf{0})$;

9. $BC^s(\mathbf{1})$ *is shifted outward relative to* $BC^s(\mathbf{0})$ *(given* $\lambda(\mathbf{1}) \leq 1$*).*

Many of the features of this proposition can be observed in Figure 4. It plots the conventional Beveridge Curve (BC) that relates vacancies v to unemployment u with the standard market tightness $\Theta = \frac{v}{u}$, for both equilibria. (Plotting the effective Beveridge Curve (BS^s) in (s, v) space looks qualitatively identical). Similar to Lemma 1 that was stated in terms of *effective* market tightness, if *conventional* market tightness under active OJS is high enough (intersecting with the bold part of the blue Beveridge Curve), then this equilibrium exists. In turn, if market tightness under passive OJS is low enough (intersecting with the bold part of the red Beveridge Curve), then the equilibrium with low search intensity exists. Thus, when multiple steady states exist, not only $\theta(\mathbf{1}) \geq \theta(\mathbf{0})$ (Lemma 1) but also $\Theta(\mathbf{1}) \geq \Theta(\mathbf{0})$ (part 1. in Proposition 3).

Vacancies are higher under active search (3.): there are relatively more effective on-the-job searchers who generate a high productivity match (5.), meaning that the pool of searchers is of better quality.

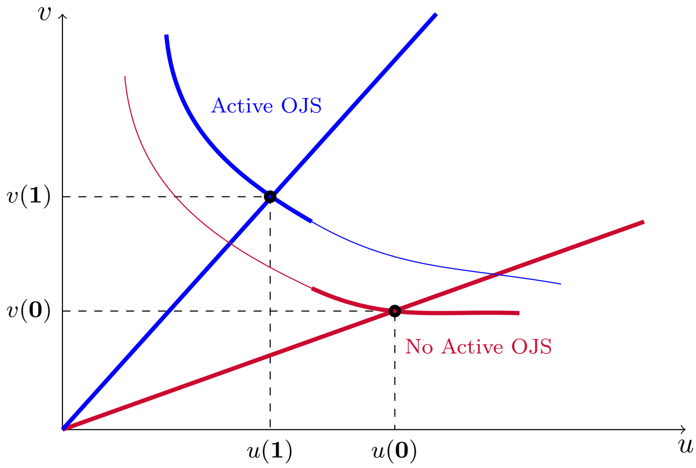


Figure 4: Multiplicity in the Beveridge Curve Diagram: The Conventional (BC) and Tightness $\Theta = \frac{v}{u}$

This increases firms' incentives to open vacancies and also leads to larger EE flows under active search (4.). Despite the lower match efficiency, unemployment is lower under active OJS (2.). This follows immediately from the flow equation for unemployment (BC) and the fact that under multiplicity $\theta(\mathbf{1}) \geq \theta(\mathbf{0})$ (from Lemma 1): The matching rate increases while job separation is unchanged.

The conventional Beveridge Curve shifts out under active OJS (8.). There are more vacant jobs under active OJS, pushing up the matching rate, but at the same time on-the-job searchers crowd out the unemployed. Hence, the match efficiency per unemployed worker is lower. Note that (BC^s) in the (s, v)-space also shifts out under active OJS (9.): Since the pool of searchers is of higher quality, the same measure of searchers encourages more vacancy posting, leading to a larger measure of vacancies.

Last, the distribution of workers across the job ladder changes with active OJS: less workers are stuck at the lowest rung of the ladder (6.), which is also the main force behind the increase in frictional wage dispersion (7.), as measured by the mean-min wage ratio (i.e. the ratio between the average wage and the lowest accepted wage or simply, Mm).

4 Dynamic Equilibrium

So far we have focussed on steady state equilibrium, showing when it is unique and when there are multiple equilibria. We now turn to the analysis of dynamic equilibria. We first ask whether, starting at initial values outside a certain steady state, there exists *a* path that leads to that steady state. Second, we want to understand whether *multiple* dynamic equilibria exist. More specifically, we are interested in whether, starting from the equilibrium with passive (active) OJS, an unanticipated permanent switch in agents' beliefs about other workers' OJS behavior puts the economy on a path to the steady state with active (passive) search.

We reduce the model's dynamic system to three equations and three unknowns, the two state variables u_t and γ_t as well as the choice variable θ_t (see Appendix A.7 for the derivations):

$$\dot{u}_t = \delta(1-u_t) - u_t m(\theta_t) \tag{12}$$

$$\dot{\gamma}_t = u_t m(\theta_t) - (\delta + \lambda_t m(\theta_t))\gamma_t \tag{13}$$

$$\dot{\theta}_t = \frac{m(\theta_t)u_t}{c\left(1-\frac{\theta_t m'(\theta_t)}{m(\theta_t)}\right)(u_t+\lambda_t\gamma_t)} \times \left[\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right)\left(-\dot{u}_t\frac{\gamma_t}{u_t} + \dot{\gamma}_t\right) - (p\underline{y} - \underline{w}_t) + \left(\frac{c}{q(\theta_t)}\frac{u_t+\lambda_t\gamma_t}{u_t} - \frac{\lambda_t\gamma_t}{u_t}\overline{J}_t\right)(r+\delta+\lambda_t m(\theta_t))\right] \tag{14}$$

where we suppress the dependence of the time-varying variables on $\mathbf{\Omega}_t$ to reduce notation. Note that our definition of equilibrium (Definition 1) applies as is, only that we have reduced the system of equilibrium equations to (12)–(14) (and we replaced the equation for $\dot{\underline{J}}_t$ by $\dot{\theta}_t$, see Appendix A.7). In addition to (12)–(14), an equilibrium must satisfy individual rationality of searchers (requirement 2. of Definition 1, i.e., there is no profitable one-shot deviation from the economy-wide search strategy) and the transversality condition (requirement 6. of Definition 1).

4.1 Local and Global Analysis

We start with the *local* analysis of the solution to (12)–(14) around each steady state. We then proceed to the *global* analysis, focusing on the solution of (12)–(14) away from each steady state.

Local Analysis. We start the local analysis by considering the linearized system of (12)–(14) around a given steady state,

$$\begin{bmatrix} \dot{u}_t \\ \dot{\gamma}_t \\ \dot{\theta}_t \end{bmatrix} = \begin{bmatrix} \frac{\partial \dot{u}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\ \frac{\partial \dot{\gamma}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\ \frac{\partial \dot{\theta}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \end{bmatrix} \begin{bmatrix} u_t - u(\mathbf{\Omega}) \\ \gamma_t - \gamma(\mathbf{\Omega}) \\ \theta_t - \theta(\mathbf{\Omega}) \end{bmatrix} \tag{15}$$

where all time-dependent variables are functions of search effort $\mathbf{\Omega}_t = \{\mathbf{0}, \mathbf{1}\}$ but we again omit this argument to economize on notation. The partial derivatives are evaluated at the steady state under consideration (indicated by vector $\mathbf{x}(\mathbf{\Omega}) \equiv (u(\mathbf{\Omega}), \gamma(\mathbf{\Omega}), \theta(\mathbf{\Omega}))$ where $\mathbf{\Omega} = \{\mathbf{0}, \mathbf{1}\}$ carries no time subscript to indicate 'steady state'), with $\dot{u}_t = \dot{\gamma}_t = \dot{\theta}_t = 0$. The eigenvalues of the Jacobian in (15) determine the stability of system (12)–(14) around the steady state.[15] Since an analytical solution for the eigenvalues of this three-dimensional linearized system is infeasible, we approach the problem numerically. We find that for typical parameter ranges, both the boom and the recession steady state are characterized by one

[15]This follows from the Hartman-Grobman Theorem (e.g. Hartman (1960)) and the topological equivalence of two linear systems with the same (non-zero) eigenvalue structure.

positive real and two complex eigenvalues, with the real part of the complex eigenvalues being negative. As a result, each of the steady states is a *stable saddle-focus.*

Following the Local Stable Manifold Theorem (e.g. Theorem 2.1. in Kuznetsov (1998)), this implies that the dynamics around each steady state $\mathbf{x}(\mathbf{\Omega})$, $\mathbf{\Omega} = \{\mathbf{0}, \mathbf{1}\}$, are characterized by a local stable manifold of dimension two, which we denote by $W^s_{loc}(\mathbf{x}(\mathbf{\Omega}))$ and an unstable manifold of dimension one, denoted by $W^u_{loc}(\mathbf{x}(\mathbf{\Omega}))$. For any values of u_t and γ_t *in the neighborhood* of steady state $\mathbf{x}(\mathbf{\Omega})$, the choice variable θ_t will adjust (or 'jump') in order to bring the economy onto stable manifold $W^s_{loc}(\mathbf{x}(\mathbf{\Omega}))$. On that manifold, the economy will then converge to steady state $\mathbf{x}(\mathbf{\Omega})$. In turn, for any initial values outside the stable manifold, the system diverges. Thus, for any initial values near steady state $\mathbf{x}(\mathbf{\Omega})$, a dynamic equilibrium exists since (i) system (12)–(14) is satisfied; (ii) θ_t is finite so the transversality condition is satisfied; and (iii) the workers' no-deviation condition holds (by Lemma 1 and a continuity argument). Moreover, since the number of negative eigenvalues is equal to the number of predetermined state variables, u_t and γ_t, this solution is unique (Acemoglu (2008), Theorem 7.18).

Global Analysis. We now analyze the solutions of dynamical system (12)–(14) away from the steady states. The global behavior of the dynamical system depends on the *shape* of the stable manifolds associated with the steady states. Since we typically find that the two steady states under consideration have one positive real eigenvalue and a pair of complex eigenvalues with negative real part, there exists in those cases for each steady state a two-dimensional stable manifold, $W^s(\mathbf{x}(\mathbf{\Omega}))$, and a one-dimensional unstable manifold, $W^u(\mathbf{x}(\mathbf{\Omega}))$, defined by

$$W^s(x(\mathbf{\Omega})) = \{\mathbf{x}_0 : \lim_{t\to\infty} \phi_t(\mathbf{x}_0) = \mathbf{x}(\mathbf{\Omega})\}$$
$$W^u(x(\mathbf{\Omega})) = \{\mathbf{x}_0 : \lim_{t\to-\infty} \phi_t(\mathbf{x}_0) = \mathbf{x}(\mathbf{\Omega})\}$$

where ϕ_t denotes the non-linear dynamic system (12)-(14), and where initial values $\mathbf{x}_0$ are not necessarily in the neighborhood of the steady state.[16] Hence, trajectories on the stable (unstable) manifold converge to the steady state in forward (backward) time. Knowledge of these manifolds is crucial to understand the global dynamics. It is well-known however that, generally, global stable and unstable manifolds cannot be found analytically – even for systems that are less complicated and of lower dimensions than ours.

We thus continue to proceed numerically: We construct the manifolds from local knowledge, that is from information near a fixed point $\mathbf{x}(\mathbf{\Omega})$, using *backward integration* (Brunner and Strulik (2002)). This method approximates the global stable manifold by choosing a set of starting points from a circle around the steady state. This circle lies in a plane spanned by the corresponding stable eigenvectors of

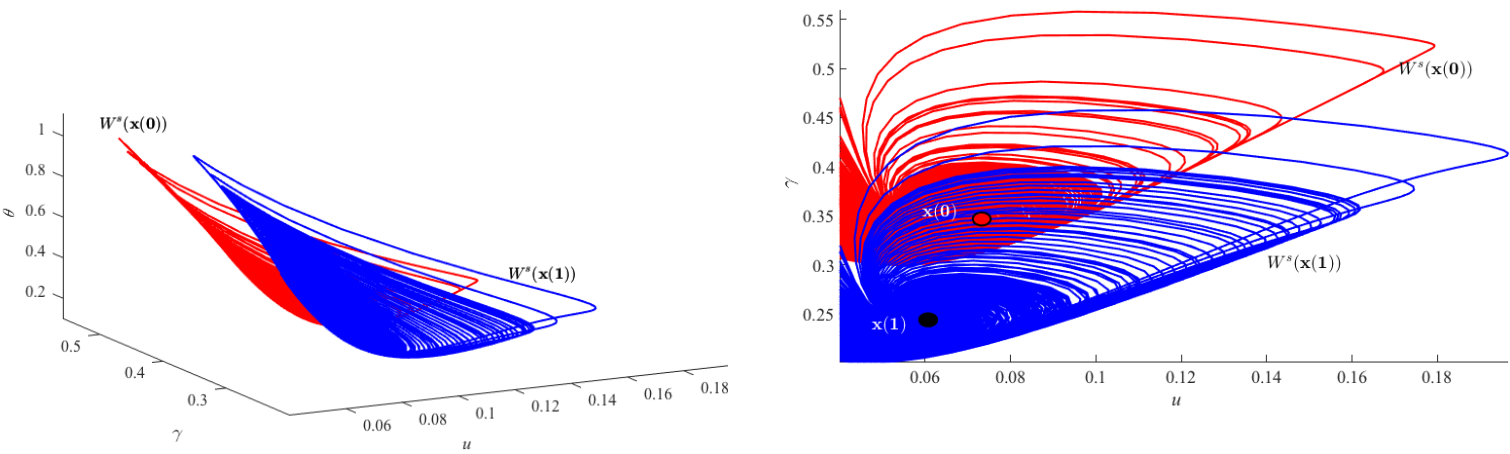
[16] One commonly uses the Local Stable Manifold Theorem to establish the existence of *local* stable manifolds, and then obtains the existence of *global* stable manifolds simply by taking unions of backward and forward iterates of local stable and unstable manifolds.

linear system (15).[17] We then evolve the dynamical system (12)–(14) backward in time, that is

$$W^s(\mathbf{x}(\mathbf{\Omega})) \approx \{\phi_t(\mathbf{x}(\mathbf{\Omega}) + \epsilon(\cos(\rho)\mathbf{v}_1 + \sin(\rho)\mathbf{v}_2)) \quad \forall t < 0\} \quad \text{where} \;\; 0 < \rho < 2\pi \;\; \text{and} \;\; \epsilon \text{ small},$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the eigenvectors corresponding to the negative eigenvalues and where the specific functional form is chosen to generate a circular structure of starting values around the steady state.

We display the resulting shape of the two stable manifolds, corresponding to the steady state with active OJS (boom; blue) and the one with passive OJS (recession; red), in Figure 5. Figure 5a shows the manifolds in the three-dimensional space and Figure 5b shows them in the two-dimensional space of state variables (u, γ).[18] Since both steady states are stable saddle-foci here, on either of the two stable manifolds the economy converges in an oscillating way to the corresponding steady state.



(a) In $u - \gamma - \theta$ space

(b) In $u - \gamma$ space

Figure 5: Stable Manifolds of Boom (Blue) and Recession (Red) Equilibrium.

But in order to understand whether these stable manifolds are indeed perfect foresight equilibria, we need to check that out-of steady state there is no profitable one-shot deviation by workers. We thus need to establish an analogue of Lemma 1 away from steady state. It turns out that in our environment with sequential auction bargaining, the same condition as in steady state guarantees that there is no profitable deviation *outside* of steady state, only that this condition needs to hold along the entire path:

Lemma 2. *There exists no profitable one-shot deviation from a dynamic equilibrium when* $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{0}\}$,

[17] This guarantees that we start constructing the stable manifold from points that lie on it. It is otherwise hard to find numerically due to its zero measure.

[18] In the figures there appear to be intersections of the paths from a given manifold, but these paths are in fact at different depths of the page.

if and only if for all $t \in [0, \infty)$

$$\theta_t(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b)}\right).$$

Likewise, there exists no profitable one-shot deviation from a dynamic equilibrium when $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{1}\}$, *if and only if for all* $t \in [0, \infty)$

$$\theta_t(\mathbf{1}) \geq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b)}\right).$$

This was the last step in establishing a dynamic Perfect Foresight Equilibrium: Starting away from steady state at some $\mathbf{x}_0 \in W^s(\mathbf{x}(\mathbf{\Omega}))$ and for a given sequence of beliefs, $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{0}\}$ or $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{1}\}$ (where $\{\mathbf{0}\}$ or $\{\mathbf{1}\}$ indicate a *sequence* of constant beliefs about other workers' OJS behavior, which is either passive or active), there exists a path such that the economy reaches steady state $\mathbf{x}(\mathbf{\Omega})$ if and only if the condition of Lemma 2 holds. This path satisfies (i) system (12)–(14), (ii) the transversality condition and (iii) the no-deviation condition by workers, and thus constitutes an equilibrium.[19] In turn, for any $\mathbf{x}_0 \notin W^s(\mathbf{x}(\mathbf{\Omega}))$ market tightness θ_t diverges to plus or minus infinity, violating transversality and thus such a trajectory does not constitute an equilibrium. We now turn to multiplicity.

4.2 Multiplicity

The previous section discusses the existence of *a* dynamic equilibrium for a given path of beliefs. We now investigate when there is *multiplicity* of dynamic equilibria. Figure 5 shows that for a given path of constant beliefs, either $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{0}\}$ or $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{1}\}$, and a range of starting values outside of steady state there is a path on the stable manifold to the steady state that corresponds to $\{\mathbf{\Omega}_t\}_{t\geq 0}$. Importantly, it also shows that the two stable manifolds have considerable overlap in the space of state variables u and γ. As can be seen in Figure 5b, there exists an upper bound $\overline{u}$, and both lower and upper bounds, $\underline{\gamma}$ and $\overline{\gamma}$, such that if $u_0 \in (0, \overline{u})$ and $\gamma_0 \in (\underline{\gamma}, \overline{\gamma})$, *two distinct dynamic paths* lead to two different steady states: one dynamic equilibrium path along the stable manifold converges to the boom steady state $\mathbf{x}(\mathbf{1})$ and another dynamic equilibrium path along the other stable manifold converges to the recession steady state $\mathbf{x}(\mathbf{0})$. Which path is selected depends on the workers' beliefs about aggregate OJS behavior.

Along those distinct paths, (i) the dynamic system (12)–(14) holds; (ii) the transversality condition is satisfied; and (iii) no worker has a profitable one-shot deviation in search effort, provided that market

[19] It is impossible to analytically pin down Lemma 2 in terms of primitives. However, we can say the following: (i) If t is large enough, then $|\theta_t(\mathbf{0}) - \theta_t^*(\mathbf{0})| < \epsilon$, ϵ small, and the passive OJS equilibrium exists under the conditions from Proposition 1. (ii) If t large enough, then $|\theta_t(\mathbf{1}) - \theta_t^*((\mathbf{1}))| < \epsilon$, ϵ small, and the active OJS equilibrium exists under the conditions from Proposition 1, where * indicates steady state.

tightness along the paths of passive (active) OJS is bounded from above (below), that is for all $t \in [0, \infty)$

$$\theta_t(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b)}\right) \leq \theta_t(\mathbf{1})$$

which follows from Lemma 2.

It follows that if the model has multiple steady states, we can find parameters and initial conditions, $u_0 \in (0, \overline{u})$ and $\gamma_0 \in (\underline{\gamma}, \overline{\gamma})$, for which our model admits *multiple* dynamic Perfect Foresight Equilibria. They are given by the two stable manifolds $W^s(x(\mathbf{0}))$ and $W^s(x(\mathbf{1}))$. These equilibria differ with respect to agents' beliefs about OJS behavior. Hence, the equilibrium dynamics of our model economy are determined not only by fundamentals (i.e., technology and preferences) but, crucially also by agents' expectations. Note that so far we have not addressed the issue of equilibrium selection. That is, we took the path of beliefs/search strategies, $\{\mathbf{\Omega}_t\}_{t\geq 0}$, as given. We will return to this issue and how agents' beliefs about the profitability of OJS change in our quantitative illustration below.

5 Quantitative Illustration

We now undertake a stylized quantitative exercise. First, we calibrate the model to the US economy during the Great Recession. We compare boom and recession steady states in terms of worker flows, worker composition and wage inequality. We also perform a simple exercise on the jobless recovery. Second, we study the dynamic equilibrium path of the economy in response to an unanticipated expectations shock that makes agents more optimistic and construct the Beveridge Curve during the recovery. Finally, to give further support for our mechanism, we provide direct evidence for its main underlying forces.

5.1 Calibration

We calibrate our model to quarterly US data from the Great Recession. The main data source for worker flows and unemployment rates is the Current Population Survey (CPS), where we aggregate the monthly series up to quarterly frequency. For vacancies, we use the JOLTS data from the Bureau of Labor Statistics. We provide details on the data, variables and quarterly aggregation in Appendix B.

We first need to parameterize the matching function and choose the telegraph matching function:

$$m(\theta) = \phi\frac{\alpha\theta}{\alpha\theta + 1}, \tag{16}$$

where ϕ is the overall matching efficiency and α is a parameter that determines the curvature of the matching technology. We use this matching function for three reasons. First, as a special case of a CES function, it has many desirable features of a matching function. Second, with the level parameter ϕ and the shape parameter α, we can closely approximate the matching functions used in the literature (e.g. that in Shimer (2005)). Finally, it allows us to explicitly solve the model for vacancies and tightness.

We set the parameters $(r, b, \delta, p, \underline{y})$ outside the model and report the values in Table 1. Our model features a constant separation rate δ across boom and recession, which we set equal to average observed quarterly separations over time. Moreover, since our model can generate multiple equilibria in the absence of any productivity changes, we normalize aggregate productivity p to one. We set b to about 70% of average labor productivity – an intermediate value considering the calibrations in the literature.

Table 1: Exogenously Set Parameters

	Value	**Parameter Description**	**Notes**
r	0.0113	discount rate	standard
$\underline{y}$	1	match-specific productivity first job	normalization
b	0.91	opportunity cost of employment	69% of average labor productivity
δ	0.052	job separation rate	average quarterly separation rate across peak and trough
p	1	aggregate productivity	aggregate productivity shifter; here normalized

We calibrate the remaining parameters using our model. They relate to active and passive OJS intensity (λ_0, λ_1), the parameters of the matching function (α, ϕ), the vacancy cost c and the cost of OJS k, as well as match productivity in highly productive jobs $\overline{y}$. We target business cycle moments from the Great Recession, i.e., moments from the previous boom (corresponding to the steady state equilibrium with active OJS, $\mathbf{\Omega} = \mathbf{1}$) and the trough of the recession ($\mathbf{\Omega} = \mathbf{0}$). We date the peak prior to the Great Recession at 2007 Q4, and the trough at 2009 Q3 (where the EE rate was at its lowest).

We choose moments as targets that strongly vary with the parameters we seek to estimate. Central to our calibration strategy is to target EE transition rates in both boom and recession, since we would like to explain business cycle fluctuations through differences in OJS. To align data and model, the targeted EE transition rates in the data are those that are associated with a wage *increase*.[20] The observed EE flows across recession and boom identify the search intensities, λ_0 and λ_1. We target unemployment rates in boom and recession as they are closely related to the matching probabilities of workers, thereby pinning down the parameters of the matching function (α, ϕ). We also target the vacancy rate in boom and recession. They determine both the vacancy cost c and productivity $\overline{y}$ through the free entry condition. Finally, we target wage dispersion in the boom as it identifies search cost k.

We use General Method of Moments to calibrate our model.[21] Targeted moments (data and model) and estimated parameter values are in Tables 2 and 3.

[20]The raw quarterly flows at peak and trough of the Great Recession are higher than those in Table 2, namely $EE(\mathbf{0}) = 0.0424$ and $EE(\mathbf{1}) = 0.0573$. Consistent with the model, we want to capture only those flows that coincide with an increase in match productivity. See Appendix B.5 for how we computed the flows associated with moves *up* the job ladder.

[21]We do not match the moments exactly – despite an equal number of moments and parameters – because we impose parameter restrictions that ensure a solution with multiple equilibria is feasible.

Table 2: Targeted Moments

	Data	Model
$EE(\mathbf{1})$	0.0351	0.0382
$EE(\mathbf{0})$	0.0223	0.0223
$u(\mathbf{1})$	0.0491	0.0601
$u(\mathbf{0})$	0.0949	0.0976
$v(\mathbf{1})$	0.0300	0.0455
$v(\mathbf{0})$	0.0176	0.0302
$\frac{\overline{w}(\mathbf{1})}{\underline{w}(\mathbf{1})}$	1.4176	1.3600

See Appendix B for construction of variables.

Table 3: Calibrated Parameters

	Estimate	Parameter Description
λ_0	0.0981	passive OJS intensity
λ_1	0.1307	active OJS intensity
α	1.1537	curvature matching function
ϕ	2.4697	overall matching efficiency
c	10.2208	vacancy posting cost
$\overline{y}$	1.6855	match-specific productivity second job
k	0.0897	search cost

We match the EE flows in boom and recession, which is the main aspect of our mechanism. Observed unemployment rates and wage dispersion are matched reasonably well while the model over-predicts the level of vacancies (but it almost exactly matches the difference between boom and recession). Our model of multiple equilibria can thus be made consistent with the observed differences in unemployment and vacancies over the cycle without necessitating large (or any) aggregate productivity shocks – something that is difficult to obtain in comparable random search models with a unique equilibrium.

The calibrated parameters suggest that on-the-job searchers are more than twice as actively searching in boom ($\lambda_0+\lambda_1 = 0.23$) compared to recession ($\lambda_0 = .098$).[22] The curvature of the matching technology α is estimated to be nearly linear, and matching efficiency ϕ is 2.4. Notice that the matching efficiency is estimated to be higher than what is suggested by the literature. This stems from using a different tightness measure v/s (which is smaller than the conventional v/u since it takes into account all searchers s, not just the unemployed u). The costs of OJS are estimated to be a relatively small fraction of the first job's output (about 9%). The match productivity difference between the two job types is large ($\overline{y} = 1.69$). Finally, the estimated cost of posting a vacancy c, which reflects the overall resources that a firm spends on hiring, are comparably high.[23]

Our calibrated economy admits multiple steady state equilibria, where labor market tightness across steady states satisfies $\theta(\mathbf{0}) < m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right) < \theta(\mathbf{1})$ (Lemma 1) or, equivalently (by Proposition 1), where aggregate productivity $p = 1 \in [p_l, p_h]$ and where p_l and p_h are computed based on our calibration.

5.2 The Economy Across Steady States

We now use the calibrated economy to analyze the volatility of equilibrium outcomes across steady states, to shed light on the cyclicality of frictional wage inequality, and to analyze the jobless recovery.

[22] The search intensity of (active) on-the-job searchers, $\lambda_0 + \lambda_1 = 0.23$, is considerably lower than the search intensity of *unemployed* workers which was normalized to 1, in line with evidence by Faberman, Mueller, Sahin, and Topa (2017).

[23] This estimate is in line with a growing literature that argues hiring costs are substantial and, depending on the worker type, can take up more than an annual wage. For evidence, see for instance Blatter, Muehlemann, and Schenker (2012) and Dube, Freeman, and Reich (2010) and the references therein.

Comparison of Labor Market Outcomes Across Boom and Recession. We start with a comparison of the two steady states, focussing on labor market moments that we did not target in our calibration. Table 4 shows these moments in data and model. Comparing labor market tightness θ (note that this statistic is not only based on v, u but also on $\lambda\gamma$ which is not targeted) and matching probability $m(\theta)$ across boom and recession shows that the model generates sizable fluctuations: the model accounts for about 84% of the observed decrease in labor market tightness and almost exactly matches the 38% drop in workers' matching probability during the Great Recession.

Table 4: Non-Targeted Moments

	Data	**Model**
$\theta(\mathbf{1})$	0.3209	0.4252
$\theta(\mathbf{0})$	0.1252	0.2096
$m(\theta(\mathbf{1}))$	0.7881	0.8128
$m(\theta(\mathbf{0}))$	0.4903	0.4810
$\frac{\lambda(\mathbf{1})\gamma(\mathbf{1})}{s(\mathbf{1})}$	0.4755	0.4387
$\frac{\lambda(\mathbf{0})\gamma(\mathbf{0})}{s(\mathbf{0})}$	0.3240	0.3224
$\frac{\gamma(\mathbf{1})}{1-u(\mathbf{1})}$	0.1533	0.2185
$\frac{\gamma(\mathbf{0})}{1-u(\mathbf{0})}$	0.2497	0.5242

See Appendix B for how m, θ, γ, $\lambda\gamma$ and s are measured in the data.

Our model also captures subtle changes in the composition of searchers as well as of workers across different rungs in the job ladder: in the data, the proportion of on-the-job searchers in overall searchers, $\lambda\gamma/s$, declined by 32% during the last recession, indicating that the quality of the pool of searchers deteriorated. In the model it declined by 26.5%, capturing 83% of the observed drop. This *shift in the composition of searchers* over the cycle is at the heart of the mechanism underlying multiplicity and – as we show below – is also crucial for generating the phenomenon of jobless recovery. Our model also predicts that there is a significant change in the distribution of workers across the job ladder over the cycle: going into the recession, there is a large increase in the proportion of employed workers on the lowest rung, $\frac{\gamma}{1-u}$. This shift towards workers at the bottom of the job ladder also occurs in the data, though it is quantitatively smaller. The *composition shift of employed workers* across different parts of the job ladder will be crucial for generating pro-cyclical wage dispersion below.

Note that these cyclical changes in labor market variables are obtained through multiple equilibria alone and without alluding to any decline in aggregate productivity p, which is held fixed in this exercise. This suggests that differences in the intensity with which workers search on-the-job in boom versus recession can have a large impact on the labor market.

Finally, because the theory is ambiguous regarding welfare, we use the calibrated economy to ascertain whether the two steady state equilibria can be Pareto-ranked. We find that the aggregate output net of search costs, $Y(\mathbf{\Omega})$, is 8% larger in the boom than in the recession: $Y(\mathbf{1}) = 0.96$ versus $Y(\mathbf{0}) = 0.89$.

Frictional Wage Dispersion. Hornstein, Krusell, and Violante (2011) argue that frictional wage dispersion in standard search models is limited. They use the 'mean-min ratio' (Mm) to quantify frictional wage dispersion and find that in a model *without* OJS, this ratio equals 1.05, that is, the average accepted wage is only 5% higher than the lowest wage a worker will accept. They also point out that the wage dispersion in a model *with* OJS is considerably larger, namely around 1.25, which is close to what we find below when using the Mm-ratio suggested by our model.[24] However, we do not want to focus on the level of frictional wage dispersion. Instead, we want to assess the implications of our model for the *cyclicality of frictional wage dispersion*.

In our job ladder model the mean-min ratio is given by:

$$Mm \equiv \frac{\frac{\gamma}{1-u}\underline{w} + \left(1 - \frac{\gamma}{1-u}\right)\overline{w}}{\underline{w}} = \frac{\gamma}{1-u} + \left(1 - \frac{\gamma}{1-u}\right)\frac{\overline{w}}{\underline{w}}$$

where we have suppressed the dependence of all variables on $\mathbf{\Omega}$. Table 5 reports this statistic in the data and the model. We find that this measure of frictional wage dispersion is highly pro-cyclical: in the data it is around 5% higher in the boom than in the recession (compare columns (1) and (2)). In our model, this measure is even more pro-cyclical with an increase of 17%. While the model overestimates the increase in frictional wage dispersion compared to the data, note that both in data and model 60-70% of the increase in wage dispersion during the boom is due to a *composition shift of employed workers* across the job ladder, captured by a change in $\gamma/(1-u)$. In contrast, pure wage dispersion itself, $\overline{w}/\underline{w}$, is relatively stable over the cycle. To see this, in column (3) we keep the distribution of workers across the job ladder at its recession level $\frac{\gamma}{1-u}(\mathbf{0})$. In this case, frictional wage dispersion increases by less than half as much as when the composition of workers changes (1.6% instead of 5% in the data; 7% instead of 17% in the model).

In sum, most of the decline in frictional wage dispersion during the Great Recession is due to the contraction of the job ladder, which leads to a large increase of workers in lower paying jobs. Our model provides an explanation for why the job ladder is contracting, namely the drop in workers' search effort and firms' response to it, and thus gives a rationale for pro-cyclical frictional wage dispersion.

Jobless Recovery. We now illustrate how the cyclical change in the composition of searchers provides an explanation for the phenomenon of *jobless recovery.* We illustrate this point through a simple exercise that highlights the impact of the recovery on unemployment.

[24]Moreover, they show that models with sequential auction wage setting can generate even larger Mm ratios.

Table 5: Frictional Wage Dispersion in Boom and Recession

	Recession $\boldsymbol{\Omega} = \mathbf{0}$	Boom $\boldsymbol{\Omega} = \mathbf{1}$	Boom $\boldsymbol{\Omega} = \mathbf{1}$ with $\frac{\gamma}{1-u}(\mathbf{0})$
	(1)	(2)	(3)
Mm Data	1.29	1.35 (+5%)	1.31 (+1.6%)
Mm Model	1.09	1.28 (+17%)	1.17 (+7%)

In parentheses: percentage difference between recession $\boldsymbol{\Omega} = \mathbf{0}$ and boom $\boldsymbol{\Omega} = \mathbf{1}$.
Column (3): we leave the composition of employed workers, $\gamma/(1-u)$, unchanged at recession level.

Consider an economy in the recession steady state, where all employed workers in low productivity jobs exert low search effort. We investigate the impact of an unexpected change in workers' beliefs, such that all those workers start searching actively for another job. On impact, the stocks u and γ do not adjust, but there is an immediate response in the search activity λ. Thus, there is a sudden increase in the measure of active searchers from $s(\mathbf{0}) = u(\mathbf{0}) + \lambda_0\gamma(\mathbf{0})$ to $s^R = u(\mathbf{0}) + (\lambda_0 + \lambda_1)\gamma(\mathbf{0})$, where the superscript R stands for 'Recovery'. This leads to a *crowding out* of unemployed workers: Conditional on forming a match, the probability that it is with an unemployed worker is now lower since more are searching on-the-job. The fraction of hires out of unemployment (denoted by κ) *decreases*:

$$\kappa(\mathbf{0}) \equiv \frac{u(\mathbf{0})}{u(\mathbf{0}) + \lambda_0\gamma(\mathbf{0})} > \frac{u(\mathbf{0})}{u(\mathbf{0}) + (\lambda_0 + \lambda_1)\gamma(\mathbf{0})} \equiv \kappa^R.$$

Based on our calibrated model, Table 6 summarizes the changes in these conditional matching rates and also in the unconditional ones between trough and recovery. As the recovery starts and conditional on match formation, the probability that the match is formed with an unemployed worker declines from $\kappa(\mathbf{0}) = 0.68$ to $\kappa^R = 0.47$ ($\Delta\kappa$ of -30%, see column (1) Table 6). Thus, the conditional likelihood that an unemployed worker is selected over an employed worker significantly drops. This is what we refer to as *crowding out* during the recovery. It stems from the *composition externality* that employed searchers impose on unemployed ones.

Certainly, what matters for job seekers is not just the conditional likelihood of being drawn. It is also important how fast the overall matching is. Under the belief that employed workers actively search for another job, the matching rate for firms goes up during recovery. In response, new vacancies are created (firms instantaneously adjust by posting vacancies so that profits are driven to zero again) and

Table 6: Jobless Recovery (Model)

		λ **unchanged**
	(1)	(2)
$\Delta\kappa$	–0.30	0
$\Delta m(\theta)$	0.04	0.37
$\Delta\kappa m(\theta)$	–0.27	0.37
$\Delta(1-\kappa)m(\theta)$	0.70	0.37

Δ: %-Change between recession and recovery.

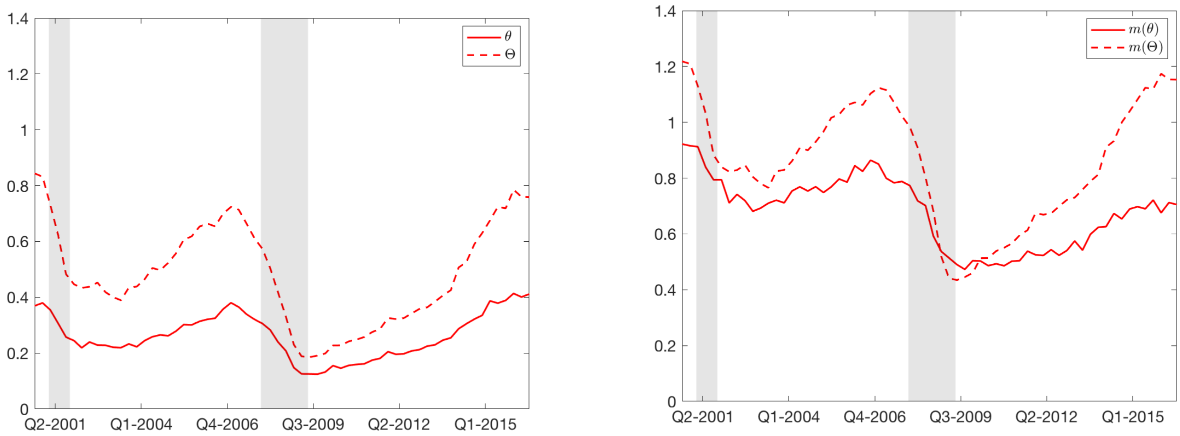
market tightness θ adjusts, as does the matching rate of workers, $m(\theta)$. Based on our calibration, for *unemployed* workers the negative composition effect dominates the positive effect of the overall matching rate: the matching rate for an unemployed worker drops from $\kappa(\mathbf{0})m(\theta(\mathbf{0})) = 0.33$ in the recession to $\kappa^R m(\theta^R) = 0.24$ in the recovery (where θ^R is the market tightness during the recovery with unadjusted stocks but adjusted vacancies and search intensity) – a drop of 27%. The implication is that the *unemployment rate initially increases during recovery*, $\dot{u} > 0$, since the separation rate δ is unchanged. In turn, the matching rate of employed workers increases by 70%, going from $(1-\kappa(\mathbf{0}))m(\theta(\mathbf{0})) = 0.16$ in the recession to $(1-\kappa^R)m(\theta^R) = 0.26$ in the recovery.

Our exercise suggests that the main force behind these shifts are changes in the *conditional* meeting rates κ and $1-\kappa$, and thus in the composition externality that kicks in during recovery. To highlight its quantitative importance, column (2) reports the changes in matching rates if only vacancies had adjusted during the recovery but with the composition of the pool of searchers unchanged (i.e. no change in search intensity λ). In this scenario, we would have observed *the same* increase in matching rates for *all* workers during the recovery. This is what a conventional random search model without multiplicity and composition externality would predict but it is clearly not supported by the data: Figure 11a (Appendix C.1) shows that the conditional matching probability of an unemployed worker κ is decreasing during recovery. This translates into a strong recovery of matching rates for the employed but at best stagnant matching rates for the unemployed (Figure 11b, Appendix C.1).[25]

Column (2) also suggests that neglecting the composition externality during recovery *overestimates* the increase of matching rate, m, after the crisis. This hints at the importance of taking the *effective* market tightness $\theta = \frac{v}{s}$ into account in order to understand the jobless recovery. Since we observe vacancies and unemployment, we can readily construct the *conventional* market tightness $\Theta = \frac{v}{u}$. We want to compare Θ to our effective market tightness θ, which we obtain from the data as $\theta = \frac{v}{u+\lambda\gamma} = \frac{v}{u+EE/m(\theta)}$. Figure 6a plots both θ and Θ. There is not only less fluctuation in θ than in Θ but, in particular after the crisis, the recovery of θ is much slower than that indicated by Θ. This is because,

[25]Note that we cannot replicate the exact exercise from Table 6 in the data, since we do not have a model-independent measure of search intensity λ.

contrary to the conventional tightness Θ, the effective market tightness θ reflects the increase in the measure of on-the-job searchers. The implications for fluctuations in matching rates follow immediately (Figure 6b): While the matching rate based on the conventional tightness measure $m(\Theta)$ shows a fast recovery, the matching rate $m(\theta)$ recovers much more slowly, fueling the jobless recovery.[26]



(a) Market Tightness $\Theta = \frac{v}{u}$ and Effective Market Tightness $\theta = v/\left(u + \frac{EE}{UE/u}\right)$; all based on data.

(b) Matching Rates: $m(\theta) = UE/u$ (data) and $m(\Theta)$ (where we use Θ from the data and the calibrated parameters of m).

Figure 6: Market Tightness and Matching Rates (Data).

All this indicates that at impact, the recovery out of the recession looks even bleaker for the unemployed than the recession itself. Due to crowding out and stagnant matching probabilities, rather than a decline, we see an increase in the unemployment rate immediately *after the recession* ends.

5.3 Dynamics in Response to an Expectations Shock

So far we have focussed on the comparison of multiple *steady states* (or of the recession steady state and the recovery at impact). We now investigate the transition dynamics, starting at the trough of the Great Recession and following the economy on its path to the boom steady state.

The transition dynamics are driven by the path of equilibrium beliefs $\{\mathbf{\Omega}_t\}_{t\geq 0}$. With slight abuse of notation, we will use $\mathbf{\Omega}_t$ to indicate both the workers' *belief* about the economy's OJS behavior as well as workers' *search strategy* itself. In order to maintain a tight link between our theoretical results under steady state and this section, we assume that agents face an *unanticipated* expectations shock that makes them change their beliefs about the aggregate OJS behavior $\mathbf{\Omega}_t$. Thereby this shock induces

[26]This model of unemployment cycles helps understand the labor dynamics of the last recession, in particular the jobless recovery. Prior to the 1990s, however, recoveries were not jobless. We propose two ways to rationalize the absence of jobless recoveries with our model. Either the economy was not in the multiplicity region, which dampens the dynamics. Or the collapse of the job ladder was milder. Thus, less employed searchers were stuck at the bottom of the job ladder and the crowding out of unemployed by employed searchers during recovery was less severe. Jobless recovery was milder/shorter. Of course, this is speculative and just highlights how our model *could* generate the absence of jobless recoveries.

a shift in their individual search behavior ω_t. This approach requires no change to our current model, definition of equilibrium, or previous theoretical results.

We are aware that there may be more plausible ways to pin down the path of beliefs $\{\mathbf{\Omega}_t\}_{t\geq 0}$ than through unanticipated expectation shocks. We lay out the model with *anticipated* expectation shocks in Online Appendix I.4, and show that the quantitative implications of the two models are similar.[27]

Here we consider an economy in the recession steady state, with low search intensity, $\mathbf{\Omega}_t = \mathbf{0}$ where all variables are are stationary, $\dot{u}(\mathbf{0}) = \dot{\gamma}(\mathbf{0}) = \dot{\theta}(\mathbf{0}) = 0$. We treat this as the initial equilibrium. We then introduce an unanticipated and permanent expectations shock, where each agent believes that all other agents become optimistic and start active OJS, i.e., $\mathbf{\Omega}_t = \mathbf{1}$ forever. This shock to workers' beliefs makes firms immediately adjust their vacancy posting upward, which brings the economy from the recession steady state $\mathbf{x}(\mathbf{0})$, through a vertical jump onto the boom stable manifold $W^s(\mathbf{x}(\mathbf{1}))$ as illustrated in Figure 7a, where as above, the vector $\mathbf{x}$ denotes $\mathbf{x}(\mathbf{\Omega}) \equiv (u(\mathbf{\Omega}), \gamma(\mathbf{\Omega}), \theta(\mathbf{\Omega}))$.[28] Once on the boom manifold, the economy will transit along an oscillating path to the boom steady state $\mathbf{x}(\mathbf{1})$ in the direction of the black arrows.[29] We thus focus on the transition dynamics of the *recovery*, i.e., from the recession steady state all the way to the boom steady state. We now investigate the Beveridge Curve along the dynamic path.

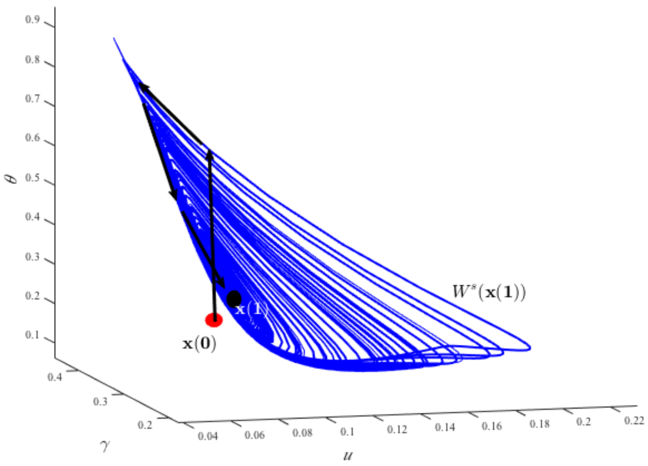
Shift of Beveridge Curve. Based on Proposition 3, our model predicts that the Beveridge Curve associated with the active OJS steady state (boom) is shifted outward compared to the Beveridge Curve associated with the passive OJS steady state (recession). Those Beveridge Curves are a hypothetical construct of which we only see one data point at a time (the equilibrium), much like a demand curve, whereas in the data we observe the transition dynamics. We therefore ask whether the transition path from the recession to the boom steady state can match the evolution of the empirically observed Beveridge Curve. Figure 7b, which plots the model transition path of θ and u from the recession to the boom steady state, indeed replicates such a shift.[30] Coming out of the recession with high unemployment and low tightness, the Beveridge Curve gets shifted outward immediately as recovery begins. It then follows a path of decreasing unemployment and increasing tightness and crawls back in towards a new

[27] In the model with *anticipated* expectation shocks (similar to Kaplan and Menzio (2016)), the agents understand that with a certain probability there is a shock to their expectations inducing them to change their search behavior. Introducing anticipated expectations shocks requires a substantial change to the model and some necessary approximation in the quantitative part, which is why we opted in the main text for the cleaner model with unanticipated shocks. Another way of pinning down $\{\mathbf{\Omega}_t\}_{t\geq 0}$ would be through a process of shocks to *fundamentals* (e.g. to p) that drive the economy out of the parameter region that admits multiplicity and hence force the economy in each t into either $\mathbf{\Omega}_t = \mathbf{0}$ or $\mathbf{\Omega}_t = \mathbf{1}$.

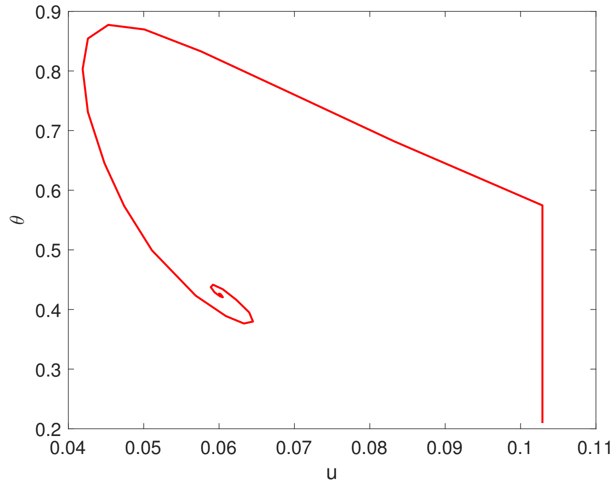
[28] This jump is feasible for the calibrated economy, as the boom stable manifold 'covers' the recession steady state in the (u, γ)-space. The manifold in Figure 7a is computed through backward integration, based on the calibrated parameters.

[29] The dynamic path is converging with oscillations, a feature consistent with the properties of the eigenvalues of the boom steady state, which indicate that it is a saddle focus.

[30] We choose to plot θ instead of v on the y-axis since θ is the jump variable in our dynamic equilibrium. Results in the u-v space look similar.



(a) Path from Recession Steady State $\mathbf{x}(\mathbf{0})$ onto Boom Stable Manifold $W^s(\mathbf{x}(\mathbf{1}))$ with Subsequent Transition to Boom Steady State $\mathbf{x}(\mathbf{1})$ (Based on Calibrated Model).



(b) Beveridge Curve in (u, θ)-Space (Model)

Figure 7: Dynamics Equilibrium: Transition Path From Recession to Boom Steady State.

steady state with active OJS. Thus, the multiplicity (inducing a jump from the recession onto the boom manifold after a belief switch) combined with the oscillating dynamics of the boom equilibrium, captures the observed shift of the Beveridge Curve fairly well. The transition takes 35 quarters or 8.75 years.

The mechanism underlying the shift of the Beveridge Curve stems from a change in the *composition of searchers*, which is key for firms who seek to hire (and closely related to the *jobless recovery* discussed above). Figure 14a in Appendix C.2 shows that the composition of searchers drastically changes during the recovery where on-the-job searchers temporarily make up the largest group in the pool of searchers, measured by $\lambda\gamma/s$. This is the main driving force in firms' sudden increase of job creation.[31] Consistent with the shift in the composition of searchers, there is also a change in the *composition of employed workers* across the job ladder: Figure 14b (Appendix C.2) shows an increase in the share of employed workers in high productivity jobs $\xi/(1-u)$ in the transition to the boom steady state, indicating that the job ladder resumes its activity during recovery.

5.4 Direct Evidence for the Model Mechanism

There are two key features underlying the model mechanism: 1. Pro-cyclical search intensity of employed workers. 2. Composition changes in the pool of searchers as well as in the pool of employed workers over the cycle. Here we aim to provide some direct evidence for them.

[31] Sniekers (2018) also explains the dynamics of the Beveridge Curve. He focuses on a limit cycle in a search model with demand externality, but without OJS and hence without the composition externality that drives our mechanism.

Pro-Cyclical Search Intensity Of The Employed. First, there is direct evidence for pro-cyclical search intensity of on-the-job searchers by Carillo-Tudela, Hobijn, Perkowski, and Visschers (2015) who use the Contingent Worker Supplement of the CPS.[32]

Second, a natural alternative source on search intensity is the American Time Use Survey (ATUS). As is well known, though, reported times for job search are extremely small in the ATUS. Therefore, and in line with the findings on search intensity of the *unemployed* (Mukoyama, Patterson, and Şahin (2018)), we find that the cyclical pattern of search intensity of the *employed* is noisy and not very pronounced. In Figure 13a in Appendix C.1, we report the time spent searching by employed workers, first unconditionally, and then conditional on reporting non-zero search activity. The latter corresponds to the the measure of search intensity in our model and it is slightly pro-cyclical.

Finally, we make an attempt to compute our own measure of search intensity directly from the CPS data, where we proceed in two steps. First, we compute a measure of the stock γ_t (which is not something we can read off the data directly). Second, we use γ_t to disentangle $\lambda_t\gamma_t$ (which we obtain from $EE_t/m(\theta_t) = \lambda_t\gamma_t$), thereby backing out search intensity λ_t.

For the first step, we rely on the dynamic flow equations. We first approximate the continuous time change in the stock of γ_t by the discrete time difference where $\dot{\gamma}_t = \gamma_{t+1} - \gamma_t$. We use the fact that the flows are given by $UE_t = m(\theta_t)u_t$ and $EE_t = \lambda_t\gamma_t m(\theta_t)$ and assume that the exogenous separations come with equal proportions from all employed workers (whether they search or not, i.e., $\delta_t = \frac{EU_t}{1-u_t}$).[33] We can then write the law of motion for γ_t, given by (10), in differences as:

$$\gamma_{t+1} = \gamma_t \frac{1 - u_t - EU_t}{1 - u_t} + UE_t - EE_t. \tag{17}$$

To obtain the stock γ_t from this flow equation, we need an initial condition γ_0, which is in principle not given. We therefore pick the initial condition that corresponds to the average γ_t in the time series.[34]

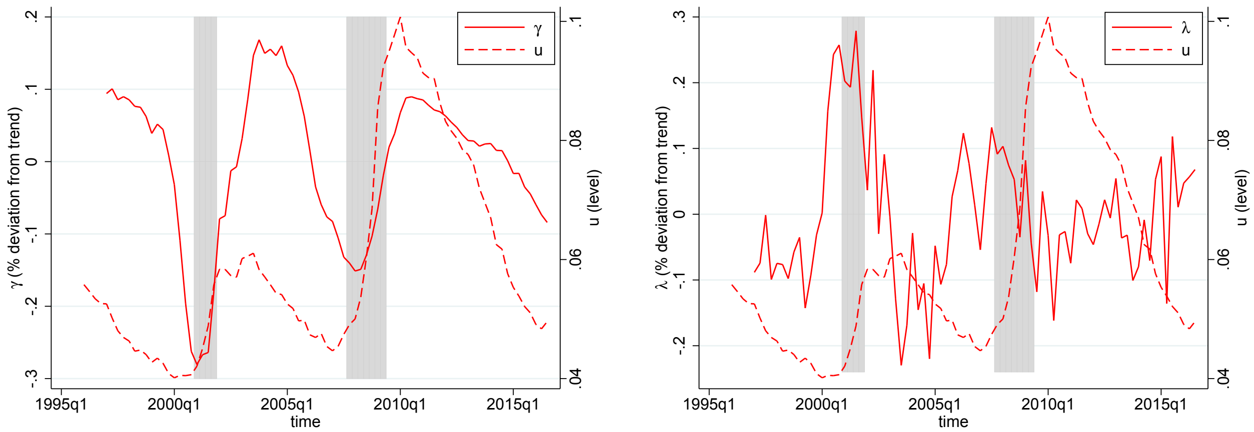
Figure 8a displays the HP de-trended time series of γ_t. In line with our calibrated model, the stock of on-the-job searchers is countercyclical (almost exactly coinciding with the unemployment rate in terms of cyclicality which we overlaid onto the figure), indicating that during downturns workers are increasingly stuck at the bottom of the job ladder. And once the boom is in full swing, γ_t starts to decrease gradually. We believe this finding that the stock of on-the-job searchers is countercyclical is new.

In the second step, once we have obtained a time-series for γ_t, we can infer the search intensity of on-

[32] This evidence on the cyclicality of employed workers' search behavior is however only suggestive because they have access to yearly data for some years only (1995, 1997, 1999, 2001, and 2005), which is too infrequent to capture the details of the business cycle. The data also does not cover the Great Recession.

[33] In the Online Appendix III, we construct γ for different assumptions on the separation rate δ and find similar results.

[34] To that end, we start from a grid of initial conditions γ_0, and compute the path γ_t for each of them. From all these paths, each corresponding to an initial condition γ_0, we pick the path where γ_0 is equal to the average of all γ_t in that path. We experimented with other ways of obtaining the initial conditions but the results were similar.



(a) On-the-Job Searchers γ (De-Trended) and Unemployment Rate u.

(b) Search Intensity λ (De-Trended) and Unemployment Rate.

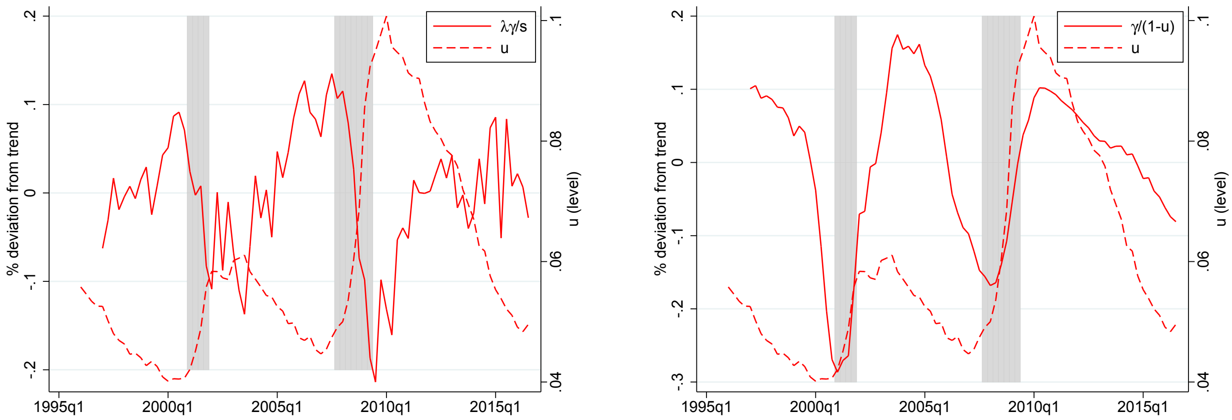
Figure 8: Active On-The-Job Searchers and Search Intensity.

the-job searchers using our framework. In fact, we know that $EE_t = \lambda_t m(\theta_t)\gamma_t$ as well as $UE_t = m(\theta_t)u_t$. Therefore, search intensity is given by $\lambda_t = \frac{EE_t u_t}{UE_t \gamma_t}$. The implied de-trended series for λ_t is *pro-cyclical*, running opposite to the unemployment rate (Figure 8b): search intensity is above trend in the boom but then falls during the recession, reaching its minimum when unemployment peaks.

Note that (as mentioned in footnote 14), if contrary to our assumption the unemployed searched with intensity $\lambda_u \neq 1$, then based on our flow equations we would obtain $\frac{\lambda_t}{\lambda_{u,t}} = \frac{EE_t}{UE_t}\frac{u_t}{\gamma_t}$. So by measuring the search intensity of the employed by $\frac{EE_t}{UE_t}\frac{u_t}{\gamma_t}$, we have effectively obtained a measure of search intensity of the employed, λ_t, *relative* to the search intensity of the unemployed, $\lambda_{u,t}$. Figure 8b indicates that this relative measure is higher during booms than recessions.

Cyclicality of the Worker Composition. There are two key composition changes in our model that are important for both the multiplicity and our results more generally. First, *the composition of searchers* shifts over the cycle. During the boom the share of employed searchers in overall searchers, $\lambda\gamma/s$, is relatively larger (while in the recession the unemployed searchers, u/s, gain importance which are just the flip slide). We showed that this composition shift leads to crowding out of the unemployed during the economy's recovery and thus gives rise to a jobless recovery. Figure 9a shows that this composition shift in the pool of searchers also exists in the data. While the fraction of unemployed searchers is counter-cyclical, the fraction of employed searchers is highly pro-cyclical, being low in the recession and starting to rise during the recovery. This is direct evidence for the crowding out of unemployed searchers during an economy's recovery.

Second, *the composition of employed workers* across the job ladder shifts over the cycle. During the



(a) Composition of Searchers (De-Trended) and Unemployment Rate.

(b) Composition of Employed Workers (De-Trended) and Unemployment Rate.

Figure 9: Composition of Searchers and Employed Workers

boom the share of employed workers in high rungs of the job ladder is relatively large (peaking at the end of the boom/beginning of recession), while in the recession the fraction of employed workers in the lowest rung grows. Figure 9b plots the de-trended shares of employed workers in the lowest rung of the job ladder (note that employed workers in the highest rung, $\xi/(1-u)$, are just the flip side of $\gamma/(1-u)$). In line with the model, the share of workers that find themselves at the bottom of the job ladder is counter-cyclical, while the share of those at the top is pro-cyclical in the data.[35]

A similar mismatch-enhancing effect of recessions was also shown in Bowlus (1995), Lazear (2014), Gertler, Huckfeldt, and Trigari (2016) and Moscarini and Postel-Vinay (2016). It has been referred to as the *sullying effect* of recessions where workers get stuck in poor matches at the bottom of the job ladder (Barlevy (2002)). This is also supported by the fact that, as our model predicts, wage growth during the boom is higher (Faberman (2015)). All this suggests that the recession negatively affects the composition of jobs, with a considerable bias towards low-productivity jobs.

Finally, these composition shifts of employed workers across the job ladder are also consistent with the finding that during recessions, the pool of unemployed workers consists of relatively more high-wage workers and therefore is improving in quality. Figure 9b shows that at the beginning of a recession, there are relatively more workers in high-paying jobs ($\xi/(1-u)$ is above trend while $\gamma/(1-u)$ is below). Hence, given a constant separation rate for all employed workers, relatively more high-wage workers enter unemployment during recessions (until the composition of employed workers flips at the end of the recession), which is why the quality of the unemployment pool improves.

[35]To compute $\gamma_t/(1-u_t)$, we use the constructed γ_t based on (17). Then, $\xi_t/(1-u_t) = (1-u_t-\gamma_t)/(1-u_t)$. Also see Figure 19, Online Appendix III, for the equivalent of Figure 9 but in terms of *percentage point* deviations from trend.

6 Conclusion

The main contribution of this paper is to develop a new theoretical mechanism that explains unemployment cycles based on endogenous search intensity of employed job seekers. We argue that the labor market behavior of the employed can have profound implications for the unemployed. In particular, even in the absence of exogenous shocks, search behavior of employed workers by itself can create multiple equilibria and hence cyclical outcomes due to a strategic complementarity in active OJS and vacancy creation. Active OJS by the employed makes it more attractive for firms to post vacancies, which in turn makes OJS more attractive. Self-fulfilling beliefs can thus give rise to either an equilibrium with high OJS activity which we interpret as a boom or an equilibrium with low OJS activity, interpreted as a recession. We show that this model qualitatively accounts for the following features in a unified way: (i) cyclicality of labor market outcomes; (ii) pro-cyclical frictional wage dispersion through a reallocation of workers from low to high productivity jobs in the boom; (iii) a jobless recovery through a novel mechanism where the employed searchers crowd out the unemployed and, as a result, (iv) an outward shift of the Beveridge Curve during the boom.

Given the stylized nature of the model, we propose a simple quantitative exercise to illustrate this mechanism in US data: First, changes in beliefs about aggregate OJS behavior are consistent with large cyclical fluctuations in vacancies, unemployment and job-to-job transitions, even without any change in aggregate productivity or other primitives. Second, we show that our calibrated model generates pro-cyclical frictional wage dispersion – in line with the data. Third, if beliefs in the passive OJS equilibrium (recession) turn optimistic and employed workers start searching actively on-the-job, then they crowd out the unemployed searchers, giving rise to a jobless recovery. Last, the model's transition dynamics from the steady state with passive OJS to the one with active OJS resemble a shift in the Beveridge curve.

In addition to this quantitative illustration of our mechanism, we provide some direct empirical evidence for two key elements that underlie it: (i) pro-cyclical search intensity of on-the-job searchers and (ii) cyclical composition shifts in the pool of searchers as well as in the pool of employed workers across the job ladder.

References

Acemoglu, D. (2008): *Introduction to modern economic growth.* Princeton University Press.

Barlevy, G. (2002): "The Sullying Effect of Recessions," *The Review of Economic Studies*, 69(1), 65–96.

Blatter, M., S. Muehlemann, and S. Schenker (2012): "The costs of hiring skilled workers," *European Economic Review*, 56(1), 20 – 35.

Bowlus, A. J. (1995): "Matching workers and jobs: Cyclical fluctuations in match quality," *Journal of Labor Economics*, pp. 335–350.

Brunner, M., and H. Strulik (2002): "Solution of perfect foresight saddlepoint problems: a simple method and applications," *Journal of Economic Dynamics and Control*, 26(5), 737 – 753.

Burdett, K., and M. Coles (1997): "Marriage and Class," *Quarterly Journal of Economics*, 112, 141–168.

Burdett, K., R. Imai, and R. Wright (2004): "Unstable relationships," *Frontiers in Macroeconomics*, 1(1).

Carillo-Tudela, C., B. Hobijn, P. Perkowski, and L. Visschers (2015): "Majority of Hires Never Report Looking for a Job," *FRBSF Economic Letters.*

Chodorow-Reich, G., and L. Karabarbounis (2016): "The Cyclicality of the Opportunity Cost of Employment," *Journal of Political Economy*, 124(6), 1563–1618.

Dey, M. S., and C. J. Flinn (2005): "An equilibrium model of health insurance provision and wage determination," *Econometrica*, 73(2), 571–627.

Diamond, P., and D. Fudenberg (1989): "Rational Expectations Business Cycles in Search Equilibrium," *Journal of Political Economy*, 97(3), 606–19.

Diamond, P. A. (1982): "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90(5), 881–94.

Dube, A., E. Freeman, and M. Reich (2010): "Employee Replacement Costs," Institute for research on labor and employment, working paper series, Institute of Industrial Relations, UC Berkeley.

Faberman, R. J. (2015): "Job switching and wage growth," *Chicago Fed Letter*, 337.

Faberman, R. J., A. I. Mueller, A. Sahin, and G. Topa (2017): "Job Search Behavior among the Employed and Non-Employed," *Working Paper.*

Fallick, B. C., and C. A. Fleischman (2004): "Employer-to-employer flows in the U.S. labor market: the complete picture of gross worker flows," Discussion paper, Finance and Economics Discussion Series.

Fujita, S., and G. Ramey (2009): "The Cyclicality of Separation and Job Finding Rates," *International Economic Review*, 50(2), 415–430.

Gertler, M., C. Huckfeldt, and A. Trigari (2016): "Unemployment fluctuations, match quality, and the wage cyclicality of new hires," *Unpublished manuscript.*

Golosov, M., and G. Menzio (2015): “Agency Business Cycle,” *Princeton mimeo.*

Hagedorn, M., and I. Manovskii (2008): “The cyclical behavior of equilibrium unemployment and vacancies revisited,” *The American Economic Review*, pp. 1692–1706.

Hall, R. E. (2005): “Employment fluctuations with equilibrium wage stickiness,” *American economic review*, pp. 50–65.

Haltiwanger, J., H. R. Hyatt, and E. McEntarfer (2015): “The job ladder over the business cycle,” *Vox EU*.

Haltiwanger, J. C., H. R. Hyatt, L. B. Kahn, and E. McEntarfer (2018): “Cyclical job ladders by firm size and firm wage,” *American Economic Journal: Macroeconomics*, 10(2), 52–85.

Hartman, P. (1960): “A lemma in the theory of structural stability of differential equations,” *Proceedings of the American Mathematical Society*, 610-620, 610–620.

Hornstein, A., P. Krusell, and G. L. Violante (2011): “Frictional wage dispersion in search models: A quantitative assessment,” *The American Economic Review*, 101(7), 2873–2898.

Howitt, P., and R. P. McAfee (1992): “Animal Spirits,” *The American Economic Review*, 82(3), 493–507.

Jovanovic, B. (1987): “Work, Rest, and Search: Unemployment, Turnover, and the Cycle,” *Journal of Labor Economics*, pp. 131–148.

Kaplan, G., and G. Menzio (2016): “Shopping externalities and self-fulfilling unemployment fluctuations,” *Journal of Political Economy*, 124(3), 771–825.

Kuznetsov, I. A. (1998): *Elements of applied bifurcation theory*, Applied mathematical sciences. Springer, New York.

Lamadon, T., J. Lise, C. Meghir, and J.-M. Robin (2013): “Matching, sorting and wages,” Discussion paper, Mimeo.

Lazear, E. P. (2014): “Structural or cyclic? Labor markets in recessions,” *IZA World of Labor*.

Lentz, R. (2014): “Optimal Employment Contracts with Hidden Search,” Discussion paper, National Bureau of Economic Research.

Lise, J., and J.-M. Robin (2017): “The macrodynamics of sorting between workers and firms,” *American Economic Review*, 107(4), 1104–35.

Moen, E. R., P. Nenov, and F. Sniekers (2015): “Buying first or selling first in housing markets,” Discussion paper, Mimeo.

Moscarini, G., and F. Postel-Vinay (2016): “Did the Job Ladder Fail after the Great Recession?,” *Journal of Labor Economics*, 34(S1), 55–93.

——— (2017): “The Cyclical Job Ladder,” *Annual Reviews of Economics*, forthcoming.

Mukoyama, T., C. Patterson, and A. Şahin (2018): “Job Search Behavior over the Business Cycle,” *American Economic Journal: Macroeconomics*, 10(1), 190–215.

Pissarides, C. A. (2000): *Equilibrium unemployment theory.* MIT press.

Pissarides, C. A., and B. Petrongolo (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39(2), 390–431.

Postel-Vinay, F., and J. M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Schaal, E., and M. Taschereau-Dumouchel (2015): "Coordinating Business Cycles," Discussion paper, NYU.

——— (2016): "Aggregate Demand and the Dynamics of Unemployment," *NYU*.

Schwartz, J. (2014): "The Job Search Intensity Supply Curve: How Labor Market Conditions Affect Job Search Effort," Discussion Paper 14-215, Loyola University Maryland Working Paper.

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

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

Shleifer, A. (1986): "Implementation cycles," *Journal of Political Economy*, 94(6), 1163–1190.

Sniekers, F. (2018): "Persistence and volatility of Beveridge cycles," *International Economic Review*, 59(2), 665–698.

Appendix A Omitted Derivations and Proofs

A.1 Equilibrium Value Functions

Firms believe workers take an individual action ω_t consistent with the equilibrium belief $\mathbf{\Omega}_t$, i.e., $\omega_t = \mathbf{\Omega}_t$. Wage setting requires that $\underline{E}_t = U_t$, which implies that $\dot{\underline{E}}_t = \dot{U}_t$. Using this and solving for U_t in Bellman equation (1), implies:

$$U_t = \frac{pb}{r} + \frac{\dot{U}_t}{r} \qquad \rightarrow \quad U = \frac{pb}{r}, \tag{18}$$

which follows from the fact that the first term $\frac{pb}{r}$ is a constant. Thus, this value is time-invariant, $U_t = U$, and we get $\dot{U} = \dot{\underline{E}} = 0$. We can thus solve for $\underline{E}_t$ in (2):

$$\underline{E}_t = \frac{\underline{w}_t(\mathbf{\Omega}_t) - \omega_t pk + \lambda(\omega_t)m(\theta_t(\mathbf{\Omega}_t))\overline{E}_t}{r + \lambda(\omega_t)m(\theta_t(\mathbf{\Omega}_t))}. \tag{19}$$

Further, solving for $\overline{E}_t$ in (3) implies:

$$\overline{E}_t = \frac{\overline{w}_t(\mathbf{\Omega}_t) + \delta\frac{pb}{r} + \dot{\overline{E}}_t}{r+\delta}.$$

The equilibrium wage for the high productivity job $\overline{w}_t$ is pinned down by the sequential auction framework, setting $\underline{J}_t = V_t = 0$ for the incumbent firm, for all $t \in [0, \infty)$. Since $V_t = 0$ by free entry, we also have $\dot{V}_t = 0$, and this implies that the wage in the high productivity job is time-invariant and independent of the equilibrium $\mathbf{\Omega}_t$. Solving for the wage from $\underline{J}_t = V = 0$ implies:

$$\overline{w}_t(\mathbf{\Omega}_t) = \overline{w} = p\underline{y}.$$

This further implies for the value of the worker in a high productivity job that

$$\overline{E} = \frac{p\underline{y} + \delta\frac{pb}{r}}{r+\delta},$$

where $\dot{\overline{E}}_t = 0$ and thus drops since all other terms in $\overline{E}_t$ are constants, which is why $\overline{E}_t = \overline{E}$.

Similarly, the equilibrium wage for the low productivity job $\underline{w}_t(\mathbf{\Omega}_t)$ is pinned down by the sequential auction framework setting $\underline{E}_t = U$. We use (18) and (19) to solve for $\underline{w}_t(\mathbf{\Omega}_t)$:

$$\underline{w}_t(\mathbf{\Omega}_t) = pb\left(\frac{r + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)) + \delta}{r+\delta}\right) - \frac{\lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))}{r+\delta}p\underline{y} + \mathbf{\Omega}_t pk. \tag{20}$$

Last, (5) and (6) can be written as (where we make use of $\dot{\overline{J}}_t = 0$):

$$\begin{aligned}
\overline{J} &= \frac{p\overline{y} - \overline{w}}{r+\delta} \\
\underline{J}_t &= \frac{p\underline{y} - \underline{w}_t(\mathbf{\Omega}_t) + \dot{\underline{J}}_t}{r + \delta + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))}.
\end{aligned}$$

Finally, from free entry, $V_t = 0$ for all t, and therefore (4) implies,

$$0 = -c + q(\theta_t(\mathbf{\Omega}_t)) \left[\frac{u_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} \underline{J}_t + \frac{\lambda(\mathbf{\Omega}_t)\gamma_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} \overline{J}_t \right]$$

Now using the fact that wages are set via sequential auctions as well as the equilibrium wages, and substituting all explicit solutions for the values from above, we can summarize the equilibrium Bellman equations as:

$$U = \frac{pb}{r} \tag{21}$$

$$\underline{E} = \frac{pb}{r} \tag{22}$$

$$\overline{E} = \frac{p\underline{y} + \delta \frac{pb}{r}}{r + \delta} \tag{23}$$

$$0 = -c + q(\theta_t(\mathbf{\Omega}_t)) \left[\frac{u_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} \left(\frac{p(\underline{y} - b)}{r + \delta} - \frac{pk\mathbf{\Omega}_t - \dot{\underline{J}}_t}{r + \delta + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))} \right) + \frac{\lambda(\mathbf{\Omega}_t)\gamma_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} \frac{p(\overline{y} - \underline{y})}{r + \delta} \right] \tag{24}$$

$$\underline{J}_t = \frac{p(\underline{y} - b)}{r + \delta} - \frac{pk\mathbf{\Omega}_t - \dot{\underline{J}}_t}{r + \delta + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t))} \tag{25}$$

$$\overline{J} = \frac{p(\overline{y} - \underline{y})}{r + \delta}. \tag{26}$$

A.2 Proof of Lemma 1

We want to specify conditions under which 1. there is no profitable one-shot deviation from the passive search steady state equilibrium; 2. there is no profitable one-shot deviation from the active search steady state equilibrium. We suppress time subscripts since we focus on steady states.

Proof. 1. No deviation when no one searches: $\underline{E}(0|\mathbf{0}) \geq \underline{E}(1|\mathbf{0})$.

In this case, when no one actively searches on-the-job ($\mathbf{\Omega} = \mathbf{0}$), a worker in a low productivity job deviating during an interval dt chooses $\omega = 1$ and gets a payoff

$$\underline{E}(1|\mathbf{0}) = \frac{1}{1 + rdt} \left[dt(\underline{w}(\mathbf{0}) - pk) + (1 - \delta dt)dt\lambda(1)m(\theta(\mathbf{0}))\overline{E} + (1 - \delta dt)(1 - dt\lambda(1)m(\theta(\mathbf{0})))\underline{E}(0|\mathbf{0}) + \delta dtU \right]$$

where $\overline{E} = \overline{E}(0|\mathbf{0})$ since that value is the same independent of the argument. There is no profitable deviation provided $\underline{E}(0|\mathbf{0}) \geq \underline{E}(1|\mathbf{0})$ or:

$$\underline{E}(0|\mathbf{0})(1 + rdt) \geq dt(\underline{w}(\mathbf{0}) - pk) + dt\lambda(1)(1 - \delta dt)m(\theta(\mathbf{0}))\overline{E} + [1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{0})) + dt^2\delta\lambda(1)m(\theta(\mathbf{0}))]\underline{E}(0|\mathbf{0}) + \delta dtU.$$

After subtracting $\underline{E}(0|\mathbf{0})$ from both sides, dividing by dt and taking the limit $dt \to 0$, we obtain:

$$r\underline{E}(0|\mathbf{0}) \geq \underline{w}(\mathbf{0}) - pk + \lambda(1)m(\theta(\mathbf{0}))\overline{E} + (-\delta - \lambda(1)m(\theta(\mathbf{0})))\underline{E}(0|\mathbf{0}) + \delta U.$$

Substituting the equilibrium values for $\underline{E}(0|\mathbf{0})$ (given by (22)), $\overline{E}$ (23), U (21) and $\underline{w}(\mathbf{0})$ (20) we obtain:

$$(\underline{y} - b)[\lambda(1) - \lambda(0)]m(\theta(\mathbf{0})) - k(r + \delta) \leq 0. \tag{27}$$

2. No deviation when all search: $\underline{E}(1|\mathbf{1}) \geq \underline{E}(0|\mathbf{1})$.

In this case, when all actively search on-the-job ($\mathbf{\Omega} = \mathbf{1}$), a worker in a low productivity job who deviates for an interval dt by choosing $\omega = 0$ gets a payoff

$$\underline{E}(0|\mathbf{1}) = \frac{1}{1+rdt}\left[dt\underline{w}(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\overline{E} + (1-\delta dt)(1-dt\lambda(0)m(\theta(\mathbf{1})))\underline{E}(1|\mathbf{1}) + \delta dtU\right].$$

There is no profitable deviation provided $\underline{E}(1|\mathbf{1}) \geq \underline{E}(0|\mathbf{1})$:

$$\underline{E}(1|\mathbf{1})(1+rdt) \geq dt\underline{w}(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\overline{E} + (1-\delta dt - dt\lambda(0)m(\theta(\mathbf{1})) + dt^2\delta\lambda(0)m(\theta(\mathbf{1})))\underline{E}(1|\mathbf{1}) + \delta dtU.$$

After subtracting $\underline{E}(1|\mathbf{1})$ from both sides, dividing by dt and taking the limit $dt \to 0$, we obtain:

$$r\underline{E}(1|\mathbf{1}) \geq \underline{w}(\mathbf{1}) + \lambda(0)m(\theta(\mathbf{1}))\overline{E} + (-\delta - \lambda(0)m(\theta(\mathbf{1})))\underline{E}(1|\mathbf{1}) + \delta U.$$

Substituting the equilibrium values for $\underline{E}(1|\mathbf{1})$ (given by (22)), $\overline{E}$ (23), U (21) and $\underline{w}(\mathbf{1})$ (20) we obtain:

$$(\underline{y} - b)[\lambda(1) - \lambda(0)]m(\theta(\mathbf{1})) - k(r+\delta) \geq 0. \tag{28}$$

Combining (27) and (28) gives the condition in the Lemma. □

A.3 Steady State in Mixed Strategies

Denote by $\underline{E}(\omega|\mathbf{\Omega})$ the value of playing ω for one instant dt while every one else pursues strategy $\mathbf{\Omega}$. This payoff is the same as the one-shot deviation payoff in Lemma 1.

For $\mathbf{\Omega} \in [0,1]$, mixing requires that $\underline{E}(0|\mathbf{\Omega}) = \underline{E}(1|\mathbf{\Omega})$, where these value functions refer to a dt-period play (after that instant the agents play $\omega = \mathbf{\Omega}$ again). If this condition is satisfied, then any mixed strategy ω (including $\mathbf{\Omega}$) is optimal from a worker's point of view. To see this, denote $\underline{E}(0|\mathbf{\Omega}) = \underline{E}(1|\mathbf{\Omega}) \equiv \underline{E}$. Then, any ω leaves the worker indifferent $\omega\underline{E} + (1-\omega)\underline{E} = \underline{E}$, i.e. there is an equilibrium in mixed strategies.

We now provide the details.

$$\begin{aligned}
\underline{E}(1|\mathbf{\Omega}) &= \frac{1}{1+rdt}\left[dt(\underline{w}(\mathbf{\Omega}) - pk) + (1-\delta dt)dt\lambda(1)m(\theta(\mathbf{\Omega}))\overline{E} + (1-\delta dt)(1-dt\lambda(1)m(\theta(\mathbf{\Omega})))\underline{E}(\omega|\mathbf{\Omega}) + \delta dtU\right] \\
\underline{E}(0|\mathbf{\Omega}) &= \frac{1}{1+rdt}\left[dt\underline{w}(\mathbf{\Omega}) + (1-\delta dt)dt\lambda(0)m(\theta(\mathbf{\Omega}))\overline{E} + (1-\delta dt)(1-dt\lambda(0)m(\theta(\mathbf{\Omega})))\underline{E}(\omega|\mathbf{\Omega}) + \delta dtU\right]
\end{aligned}$$

Set these values equal to each other and simplify (divide by dt and let $dt \to 0$) to obtain:

$$\lambda(0)m(\theta(\mathbf{\Omega}))\overline{E} - \lambda(0)m(\theta(\mathbf{\Omega}))\underline{E}(\omega|\mathbf{\Omega}) = -pk + \lambda(1)m(\theta(\mathbf{\Omega}))\overline{E} - \lambda(1)m(\theta(\mathbf{\Omega}))\underline{E}(\omega|\mathbf{\Omega})$$

Note that (as any equilibrium value of employment in the low-productivity job), $\underline{E}(\omega|\mathbf{\Omega}) = U = \frac{bp}{r}$. Using this, we obtain a necessary and sufficient condition for the mixing steady state to exist,

$$\theta(\mathbf{\Omega}) = m^{-1}\left(\frac{k(\delta+r)}{\lambda_1(\underline{y}-b)}\right)$$

where the RHS is the same constant as in the condition of Lemma 1. In sum, there co-exist three steady states iff

$$\theta(\mathbf{0}) \leq \theta(\mathbf{\Omega}) = m^{-1}\left(\frac{k(\delta+r)}{\lambda_1(\underline{y}-b)}\right) \leq \theta(\mathbf{1}).$$

The mixing probability $\mathbf{\Omega}$ can be found by plugging $\theta(\mathbf{\Omega}) = m^{-1}\left(\frac{k(\delta+r)}{\lambda_1(\underline{y}-b)}\right)$ into the FE condition of the firm and solving for $\mathbf{\Omega}$. We obtain the following result.

Proposition A1 *(Existence of Mixed Strategy Steady State).*
If there exist both active and passive search steady states, then there also exists a steady state in mixed strategies.

Proof. We showed in Lemma 1 that the active OJS steady state exists if

$$E(1|\mathbf{1}) \geq E(0|\mathbf{1}). \tag{29}$$

In turn, the passive OJS steady state exists if

$$E(0|\mathbf{0}) \geq E(1|\mathbf{0}). \tag{30}$$

We provided conditions in terms of exogenous parameters such that both (29) and (30) hold. So, for $\mathbf{\Omega}$ close to one,

$$E(1|\mathbf{\Omega}) \geq E(0|\mathbf{\Omega}) \tag{31}$$

but not

$$E(0|\mathbf{\Omega}) > E(1|\mathbf{\Omega}). \tag{32}$$

In turn, for $\mathbf{\Omega}$ close to zero, (32) holds (with weak inequality) but not (31) (with strict inequality). Since $E(\omega|\mathbf{\Omega}) - E(\omega|\mathbf{\Omega})$ is continuous in $\mathbf{\Omega}$, there exist a $\mathbf{\Omega} \in (0,1)$, such that $E(0|\mathbf{\Omega}) = E(1|\mathbf{\Omega})$. $\square$

A.4 Proof of Proposition 1

Proof. We first derive necessary and sufficient bounds for aggregate productivity, $p \in [p_l, p_h]$, in order for multiple steady states to exist.

Based on Lemma 1, the no-deviation conditions (27) and (28) at equality define the θ-bounds for multiplicity,

$$\theta_l = m^{-1}\left(\frac{k(\delta + r)}{\lambda_1(\underline{y} - b)}\right) = \theta_h$$

where θ_l is the lowest tightness that sustains the equilibrium with active OJS and θ_h is the highest tightness that sustains the equilibrium with passive OJS.

To obtain these bounds in terms of productivity p, we evaluate free entry condition (24) in the steady state of active OJS at θ_l to obtain a lower bound on aggregate productivity, denoted by p_l:

$$\begin{aligned} p_l = \; & [c\lambda_1(b-\underline{y})(k(\lambda_0+\lambda_1)+\lambda_1(-b+\underline{y}))(2\delta k(\lambda_0+\lambda_1)+2k(\lambda_0+\lambda_1)r+\delta\lambda_1(-b+\underline{y}))m^{-1}((k(\delta+r))/(\lambda_1(-b+\underline{y})))]/ \\ & [k(b^3\delta\lambda_1^2 - k^2(\lambda_0+\lambda_1)^2(\delta+r)\overline{y} + k\lambda_1(\lambda_0+\lambda_1)(\delta+r)(k-\overline{y})\underline{y} - \delta k\lambda_0\lambda_1\underline{y}^2 - \delta\lambda_1^2\underline{y}^3 - \\ & b^2\lambda_1(\delta k(2\lambda_0+\lambda_1)+k(\lambda_0+\lambda_1)r+3\delta\lambda_1\underline{y}) + b(k(\lambda_0+\lambda_1)(\delta+r)(k\lambda_0+\lambda_1\overline{y}) + \\ & k\lambda_1(\delta(3\lambda_0+\lambda_1)+(\lambda_0+\lambda_1)r)\underline{y} + 3\delta\lambda_1^2\underline{y}^2))]. \end{aligned}$$

And similarly to obtain an upper bound on aggregate productivity, p_h, (where we evaluate free entry condition (24) under the passive OJS steady state at θ_h):

$$p_h = \frac{c\lambda_1(\underline{y}-b)(2\delta k\lambda_0 + 2k\lambda_0 r + \delta\lambda_1(\underline{y}-b))m^{-1}((k(\delta+r))/(\lambda_1(\underline{y}-b)))}{k(b^2\delta\lambda_1 - bk\lambda_0(\delta+r) + k\lambda_0(\delta+r)\overline{y} - 2b\delta\lambda_1\underline{y} + \delta\lambda_1\underline{y}^2)}.$$

We still need to show that $p_h > p_l$ for an open set of remaining parameters $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$. Solving $p_h - p_l > 0$ for $\overline{y}$, we obtain a sufficient condition on $\overline{y}$ under which $p_h - p_l > 0$:

$$\begin{aligned}\overline{y} \;&>\; K := [2k\lambda_0(\lambda_0+\lambda_1)r^2 + \delta^2(2k\lambda_0 + \lambda_1(3\underline{y} - 2b))(k(\lambda_0+\lambda_1) + \lambda_1(\underline{y} - b)) \\ &\quad + \delta r(4k^2\lambda_0(\lambda_0+\lambda_1) - \lambda_1^2(2\underline{y} - b)(\underline{y} - b) + k\lambda_1(\underline{y}(5\lambda_0 + 3\lambda_1) - b(4\lambda_0 + 2\lambda_1)))]/ \\ &\quad [\delta\lambda_1(\delta + r)(k(\lambda_0+\lambda_1) + \lambda_1(\underline{y} - b))].\end{aligned}$$

Thus, for fixed $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$, $\overline{y} > K$ is sufficient for $p_h - p_l > 0$. Further note that for fixed $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$, K is finite and thus the result holds for all $\overline{y} \in (K, \infty)$. Finally, since K is continuous in $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$, the result not only holds for a fixed vector of remaining parameters $(\lambda_0, \lambda_1, \underline{y}, \overline{y}, k, c, b, r, \delta)$ but for an open set of them. □

A.5 Proof of Proposition 2

Proof. We first derive necessary and sufficient bounds for match productivity, $\overline{y} \in [\overline{y}_l(\underline{y}), \overline{y}_h(\underline{y})]$ (for any given $\underline{y}$), in order for multiple steady states to exist. We can explicitly compute these bounds from a system of two equations (per bound), namely free entry (24) under active (passive) OJS and the no-deviation condition (28) (and (27)) from active (passive) OJS, that we solve for $\overline{y}$ in each case:

$$\begin{aligned}\overline{y}_l(\underline{y}) \;&=\; \underline{y} + \frac{(b-\underline{y})(k\lambda_0 + \lambda_1(-b+\underline{y}))(\delta k(\lambda_0+\lambda_1) + k(\lambda_0+\lambda_1)r + \delta\lambda_1(-b+\underline{y}))}{k(\lambda_0+\lambda_1)(\delta+r)(k(\lambda_0+\lambda_1) + \lambda_1(-b+\underline{y}))} \\ &\quad + \frac{c\lambda_1(-b+\underline{y})(2\delta k(\lambda_0+\lambda_1) + 2k(\lambda_0+\lambda_1)r + \delta\lambda_1(-b+\underline{y}))m^{-1}((k(\delta+r))/(\lambda_1(-b+\underline{y})))}{k^2(\lambda_0+\lambda_1)p(\delta+r)} \\ \overline{y}_h(\underline{y}) \;&=\; \frac{kp(-b^2\delta\lambda_1 + bk\lambda_0(\delta+r) + 2b\delta\lambda_1\underline{y} - \delta\lambda_1\underline{y}^2) + c\lambda_1(-2\delta k\lambda_0 - 2k\lambda_0 r + \delta\lambda_1(b-\underline{y}))(b-\underline{y})m^{-1}((k(\delta+r))/(\lambda_1(-b+\underline{y})))}{k^2\lambda_0 p(\delta+r)}.\end{aligned}$$

We still need to show that $\overline{y}_h(\underline{y}) > \overline{y}_l(\underline{y})$ for an open set of parameters $(\lambda_0, \lambda_1, \underline{y}, p, k, c, b, r, \delta)$. Solving $\overline{y}_h(\underline{y}) - \overline{y}_l(\underline{y}) > 0$ for c, we obtain:

$$c \;>\; \hat{K} := \frac{kp(k^2\lambda_0(\lambda_0+\lambda_1)r + \delta(k\lambda_0 + \lambda_1(-b+\underline{y}))(k(\lambda_0+\lambda_1) + \lambda_1(-b+\underline{y})))}{\delta\lambda_1^2(-k(\lambda_0+\lambda_1) + \lambda_1(b-\underline{y}))(b-\underline{y})m^{-1}((k(\delta+r))/(\lambda_1(-b+\underline{y})))}$$

Thus, for fixed $(\lambda_0, \lambda_1, \underline{y}, p, k, b, r, \delta)$, $c > \hat{K}$ is sufficient for $\overline{y}_h(\underline{y}) > \overline{y}_l(\underline{y})$. Further note that for fixed $(\lambda_0, \lambda_1, \underline{y}, p, k, b, r, \delta)$, $\hat{K}$ is finite and thus the result holds for all $c \in (\hat{K}, \infty)$. Finally, since $\hat{K}$ is continuous in $(\lambda_0, \lambda_1, \underline{y}, p, k, b, r, \delta)$, the result not only holds for a fixed vector of parameters $(\lambda_0, \lambda_1, \underline{y}, p, k, c, b, r, \delta)$ but for an open set of them. □

A.6 Proof of Proposition 3

Proof. Each of the items in the proposition hinges on the fact that $\theta(\mathbf{1}) \geq \theta(\mathbf{0})$, which follows from Lemma 1.

1. This follows from Lemma 1 (i.e. necessary condition for multiplicity $\theta(\mathbf{1}) \geq \theta(\mathbf{0})$) and

$$\theta(\mathbf{\Omega})\frac{u + \lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega})}{u} = \Theta(\mathbf{\Omega})$$

where $\frac{u+\lambda\gamma}{u} = 2 - \frac{\delta}{\delta + \lambda m(\theta)}$ is increasing in $\mathbf{\Omega}$ since $\lambda(\mathbf{1})m(\theta(\mathbf{1})) \geq \lambda(\mathbf{0})m(\theta(\mathbf{0}))$.

2. From (BC), $u(\mathbf{1}) \leq u(\mathbf{0})$ immediately follows from $\theta(\mathbf{1}) \geq \theta(\mathbf{0})$ and since $m(\theta)$ is increasing in $\mathbf{\Omega}$.
3. We return to 3. after proving 8.

4. EE flows are defined as:

$$\begin{aligned} EE(\mathbf{\Omega}) &= \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))\gamma(\mathbf{\Omega}) \\ &= \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))\frac{\delta m(\theta(\mathbf{\Omega}))}{(\delta + m(\theta(\mathbf{\Omega})))(\delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))}, \end{aligned}$$

where we used (γC). Then $EE(\mathbf{1}) \geq EE(\mathbf{0})$ provided

$$\begin{aligned} \frac{\delta\lambda(\mathbf{1})m(\theta(\mathbf{1}))^2}{(\delta + m(\theta(\mathbf{1})))(\delta + \lambda(\mathbf{1})m(\theta(\mathbf{1})))} - \frac{\delta\lambda(\mathbf{0})m(\theta(\mathbf{0}))^2}{(\delta + m(\theta(\mathbf{0})))(\delta + \lambda(\mathbf{0})m(\theta(\mathbf{0})))} &\geq 0 \\ \delta^2\left(\lambda(\mathbf{1})m(\theta(\mathbf{1}))^2 - \lambda(\mathbf{0})m(\theta(\mathbf{0}))^2\right) + \lambda(\mathbf{0})\lambda(\mathbf{1})m(\theta(\mathbf{0}))m(\theta(\mathbf{1}))\left[m(\theta(\mathbf{1})) - m(\theta(\mathbf{0}))\right] & \\ +m(\theta(\mathbf{0}))m(\theta(\mathbf{1}))\delta\left[\lambda(\mathbf{1})m(\theta(\mathbf{1})) - \lambda(\mathbf{0})m(\theta(\mathbf{0}))\right] &\geq 0, \end{aligned}$$

which is holds since $\lambda(\mathbf{1}) > \lambda(\mathbf{0})$ and under multiplicity $m(\theta(\mathbf{1})) \geq m(\theta(\mathbf{0}))$.

5. Inequality $\lambda(\mathbf{1})\gamma(\mathbf{1})/s(\mathbf{1}) > \lambda(\mathbf{0})\gamma(\mathbf{0})/s(\mathbf{0})$ follows from $\lambda(\mathbf{1})m(\theta(\mathbf{1})) > \lambda(\mathbf{0})m(\theta(\mathbf{0}))$ and $\lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega})/s(\mathbf{\Omega}) = 1 - u(\mathbf{\Omega})/s(\mathbf{\Omega}) = 1 - (\delta^2 + \delta\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))/(\delta^2 + 2\delta\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))$ where

$$\frac{\partial(u(\mathbf{\Omega})/s(\mathbf{\Omega}))}{\partial\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))} = -\frac{\delta^3}{(\delta^2 + 2\delta\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))^2} < 0.$$

6. Using (BC) and (γC), we obtain $\frac{\gamma(\mathbf{\Omega})}{1-u(\mathbf{\Omega})} = \frac{\delta}{\delta+\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))}$. By Lemma 1, $\theta(\mathbf{\Omega})$ is increasing in $\mathbf{\Omega}$ (and also $\lambda(\mathbf{1}) > \lambda(\mathbf{0})$ by assumption), and thus $\frac{\gamma(\mathbf{\Omega})}{1-u(\mathbf{\Omega})}$ is lower when employed workers search actively, $\mathbf{\Omega} = 1$, compared to when they do not, $\mathbf{\Omega} = 0$.

7. The mean-min wage ratio can be re-formulated as $Mm(\mathbf{\Omega}) = \frac{\gamma(\mathbf{\Omega})}{1-u(\mathbf{\Omega})}\left(1 - \frac{\overline{w}(\mathbf{\Omega})}{\underline{w}(\mathbf{\Omega})}\right) + \frac{\overline{w}(\mathbf{\Omega})}{\underline{w}(\mathbf{\Omega})}$. We want to provide conditions under which Mm is increasing in $\mathbf{\Omega}$. We have (treating $\mathbf{\Omega}$ with some abuse as continuous here)

$$\frac{\partial Mm(\mathbf{\Omega})}{\partial\mathbf{\Omega}} = \frac{\partial\frac{\gamma(\mathbf{\Omega})}{1-u(\mathbf{\Omega})}}{\partial\mathbf{\Omega}}\left(1 - \frac{\overline{w}(\mathbf{\Omega})}{\underline{w}(\mathbf{\Omega})}\right) + \left(\frac{\gamma(\mathbf{\Omega})}{1-u(\mathbf{\Omega})}\right)\frac{\partial\frac{\overline{w}(\mathbf{\Omega})}{\underline{w}(\mathbf{\Omega})}}{\partial\mathbf{\Omega}}.$$

The first term is positive by [5.] of this Proposition. So we need to discipline the second term, where

$$\frac{\partial\frac{\overline{w}(\mathbf{\Omega})}{\underline{w}(\mathbf{\Omega})}}{\partial\mathbf{\Omega}} = -\frac{(\delta+r)\underline{y}[k(\delta+r) + (b-\underline{y})(\lambda\frac{\partial m}{\partial\theta}\frac{\partial\theta}{\partial\mathbf{\Omega}} + \frac{\partial\lambda}{\partial\mathbf{\Omega}}m)]}{((\delta+r)(b+k\mathbf{\Omega}) + (b-\underline{y})\lambda m(\theta))^2}$$

which has ambiguous sign. A sufficient condition for this expression to be positive is that k is small ($k \to 0$), since then $0 \geq (b-\underline{y})(\lambda(\partial m/\partial\theta)(\partial\theta/\partial\mathbf{\Omega}) + \partial\lambda/\partial\mathbf{\Omega}m)$, provided that the value of a job, $\underline{J}_t$, in (25) is non-negative (an assumption that we maintain throughout).

8. To show that the conventional Beveridge Curve (BC), which gives v as a function of u, shifts out in the boom, it suffices that v increases in $\lambda(\mathbf{\Omega})$ *for any given* u. Differentiating (BC) w.r.t. λ while keeping u fixed and solving for $\partial v/\partial\lambda$ yields (assuming $u \in (0,1]$ such that $m' > 0$)

$$\frac{\partial v}{\partial\lambda(\mathbf{\Omega})} = \frac{\gamma v}{\gamma\lambda(\mathbf{\Omega}) + u},$$

which is positive.

3. We know from [2.] of this Proposition that $u(\mathbf{1}) \leq u(\mathbf{0})$, and from [8.] that for any given u, v is higher under active OJS (outward shift of Beveridge Curve). Because (BC) is downward sloping, it

follows that also for $u(\mathbf{1}) \leq u(\mathbf{0})$ it must be the case that $v(\mathbf{1}) > v(\mathbf{0})$.

9. Since $\lambda(\boldsymbol{\Omega})$ is increasing in $\boldsymbol{\Omega}$, it suffices to show that the derivative of v w.r.t $\lambda(\boldsymbol{\Omega})$ is non-negative *for any given s*. Differentiating (BC^s) w.r.t. $\lambda(\boldsymbol{\Omega})$ while keeping s fixed and solving for $\partial v/\partial\lambda$ yields

$$\frac{\partial v}{\partial \lambda(\boldsymbol{\Omega})} = \frac{\delta s m(\theta(\boldsymbol{\Omega}))(\delta + m(\theta(\boldsymbol{\Omega})))}{(\delta^2(1-\lambda(\boldsymbol{\Omega})) + 2\delta\lambda(\boldsymbol{\Omega})m(\theta(\boldsymbol{\Omega})) + 2\lambda^2 m(\theta(\boldsymbol{\Omega}))^2)m'(\theta(\boldsymbol{\Omega}))},$$

which is positive for $\lambda(\boldsymbol{\Omega}) \leq 1$. □

A.7 Dynamic Equilibrium

A.7.1 Local Stability: Derivations

To analyze the dynamic properties, we take the following dynamic equilibrium equations into account,

$$\dot{u}_t = \delta(1-u_t) - u_t m(\theta_t) \tag{33}$$

$$\dot{\gamma}_t = u_t m(\theta_t) - (\delta + \lambda_t m(\theta_t))\gamma_t \tag{34}$$

$$\dot{\underline{J}}_t = -(p\underline{y} - \underline{w}_t) + \underline{J}_t(r + \delta + \lambda_t m(\theta_t)) \tag{35}$$

where (33) describes unemployment dynamics, (34) gives the dynamics for employed workers after a UE transition and (35) describes how the value of a filled job evolves over time. All time-varying values and variables in this system depend on agents' beliefs about how profitable OJS is, i.e. on the path of $\{\boldsymbol{\Omega}\}_{t\geq 0}$, but we suppress this dependence to simplify notation.

It will be more convenient to work with $\dot{\theta}_t$ instead of $\dot{\underline{J}}_t$, so we first transform the equation for $\dot{J}_t$ into an equation in $\dot{\theta}_t$. Notice that from the free entry condition we can find an expression for $\underline{J}_t$:

$$\underline{J}_t = \frac{c}{q(\theta_t)}\frac{u_t + \lambda_t\gamma_t}{u_t} - \frac{\lambda_t\gamma_t}{u_t}\overline{J}_t \tag{36}$$

Take the time derivative of $\underline{J}_t$ (taking into account $q(\theta_t) = m(\theta_t)/\theta_t)$) to obtain:

$$\begin{aligned}
\dot{\underline{J}}_t &= \dot{\theta}_t\frac{c}{m(\theta_t)^2}(m(\theta_t) - \theta_t m'(\theta_t))\frac{u_t + \lambda_t\gamma_t}{u_t} + \dot{u}_t\frac{\lambda_t\gamma_t}{u_t^2}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right) - \dot{\gamma}_t\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right) \\
&= \dot{\theta}_t\frac{c}{m(\theta_t)}(1 - \eta(\theta_t))\frac{u_t + \lambda_t\gamma_t}{u_t} + \dot{u}_t\frac{\lambda_t\gamma_t}{u_t^2}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right) - \dot{\gamma}_t\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right)
\end{aligned} \tag{37}$$

where we define the elasticity of the matching function as $\eta(\theta) = \frac{\theta m'(\theta)}{m(\theta)}$.

Plug the expressions for $\dot{\underline{J}}_t$, (37), and for $\underline{J}_t$ from free entry (36), into (35) to obtain,

$$\begin{aligned}
&\dot{\theta}_t\frac{c}{m(\theta_t)}(1 - \eta(\theta_t))\frac{u_t + \lambda_t\gamma_t}{u_t} + \dot{u}_t\frac{\lambda_t\gamma_t}{u_t^2}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}\right) - \dot{\gamma}_t\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right) \\
= \; & -(p\underline{y} - \underline{w}_t) + \left(\frac{c}{q(\theta_t)}\frac{u_t + \lambda_t\gamma_t}{u_t} - \frac{\lambda_t\gamma_t}{u_t}\overline{J}_t\right)(r + \delta + \lambda_t m(\theta_t))
\end{aligned}$$

and solve for $\dot{\theta}$, to obtain:

$$\begin{aligned}
\dot{\theta}_t = \frac{m(\theta_t)u_t}{c(1-\eta(\theta_t))(u_t + \lambda_t\gamma_t)} \times \; & \left[\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right)\left(-\dot{u}_t\frac{\gamma_t}{u_t} + \dot{\gamma}_t\right) - (p\underline{y} - \underline{w}_t)\right. \\
& \left. + \left(\frac{c}{q(\theta_t)}\frac{u_t + \lambda_t\gamma_t}{u_t} - \frac{\lambda_t\gamma_t}{u_t}\overline{J}_t\right)(r + \delta + \lambda_t m(\theta_t))\right].
\end{aligned}$$

So our dynamic system is given by (12)-(14) in the main text, which we re-state here for convenience:

$$
\begin{aligned}
\dot{u}_t = &\quad \delta(1-u_t) - u_t m(\theta_t) && (38)\\
\dot{\gamma}_t = &\quad u_t m(\theta_t) - (\delta + \lambda_t m(\theta_t))\gamma_t && (39)\\
\dot{\theta}_t = &\quad \frac{m(\theta_t)u_t}{c\,(1-\eta(\theta_t))\,(u_t+\lambda_t\gamma_t)} \times \Bigg[\frac{\lambda_t}{u_t}\left(-\frac{\theta_t c}{m(\theta_t)} + \overline{J}_t\right)\left(-\dot{u}_t\frac{\gamma_t}{u_t} + \dot{\gamma}_t\right) - (p\underline{y} - \underline{w}_t) \\
&\quad + \left(\frac{c}{q(\theta_t)}\frac{u_t+\lambda_t\gamma_t}{u_t} - \frac{\lambda_t\gamma_t}{u_t}\overline{J}_t\right)(r+\delta+\lambda_t m(\theta_t))\Bigg]. && (40)
\end{aligned}
$$

To analyze the local stability of system (38)-(40), we further have to specify the system's Jacobian,

$$
J^*(\mathbf{\Omega}) = \begin{bmatrix}
\frac{\partial \dot{u}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\
\frac{\partial \dot{\gamma}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\
\frac{\partial \dot{\theta}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})}
\end{bmatrix}
$$

whose entries (which are evaluated at steady state $\mathbf{x}(\mathbf{\Omega}) := (u(\mathbf{\Omega}), \gamma(\mathbf{\Omega}), \theta(\mathbf{\Omega}))$) are given by:

$$
\begin{aligned}
\frac{\partial \dot{u}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= -(\delta + m(\theta(\mathbf{\Omega})))\\
\frac{\partial \dot{u}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= 0\\
\frac{\partial \dot{u}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= -u(\mathbf{\Omega})m'(\theta(\mathbf{\Omega}))\\
\frac{\partial \dot{\gamma}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= m(\theta(\mathbf{\Omega}))\\
\frac{\partial \dot{\gamma}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= -(\delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))\\
\frac{\partial \dot{\gamma}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= m'(\theta(\mathbf{\Omega}))(u(\mathbf{\Omega}) - \lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega}))\\
\frac{\partial \dot{\theta}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= \frac{m(\theta(\mathbf{\Omega}))\lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega})}{c(1-\eta(\theta(\mathbf{\Omega})))(u(\mathbf{\Omega})+\lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega}))} \times \Bigg[\left(-\frac{\theta(\mathbf{\Omega})c}{m(\theta(\mathbf{\Omega}))} + \overline{J}\right)\Bigg(\frac{\gamma(\mathbf{\Omega})}{u(\mathbf{\Omega})}[r + 2\delta + (\lambda(\mathbf{\Omega})+1)m(\theta(\mathbf{\Omega}))] + m(\theta(\mathbf{\Omega}))\Bigg)\\
\frac{\partial \dot{\theta}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= \frac{m(\theta(\mathbf{\Omega}))\lambda(\mathbf{\Omega})}{c(1-\eta(\theta(\mathbf{\Omega})))(u(\mathbf{\Omega})+\lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega}))} \times \left[-\left(-\frac{\theta(\mathbf{\Omega})c}{m(\theta(\mathbf{\Omega}))} + \overline{J}\right)(r + 2\delta + 2\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega})))\right]\\
\frac{\partial \dot{\theta}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} &= \frac{m(\theta(\mathbf{\Omega}))u(\mathbf{\Omega})}{c(1-\eta(\theta(\mathbf{\Omega})))(u(\mathbf{\Omega})+\lambda(\mathbf{\Omega})\gamma(\mathbf{\Omega}))} \times \Bigg[\lambda(\mathbf{\Omega})m'(\theta(\mathbf{\Omega}))\Bigg(-\frac{pk\mathbf{\Omega}}{r+\delta+\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))}\\
&\qquad + \frac{\gamma(\mathbf{\Omega})(1-\lambda(\mathbf{\Omega})) + u(\mathbf{\Omega})}{u(\mathbf{\Omega})}\left(-\frac{\theta(\mathbf{\Omega})c}{m(\theta(\mathbf{\Omega}))} + \overline{J}\right)\Bigg)\Bigg] + r + \delta + \lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))
\end{aligned}
$$

The linearized system (around the steady state) of differential equations is then given by

$$
\begin{bmatrix} \dot{u}_t \\ \dot{\gamma}_t \\ \dot{\theta}_t \end{bmatrix} = \begin{bmatrix}
\frac{\partial \dot{u}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{u}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\
\frac{\partial \dot{\gamma}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\gamma}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} \\
\frac{\partial \dot{\theta}_t}{\partial u_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \gamma_t}\Big|_{\mathbf{x}(\mathbf{\Omega})} & \frac{\partial \dot{\theta}_t}{\partial \theta_t}\Big|_{\mathbf{x}(\mathbf{\Omega})}
\end{bmatrix}
\begin{bmatrix} u_t - u(\mathbf{\Omega}) \\ \gamma_t - \gamma(\mathbf{\Omega}) \\ \theta_t - \theta(\mathbf{\Omega}) \end{bmatrix}.
$$

A.7.2 Proof of Lemma 2

Proof. This proof closely follows the proof from Lemma 1.

1. No deviation when no one searches: given $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{0}\}$, it must be that $\underline{E}_t(0|\mathbf{0}) \geq \underline{E}_t(1|\mathbf{0})$ for all $t \in [0, \infty)$. In this case, when no one actively searches on-the-job, a worker, who is in a low productivity job and deviates during an interval dt by choosing $\omega = 1$, gets a payoff

$$\underline{E}_t(1|\mathbf{0}) = \frac{1}{1+rdt} \Big[dt(\underline{w}_t(\mathbf{0}) - pk) + (1-\delta dt)dt\lambda(1)m(\theta_t(\mathbf{0}))\overline{E}_{t+dt} \\ +(1-\delta dt)(1-dt\lambda(1)m(\theta_t(\mathbf{0})))\underline{E}_{t+dt}(0|\mathbf{0}) + \delta dt U_{t+dt}\Big]$$

where the only difference compared to the expression in Lemma 1 is that the variables carry time subscripts indicating that we allow for the economy to be out of steady state. Noticing, however, that under sequential auctions bargaining, $\overline{E}_{t+dt}(0|\mathbf{0}) = \overline{E}(0|\mathbf{0})$ is a constant and independent of t, and similarly that $\underline{E}_{t+dt}(0|\mathbf{0}) = U_{t+dt} = U$ is a time-independent constant (which we can hence write as $\underline{E}(0|\mathbf{0})$), the no-deviation condition $\underline{E}_t(0|\mathbf{0}) \geq \underline{E}_t(1|\mathbf{0})$ is almost identical to the one in steady state $\underline{E}(0|\mathbf{0}) \geq \underline{E}(1|\mathbf{0})$ (except that θ_t and thus also w_t are time-dependent), namely

$$\underline{E}(0|\mathbf{0})(1+rdt) \geq dt(\underline{w}_t(\mathbf{0}) - pk) + dt\lambda(1)(1-\delta dt)m(\theta_t(\mathbf{0}))\overline{E} + [1 - \delta dt - dt\lambda(1)m(\theta_t(\mathbf{0})) + dt^2\delta\lambda(1)m(\theta_t(\mathbf{0}))]\underline{E}(0|\mathbf{0}) + \delta dt U.$$

After subtracting $\underline{E}(0|\mathbf{0})$ from both sides, dividing by dt and taking the limit $dt \to 0$, we obtain:

$$r\underline{E}(0|\mathbf{0}) \geq \underline{w}_t(\mathbf{0}) - pk + \lambda(1)m(\theta_t(\mathbf{0}))\overline{E} + (-\delta - \lambda(1)m(\theta_t(\mathbf{0})))\underline{E}(0|\mathbf{0}) + \delta U.$$

This leads to the same upper bound on $\theta_t(\mathbf{0})$ as in steady state, but here needs to hold for all $t \in [0, \infty)$:

$$\theta_t(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y} - b)}\right). \tag{41}$$

If and only if (41) holds for all $t \in [0, \infty)$, the equilibrium with passive OJS exists.

2. No deviation when everyone searches: given $\{\mathbf{\Omega}_t\}_{t\geq 0} = \{\mathbf{1}\}$, it must be that $\underline{E}_t(1|\mathbf{1}) \geq \underline{E}_t(0|\mathbf{1})$ for all $t \in [0, \infty)$. We proceed in the analogous way to point 1. and arrive at the following no-deviation condition, which gives a lower bound for $\theta_t(\mathbf{1})$ for all $t \in [0, \infty)$:

$$\theta_t(\mathbf{1}) \geq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y} - b)}\right). \tag{42}$$

If and only if (42) holds for all $t \in [0, \infty)$, the equilibrium with active OJS exists. □

Appendix B Data Appendix

B.1 Monthly flow rates and flows

We use data provided by IPUMS-CPS from 1996 to 2016. We follow the dating convention as in Fallick and Fleischman (2004) and refer to a flow from month t to month $t+1$ as a month t flow. We denote by $EE_{t,t+1}$ the EE flow from t to $t+1$ and similarly for $UE_{t,t+1}$ and $EU_{t,t+1}$. Labor market *flow rates* are defined as $\frac{\sum EE_{t,t+1}}{\sum E_t}$, $\frac{\sum UE_{t,t+1}}{\sum U_t}$ and $\frac{\sum EU_{t,t+1}}{\sum E_t}$ respectively, where $\sum$ denotes the sum of sample weights for all observations in the respective categories. In our model, however, all flows are relative to the labor force. We therefore report the *normalized flows* (by the labor force) $\frac{\sum EE_{t,t+1}}{\sum LF_t}$, $\frac{\sum UE_{t,t+1}}{\sum LF_t}$ and $\frac{\sum EU_{t,t+1}}{\sum LF_t}$.

B.2 Quarterly Flow Rates, Flows, Seasonal Adjustment and De-Trending

Quarterly Flows: In order to obtain quarterly estimates of our monthly flows we first convert them to instantaneous rates. Let X_t be the flow of interest in month t. Then the instantaneous flow is given by $X_t^{inst} = -\log(1 - X_t)$. This gives us an instantaneous rate for each month. Thereafter we convert these into quarterly estimates using $X_t^{quarterly} = 1 - \exp(-3X_t^{inst})$. Since, we have this quarterly estimate for *each* month t, we average $X_t^{quarterly}$ for the 3 months of a quarter to arrive at the final quarterly rate.

Seasonal Adjustment: Once we have the monthly and quarterly data on flows, flow rates and wage measures, we seasonally adjust all time series using the X-13 ARIMA program of the US Census Bureau.

De-trending: All de-trended time series are computed using the HP filter with smoothing parameter 1600 for quarterly data and 129,600 for monthly data.

B.3 Construction of the Variables Used in Calibration and Quantitative Exercises

- Worker flows: In the model we denote the flows by EE_t, UE_t, EU_t and they correspond to $\frac{\sum EE_{t,t+1}}{\sum LF_t}$, $\frac{\sum UE_{t,t+1}}{\sum LF_t}$ and $\frac{\sum EU_{t,t+1}}{\sum LF_t}$ in the data, see Section B.1 above for how we construct them.
- Unemployment rate u_t is the quarterly average of the CPS unemployment rate (which we seasonally adjusted). We obtain the vacancy rate v_t from the JOLTS data (again seasonally adjusted).
- Matching rate of workers m: $m(\theta_t) = UE_t/u_t$.
- Effective measure of on-the-job searchers, $\gamma_t\lambda_t$: $\gamma_t\lambda_t = EE_t/m(\theta_t)$, where we use both EE_t and $m(\theta_t) = UE_t/u_t$, both obtained above.
- Effective labor market tightness $\theta_t = v_t/s_t$: v_t is vacancy data from JOLTS and $s_t = u_t + \lambda_t\gamma_t$ can be computed from the effective measures of on-the-job searchers and the unemployment rate.
- Stock of on-the-job searchers γ_t: See Section 5.4, and in particular equation (17) (main text).
- Stock of employed workers in high productivity jobs: $\xi_t = 1 - u_t - \gamma_t$.
- Conditional matching rate for unemployed workers $\kappa_t = u_t/(u_t + \lambda_t\gamma_t)$, where $\gamma_t\lambda_t = EE_t/m(\theta_t)$ (see above).
- Wage dispersion $\frac{\overline{w}}{\underline{w}}$: see Section B.4 below.
- EE flows that are associated with wage growth $EE_{t,t+1}^{(Wage_{t+1} > Wage_t)}$: see Section B.5 below.

B.4 The Mean-Min Wage Ratio

In the CPS, in each period t, a new cohort enters for 16 sample periods. We distinguish calendar time t and sample time τ. At any calendar time t, a cohort may be in any of possible the sample times $\tau \in \{1, ..., 16\}$. Each individual answers the CPS questions in sample months $\tau = 1, ..., 4$ and $\tau = 13, ..., 16$, and during times $\tau = 5, ..., 12$, they do not. So for each participant, we have 8 monthly observations over a 16 month period. Let X denote the variable of interest and let $X_{t,\tau}$ be the time and sample specific variable for each cohort (or if the variable of interest is a flow, it is the change between t and $t+1$ and τ and $\tau+1$). A flow $X_{t,\tau}$ between time t and $t+1$ can comprise of observations from 6 cohorts, who transition between the sample months 1–2, 2–3, 3–4, 13–14, 14–15, 15–16. Earnings are

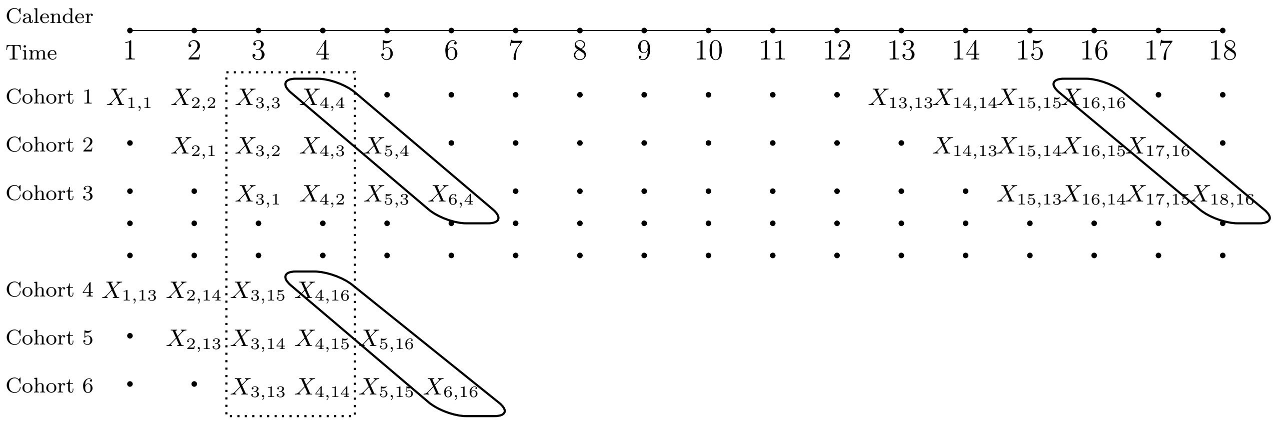


Figure 10: Timeline.

only recorded in sample periods $\tau = 4$ and 16 (variable *earnweek*), which corresponds to the two months when each cohort is in the outgoing rotation (earner study) group in the CPS sample.[36]

Figure 10 illustrates this structure. Calendar time t is given by the top row. Sample time is given by τ in $X_{t,\tau}$. Consider for example $X = EE$, the EE flow. Then EE_t equals the sum of flows for the 6 cohorts marked by the dotted rectangle and is given by $\sum_{\tau \in T} EE_{3,\tau}$ where $T = \{1, 2, 3, 13, 14, 15\}$. Wages are observed in the diagonal sections with $\tau = \{4, 16\}$.

To compute the Mm wage ratio, we need an estimate for wage dispersion between high and low-productivity jobs, $\frac{\overline{w}_{t+1}}{\underline{w}_{t+1}}$, for all t. Guided by our model, we compute $\overline{w}_{t+1}$ as the average wage associated with an EE move, EE_t. In turn, we compute $\underline{w}_{t+1}$ as the average wage associated with an UE move, UE_t. Here, we focus on the EE moves (and $\overline{w}$), the logic for UE moves (and $\underline{w}$) is similar. In any given month t we first find all individuals who make an EE_t switch. The challenge however is that in the CPS, among all these individuals only one fourth of the employed are in the Earnings Sample in period $t + 1$. However, we can impute the wages of those who have an EE_t switch but for whom we do not directly observe wages in period $t + 1$.

B.4.1 Imputing Wages for EE Movers Who Are Not in The Earnings Study

Let the set of all EE_t movers for whom we do not observe wages in $t + 1$ be denoted by $EE^{w'}_{t,\tau}$ where w' denotes unobserved wages in period $t + 1$. We observe wages only for the $EE_{t,\tau}$ movers who have $\tau + 1 = 4$ or 16, and denote this set by $EE^{w}_{t,\tau}$ where w denotes observed wages in period $t + 1$. For illustration, we focus on only 4 consecutive calendar months $t \in \{3, 4, 5, 6\}$ and on $EE_{3,\tau}$. Then, the set $EE^{w}_{3,\tau}$ comprises of members of cohort 1 and 4 who are in the Earning Study at the time of the flow $t = 4$. On the other hand, the set $EE^{w'}_{3,\tau}$ comprises of 2 subsets.

1. Individuals in cohort 2 and 5 with $EE_{3,\tau}$ and $\tau + 1 = 3$ or 15 whose wages may be observed in the next period when they enter the earner study in period $t = 5$.

2. Individuals in cohort 3 and 6 with $EE_{3,\tau}$ and $\tau + 1 = 2$ or 14 whose wages may be observed two periods ahead when they enter the earner study in period $t = 6$.

As we can see in the timeline in Figure 10 we can find wages for subset 1 and subset 2 in calendar

[36]We use variable *earnweek* as our measure of wage which is a part of the earner study. This variable tracks wages and reports how much the respondent usually earns per week at their current job, before deductions. Approximately one quarter of the CPS sample is in the earner study each month.

time 5 and 6 respectively. But in order to use those wage observations, we need to ensure that they are representative of the wages in calendar time $t = 4$, when the $EE_{3,\tau}$ switch was made.

To do so we create two subsets $EE^{w' \to w} \subset EE^{w'}_{3,\tau}$ and $EE^{w' \to \neq EE} \subset EE^{w'}_{3,\tau}$, where the first subset denotes individuals for whom we observe a wage in the future and the second subset denotes individuals who do not make another EE switch before this wage is observed.

Then $EE^{w' \to w} \cap EE^{w' \to \neq EE}$ is the required set of observations who made a $EE_{3,\tau}$ switch without observed wages in $t = 4$ but whose wages when observed in the future can be imputed as a wage for $t = 4$ (the underlying assumption is that there was no wage change between $t = 4$ and the time when the wage was observed). We append these observations to the set $EE^{w}_{3,\tau}$. Then the final set of moves where wages are either observed or imputed for calendar time $t = 4$ is given by:

$$\text{Expanded Set } EE_3 = EE^{w}_{3,\tau} \cup \left(EE^{w' \to w} \cap EE^{w' \to \neq EE}\right)$$

Once we have this set of all individuals with an $EE_{3,\tau}$ move and a wage in $t = 4$, we simply take the weighted average of these wages using the earnings weight in the earner study as our sample weights.

B.5 Measure of EE Flows to a Higher Wage

In order to account for the EE transitions that are consistent with the model, our objective in this section is to obtain a measure of EE_t flows that are associated with *higher* wages. Given the limited wage data in the CPS, assessing the wage before and after an EE move requires some assumptions in order to use the two available wage observations as the wage before and after such a move.

B.5.1 Focus On EE Moves Where Both Previous Wage and New Wage Are Observed

We illustrate our method for obtaining EE flows to higher wages between any two calendar months t and $t + 1$. To make the exposition clear, consider flows $EE_{15,\tau}$. First, unlike the previous section note that this computation requires a measure of w in both time t and $t + 1$. Second, note that only cohorts with $\tau \in \{13, 14, 15, 16\}$ are relevant for our purpose, as only for these cohorts we have two wage observations. As a result, we only focus on Cohort 1, 2 and 3 from Figure 10. We keep our notation $X_{t,\tau}$, and we need data for sample months, $\tau \in \{4, 13, 14, 15, 16\}$ for these 3 cohorts which corresponds to calendar time $t \in \{4, 5, 6, 13, 14, 15, 16, 17, 18\}$. This is because for cohort 1, $w_{4,4}$ and $w_{16,16}$ are the two instances we observe their wages. Similarly for cohort 2, $w_{5,4}$ and $w_{17,16}$ are the two times we observe their wages. Finally, for cohort 3, $w_{6,4}$ and $w_{18,16}$ are the two times we observe their wages. It is useful to split the calendar times t in two subsets $t \in \{4, 5, 6\}$ and $t \in \{13, 14, 15, 16, 17, 18\}$. The first set records the first calendar time that we observe the wages of the three cohorts. These give us the measure of the previous wage (i.e. prior to the EE move) for each cohort. The lower bound of set 2 is given by the earliest possible EE move made by cohort 1, while the upper bound of set 2 is given by the last possible EE move made by cohort 3.

B.5.2 Imputing Previous Wage and New Wage

We define the previous wage to be the wage in the job in E_t before making a move to job in E_{t+1} where the wage is that in the new job E_{t+1} after the job switch. For each cohort at the time of the EE_t switch we want the first observed wage to be representative of the wage w_t in job E_t. At the same time we want the second wage observed to be representative of the wage w_{t+1} in the job E_{t+1}. This entails different restrictions on the 3 cohorts which are explained below, again for simplicity consider the $EE_{15,\tau}$ move.

Our key assumption is the following: for all three cohorts we assume that there is no EE move in the 8 months between the two rotation phases. That is, we assume there is no EE_t move for

$t \in \{4,5,6,7,8,9,10,11,12\}$ for cohort 1 and no EE_t move for $t \in \{6,7,8,9,10,11,12,13,14\}$ for cohort 3. This is a strong assumption but it is unavoidable for recovering the wage before and after an EE move from the CPS data. In addition, in order to correctly impute wages for the two months before and after the EE switch, we restrict our set of observations for each cohort to individuals for whom we have the data for all 5 relevant sample months.

For cohort 1, the $EE_{15,15}$ move coincides with $\tau+1 = 16$ when they are in the earner study. Therefore, the second observed wage $w_{16,16}$ is representative of the job in $E_{16,16}$ for all EE switchers. Now, for the first observed $Wage_{4,4}$ to be representative of the job $E_{15,15}$, we restrict our sample only to observations where we do not observe an $EE_{13,13}$ switch or an $EE_{14,14}$, but we do observe an $EE_{15,15}$ switch.

For cohort 2, the $EE_{15,14}$ move does not have a wage observation in $\tau + 1$ but only in $\tau + 2 = 16$. Therefore, the second observed $w_{17,16}$ may not be representative of the job in $E_{16,15}$ for all EE switchers. Moreover, the first observed wage $w_{5,4}$ may not be representative of the job in $E_{15,14}$. Now, for $w_{5,4}$ to be representative of the job $E_{15,14}$ before the switch and $w_{17,16}$ to be representative of the job $E_{16,15}$ after the switch, we restrict our sample only to observations where we do not observe an $EE_{14,13}$ switch or an $EE_{16,15}$, but we do observe an $EE_{15,14}$ move.

Similarly, for cohort 3, the $EE_{15,13}$ move does not satisfy $\tau+1 = 16$. Therefore, the second observed wage $w_{18,16}$ may not be representative of the job in $E_{16,14}$ for all EE switchers. However, given our assumptions the first observed $w_{6,4}$ is representative of the job in $E_{15,13}$. Now, for the second observed wage $w_{18,16}$ to be representative of the job $E_{16,14}$, we restrict our sample only to observations where we do not observe an $EE_{16,14}$ switch or an $EE_{17,15}$, but we do observe an $EE_{15,13}$ move.

B.5.3 Collecting the Observations

Once we have the $EE_{15,\tau}$ moves $\forall \tau \in \{13, 14, 15\}$ with the property that the two wages observed for each cohort can be considered as the previous wage in job $E_{15,\tau}$ and new wage in job $E_{16,\tau+1}$, we append the observations from these three cohorts to have a larger dataset with EE flows for calendar time 15 and 16, with wage observations before and after the EE move. Note that this larger dataset for wage comparisons for EE moves can be created for $\forall t$. Finally, we use this set to compute $EE_t^{(w_{t+1}>w_t)}$, i.e. the set of EE moves to higher wages $\forall t$.

Our notation in the *model* for EE rates, EE_t, should be read as capturing the normalization by the labor force. The adjustment to normalize relative to the labor force is done as follows. Because we do not have a relevant measure of labor force here, we cannot use $\frac{\sum EE_{t,t+1}}{\sum LF_t}$ to compute the normalized EE flows.[37] However, what we can compute is the fraction of the observed EE moves that transition to a higher wage out of all EE moves for which we can observe pre and post switch wages. This is given by:

$$\%EE_{t,t+1}^{(w_{t+1}>w_t)} = \frac{\sum \mathbf{1}_{(w_{t+1}>w_t)} EE_{t,t+1}}{\sum \mathbf{1}_{(w_{t+1}<w_t)} EE_{t,t+1} + \sum \mathbf{1}_{(w_{t+1}>w_t)} EE_{t,t+1}}$$

Where $\mathbf{1}$ is an indicator variable that takes a value 1 when the condition stated is true and 0 otherwise. Thereafter, we compute the economy-wide $EE_{t,t+1}^{(w_{t+1}>w_t)}$ switches to a higher wage by

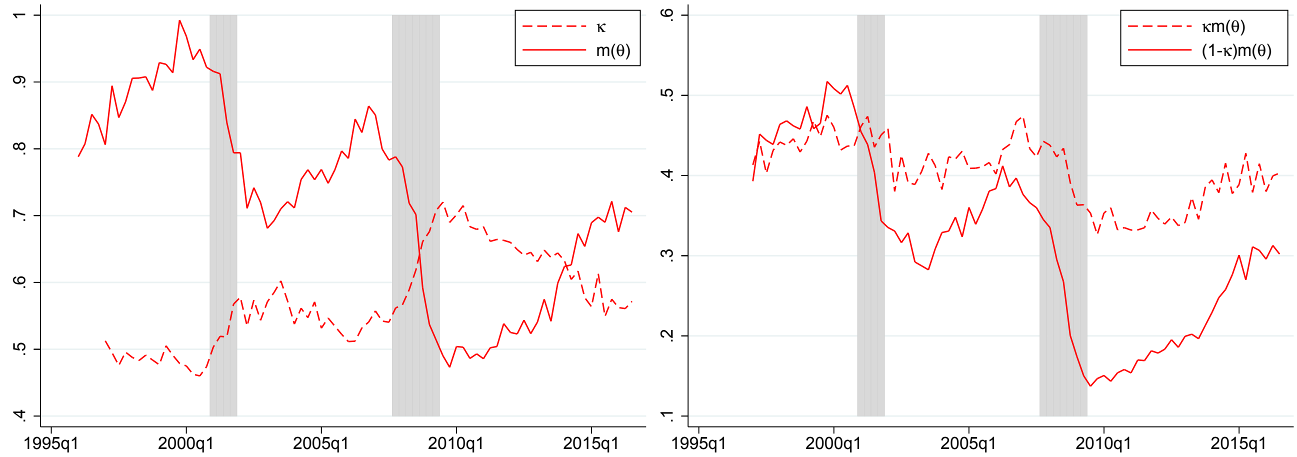
$$EE_{t,t+1}^{(w_{t+1}>w_t)} = \%EE_{t,t+1}^{(w_{t+1}>w_t)} \times EE_{t,t+1}$$

where $EE_{t,t+1}$ are *all* EE moves we observe between month t and $t+1$. $EE_{t,t+1}^{(w_{t+1}>w_t)}$ is the measure of EE flows that we target in the calibration.

[37] We cannot properly measure the labor force here as there exist $EE_{t,t+1}$ switchers for whom we do not observe the relevant wages, which is why we dropped them from our dataset. It is therefore not clear how to measure $LF = E + U$.

Appendix C Additional Empirical and Simulation Results

C.1 Flow Rates, Search Intensity in ATUS and of Those NILF



(a) Overall Matching Rate m and Matching Rate of Unemployed Workers Conditional on Matching κ.

(b) Matching Rates of Employed Workers $(1-\kappa)m$ and of Unemployed Workers κm

Figure 11: Matching Rates of Employed and Unemployed Workers.

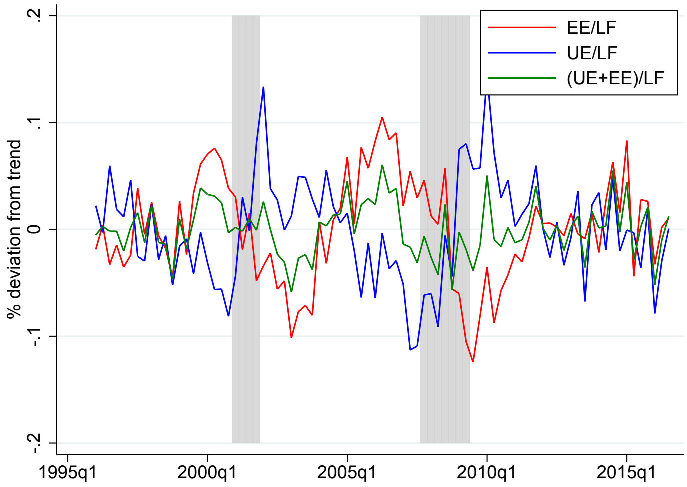
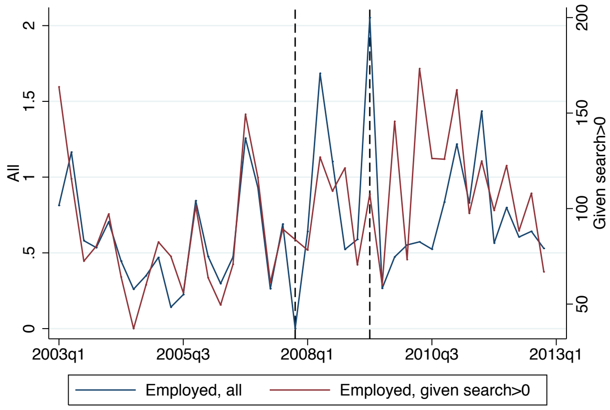
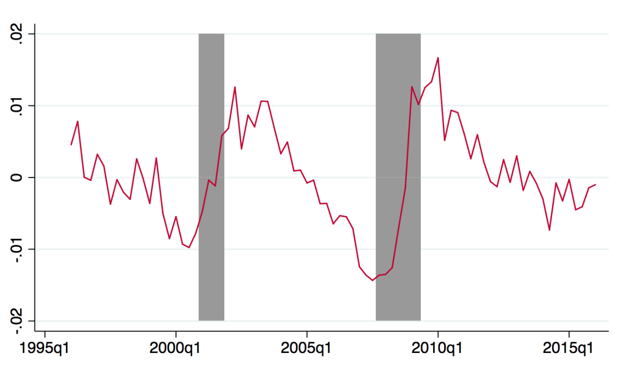


Figure 12: Labor Market Flows (De-Trended).

Search Intensity of those Not in the Labor Force Similar to how we calculate the search intensity of the employed, we calculate the search intensity of those Not in the Labor Force. Denote their stock by N and their flow into employment by NE. Then $NE = \lambda_N m(\theta) N$ where λ_N is the search intensity and $m(\theta) = UE/U$ is the matching probability derived from the Unemployment to Employment flow. Therefore we calculate search intensity as $\lambda_N = \frac{NE \cdot U}{UE \cdot N}$.



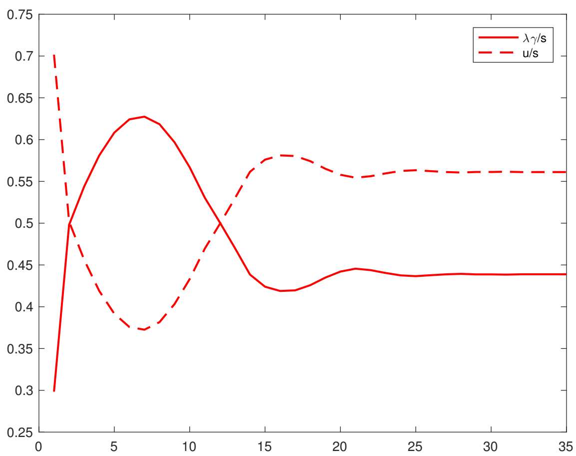
(a) Search Intensity of the Employed from the American Time Use Survey (ATUS); In minutes per day.



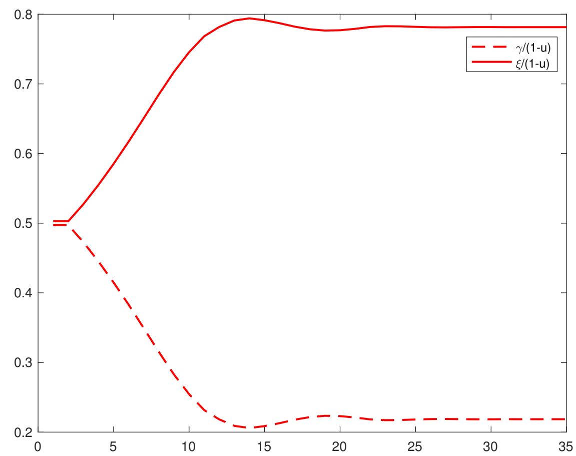
(b) Search Intensity of those Not in the Labor Force (log-deviations from trend; de-trending by HP filter). See detailed explanations above.

Figure 13: Search Intensity.

C.2 Additional Figures: Transition Dynamics



(a) Composition of Searchers over Time.



(b) Composition of Employed Workers over Time.

Figure 14: Transition Dynamics of Worker Composition (Based on Calibrated Model; Time Measured in Quarters).

Online Appendix – Not for Publication

Online Appendix I: Alternative Model Specifications

We will analyze four generalizations of the model: In the first three we show that the multiplicity of steady state equilibria is not specific to our simplified baseline model. In the fourth one, we show that a dynamic model with *anticipated* expectation shocks has similar quantitative implications as our model with unanticipated expectation shocks.

1. We generalize our baseline model in two dimensions: first, we allow for *stochastic* match-specific productivity upgrades (i.e. when employed or unemployed workers search, they all have the same probability of obtaining a $\overline{y}$ match (with probability π) or $\underline{y}$ match (with probability $1-\pi$). Second, the number of rounds of OJS is no longer restricted to one. Workers can search as many rounds as they like.
2. We keep the deterministic match-specific productivity upgrade from OJS (as in baseline model) but allow for an unrestricted number of rounds of OJS.
3. We introduce ex-ante heterogeneity of firms (there is free entry into low productivity vacancies that produce $\underline{y}$ and free entry into high productivity vacancies that produce $\overline{y}$ once matched). Both unemployed and employed workers meet these different vacancies with the same probabilities. We keep the assumption of a single round of OJS from the baseline model.
4. We introduce *anticipated* expectation shocks into the dynamic version of our baseline model.

1. Stochastic Match-Specific Types with Unrestricted Number of Search Rounds

Suppose that any realized match is of productivity $\overline{y}$ with probability π and of productivity $\underline{y}$ with probability $1-\pi$. Once matched, the worker decides whether to continue to search with high or low intensity. Now the exact history matters for the continuation. We focus on steady state equilibrium, which is why we drop time-subscripts. There are 6 possible states that workers can be in, depending on their history:

1. u: unemployment
2. γ_L: employed out of u in a $\underline{y}$ job
3. γ_H: employed out of u in a $\overline{y}$ job
4. γ_{LL}: employed in an $\underline{y}$ job after being employed in $\underline{y}$ in a previous period
5. γ_{LH}: employed in a $\overline{y}$ job after being employed in a $\underline{y}$ job or after having received at least once an outside offer from a $\underline{y}$ job.
6. γ_{HH}: employed in a $\overline{y}$ job after being employed in $\overline{y}$ in a previous period.

We denote the corresponding values of (un)employment and wages by:

1. U: Value of being unemployed.
2. E_L: Value of being employed at $\underline{y}$ out of U, i.e. coming from u into γ_L. Get wage w_L.

3. E_H: Value of being employed at $\overline{y}$ out of U, i.e. coming from u into γ_H. Get wage w_H.
4. E_{LL}: Value of being employed at $\underline{y}$ after a match with at least one other $\underline{y}$ and no $\overline{y}$, i.e. coming from γ_L into γ_{LL}. Get wage w_{LL}.
5. E_{LH}: Value of being employed at $\overline{y}$ after having matched with at least one $\underline{y}$ and not matched with another $\overline{y}$, i.e. coming from γ_L or γ_{LL}, or after not having been matched with any $\underline{y}$ before but having received an outside offer from $\underline{y}$ while matched with a $\overline{y}$-job (i.e. coming from γ_H). Get wage w_{LH}.
6. E_{HH}: Value of being employed at $\overline{y}$ after matching with at least one other $\overline{y}$, i.e. coming from γ_H or γ_{LH}. Get wage w_{HH}.

Apart from the stochastic (as opposed to deterministic) upgrading of productivity as workers search, there is a second change compared to the baseline model: We now leave the number of rounds of OJS unrestricted, i.e. workers will stop searching only once they extract the entire rents from a match.

As in the baseline model, ties are broken in favor of the incumbent: in case of a tie, assume the worker does not move or equivalently, that there is an ε moving cost. For the values of a filled job, we adopt a similar notation below.

We now make a distinction between the action of the different actively searching workers, depending on which state they are in. Denote by $\mathbf{\Omega}_L$ the search intensity of all workers in low productivity jobs in state γ_L (and the individual search intensity is ω_L). Likewise for $\mathbf{\Omega}_H, \mathbf{\Omega}_{LL}$ and $\mathbf{\Omega}_{LH}$ (note that $\mathbf{\Omega}_{HH} = 0$ since at this stage the worker extracts the entire match surplus). The overall search intensity that enters market tightness θ is a vector $\mathbf{\Omega} = [\mathbf{\Omega}_L, \mathbf{\Omega}_H, \mathbf{\Omega}_{LL}, \mathbf{\Omega}_{LH}]$. Then we will determine the number of searchers as $s = u + (\lambda_0 + \lambda_1 \mathbf{\Omega})\gamma^T$ where γ^T is the (transposed) vector $[\gamma_L, \gamma_H, \gamma_{LL}, \gamma_{LH}]$.

Steady State Equilibrium. The laws of motion in steady state satisfy:

$$
\begin{aligned}
1 &= u + \gamma_L + \gamma_H + \gamma_{LL} + \gamma_{LH} + \gamma_{HH} \\
0 &= um(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_L[\delta + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))] \\
0 &= um(\theta(\mathbf{\Omega}))\pi - \gamma_H[\delta + \lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))] \\
0 &= \gamma_L\lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_{LL}[\delta + \lambda(\mathbf{\Omega}_{LL})m(\theta(\mathbf{\Omega}))\pi] \\
0 &= \gamma_L\lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))\pi + \gamma_{LL}\lambda(\mathbf{\Omega}_{LL})m(\theta(\mathbf{\Omega}))\pi + \gamma_H\lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_{LH}[\delta + \lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))\pi] \\
0 &= \gamma_H\lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))\pi + \gamma_{LH}\lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))\pi - \gamma_{HH}\delta
\end{aligned}
$$

We can write the value functions of the worker in steady state as:

$$
\begin{aligned}
rU &= pb + m(\theta(\mathbf{\Omega}))((1-\pi)E_L + \pi E_H - U) \\
rE_L &= w_L(\mathbf{\Omega}) - \omega_L pk + \lambda(\omega_L)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LL} + \pi E_{LH} - E_L) - \delta(E_L - U) \\
rE_H &= w_H(\mathbf{\Omega}) - \omega_H pk + \lambda(\omega_H)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LH} + \pi E_{HH} - E_H) - \delta(E_H - U) \\
rE_{LL} &= w_{LL}(\mathbf{\Omega}) - \omega_{LL} pk + \lambda(\omega_{LL})m(\theta(\mathbf{\Omega}))(\pi(E_{LH} - E_{LL})) - \delta(E_{LL} - U) \\
rE_{LH} &= w_{LH}(\mathbf{\Omega}) - \omega_{LH} pk + \lambda(\omega_{LH})m(\theta(\mathbf{\Omega}))(\pi(E_{HH} - E_{LH})) - \delta(E_{LH} - U) \\
rE_{HH} &= w_{HH}(\mathbf{\Omega}) - \delta(E_{HH} - U)
\end{aligned}
$$

The steady state values of a filled job, high or low productivity (and depending on the workers previous position), are given by:

$$
\begin{aligned}
rJ_L &= p\underline{y} - w_L(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))[(1-\pi)(J_{LL}-J_L) + \pi(V-J_L)] - \delta(J_L - V) \\
rJ_H &= p\overline{y} - w_H(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))[(1-\pi)(J_{LH}-J_H) + \pi(J_{HH}-J_H)] - \delta(J_H - V) \\
rJ_{LL} &= p\underline{y} - w_{LL}(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_{LL})m(\theta(\mathbf{\Omega}))\pi(V-J_{LL}) - \delta(J_{LL} - V) \\
rJ_{LH} &= p\overline{y} - w_{LH}(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))(\pi(J_{HH}-J_{LH})) - \delta(J_{LH} - V) \\
rJ_{HH} &= p\overline{y} - w_{HH}(\mathbf{\Omega}) - \delta(J_{HH} - V)
\end{aligned}
$$

The value of a vacancy to the firm is

$$
rV = -c + q(\theta(\mathbf{\Omega}))\left[\frac{u}{s(\mathbf{\Omega})}(1-\pi)J_L + \frac{u}{s(\mathbf{\Omega})}\pi J_H + \frac{\lambda(\mathbf{\Omega}_L)\gamma_L\pi + \lambda(\mathbf{\Omega}_{LL})\gamma_{LL}\pi}{s(\mathbf{\Omega})}J_{LH} - V\right]
$$

where s denotes the number of searchers: $s(\mathbf{\Omega}) = u + \lambda(\mathbf{\Omega}_L)\gamma_L + \lambda(\mathbf{\Omega}_H)\gamma_H + \lambda(\mathbf{\Omega}_{LL})\gamma_{LL} + \lambda(\mathbf{\Omega}_{LH})\gamma_{LH}$. The value of a vacancy V reflects that workers stay with the incumbent firm in case the worker draws the same match-specific productivity.

Then the equilibrium tightness can be written as:

$$
\theta(\mathbf{\Omega}) = \frac{v}{s(\mathbf{\Omega})}.
$$

We now derive the steady state equilibrium values (where $\omega = \mathbf{\Omega}$):

$$
\begin{aligned}
U &= \frac{pb}{r} \\
E_L &= \frac{w_L(\mathbf{\Omega}) - \omega_L pk + \lambda(\omega_L)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LL} + \pi E_{LH}) + \delta U}{r + \lambda(\omega_L)m(\theta(\mathbf{\Omega})) + \delta} \\
E_H &= \frac{w_H(\mathbf{\Omega}) - \omega_H pk + \lambda(\omega_H)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LH} + \pi E_{HH}) + \delta U}{r + \lambda(\omega_H)m(\theta(\mathbf{\Omega})) + \delta} \\
E_{LL} &= \frac{w_{LL}(\mathbf{\Omega}) - \omega_{LL} pk + \lambda(\omega_{LL})m(\theta(\mathbf{\Omega}))\pi E_{LH} + \delta U}{r + \lambda(\omega_{LL})m(\theta(\mathbf{\Omega}))\pi + \delta} \\
E_{LH} &= \frac{w_{LH}(\mathbf{\Omega}) - \omega_{LH} pk + \lambda(\omega_{LH})m(\theta(\mathbf{\Omega}))\pi E_{HH} + \delta U}{r + \lambda(\omega_{LH})m(\theta(\mathbf{\Omega}))\pi + \delta} \\
E_{HH} &= \frac{w_{HH}(\mathbf{\Omega}) + \delta U}{r + \delta}
\end{aligned}
$$

The equilibrium wage is set to the maximum amount that the 'losing' firm (which is the firm who loses the worker to another firm or that unsuccessfully tries to poach the worker) would be able to pay for the worker. In turn, if the worker is hired out of unemployment than the wage is set such that he is indifferent between remaining unemployed and taking the job. We assume that firms commit to this wage setting protocol. That implies:

$$
\begin{aligned}
w_L(\mathbf{\Omega}) &: \quad E_L = U \\
w_H(\mathbf{\Omega}) &: \quad E_H = U \\
w_{LL}(\mathbf{\Omega}) &: \quad J_{LL} = V \quad \rightarrow \quad w_{LL} = p\underline{y} \\
w_{LH}(\mathbf{\Omega}) &: \quad J_L = V \quad \rightarrow \quad w_{LH} = p\underline{y} \\
w_{HH}(\mathbf{\Omega}) &: \quad J_{HH} = V \quad \rightarrow \quad w_{HH} = p\overline{y}
\end{aligned}
$$

This implies that the values for the firm (imposing free entry of firms, $V = 0$) are given by:

$$0 = -c + q(\theta(\boldsymbol{\Omega}))\left[\frac{u}{s}(1-\pi)J_L + \frac{u}{s}\pi J_H + \frac{\lambda(\boldsymbol{\Omega}_L)\gamma_L\pi + \lambda(\boldsymbol{\Omega}_{LL})\gamma_{LL}\pi}{s}J_{LH}\right] \tag{43}$$

$$J_L = \frac{p\underline{y} - w_L(\boldsymbol{\Omega}) + \lambda(\boldsymbol{\Omega}_L)m(\theta(\boldsymbol{\Omega}))(1-\pi)J_{LL}}{r + \lambda(\boldsymbol{\Omega}_L)m(\theta(\boldsymbol{\Omega})) + \delta} \tag{44}$$

$$J_H = \frac{p\overline{y} - w_H(\boldsymbol{\Omega}) + \lambda(\boldsymbol{\Omega}_H)m(\theta(\boldsymbol{\Omega}))((1-\pi)J_{LH} + \pi J_{HH})}{r + \lambda(\boldsymbol{\Omega}_H)m(\theta(\boldsymbol{\Omega})) + \delta} \tag{45}$$

$$J_{LL} = 0$$

$$J_{LH} = \frac{p\overline{y} - w_{LH}(\boldsymbol{\Omega})}{r + \lambda(\Omega_{LH})m(\theta(\boldsymbol{\Omega}))\pi + \delta} \tag{46}$$

$$J_{HH} = 0$$

Multiple Steady State Equilibria. In this model, there are four states at which workers make search decisions. Hence, there are $2^4 = 16$ potential equilibria. Multiplicity obtains if any two out of 16 equilibria coexist for some parameter values. We can verify numerically that the following two equilibria can coexist: one where workers search intensively in states γ_L, γ_H (this we will interpret as a 'boom') and one where workers search intensively only in γ_H (which we interpret as a 'recession'). That is workers tend to search more intensively at the lower rungs of the job ladder (in states γ_H and also γ_L) compared to the upper rungs (in states γ_{LH} and how γ_{LL}). The intuition is that coming out of unemployment, workers have a particularly low wage, which is why they put more search effort. In short we will indicate the equilibrium strategy in the first equilibrium by 1100 and the strategy in the second equilibrium by 0100. (Of course, other equilibria may coexist as well. Establishing multiplicity of all other equilibria is beyond the purpose of this exercise which is to show that the mechanism that leads to multiplicity does not hinge on the particular job ladder that we assume in the baseline model.)

We need to verify four no-deviation conditions (corresponding to the four states), where, as in the baseline model, it suffices to consider *one-shot deviations*. This implies that we need to check the conditions where in each state we have to take the two equilibria into account that we are focusing on:

1. In state γ_L: $E_L(0|\mathbf{0100}) \geq E_L(1|\mathbf{0100})$ and $E_L(1|\mathbf{1100}) \geq E_L(0|\mathbf{1100})$
2. In state γ_H: $E_H(1|\mathbf{0100}) \geq E_H(0|\mathbf{0100})$ and $E_H(1|\mathbf{1100}) \geq E_H(0|\mathbf{1100})$
3. In state γ_{LL}: $E_{LL}(0|\mathbf{0100}) \geq E_{LL}(1|\mathbf{0100})$ and $E_{LL}(0|\mathbf{1100}) \geq E_{LL}(1|\mathbf{1100})$
4. In state γ_{LH}: $E_{LH}(0|\mathbf{0100}) \geq E_{LH}(1|\mathbf{0100})$ and $E_{LH}(0|\mathbf{1100}) \geq E_{LH}(1|\mathbf{1100})$

In more detail:

1. Check Deviations from Search Decisions in γ_L:

1.1 $E_L(0|\mathbf{0100}) \geq E_L(1|\mathbf{0100})$ (no deviation from 'no search') if

$$\begin{aligned} E_L(0|\mathbf{0100})(1+rdt) \geq dt(w_L(\mathbf{0100}) - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{0100}))[(1-\pi)E_{LL}(0|\mathbf{0100}) + \pi E_{LH}(0|\mathbf{0100})] \\ +[1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{0100})) + dt^2\delta\lambda(1)m(\theta(\mathbf{0100}))]E_L(0|\mathbf{0100}) + \delta dtU. \end{aligned}$$

After subtracting $E_L(0|\mathbf{0100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$\begin{aligned} rE_L(0|\mathbf{0100}) \geq w_L(\mathbf{0100}) - pk + \lambda(1)m(\theta(\mathbf{0100}))[(1-\pi)E_{LL}(0|\mathbf{0100}) + \pi E_{LH}(0|\mathbf{0100})] \\ +(-\delta - \lambda(1)m(\theta(\mathbf{0100})))E_L(0|\mathbf{0100}) + \delta U. \end{aligned}$$

1.2 $E_L(1|\mathbf{1100}) \geq E_L(0|\mathbf{1100})$ (no deviation from 'search') if

$$E_L(1|\mathbf{1100})(1+rdt) \geq dtw_L(\mathbf{1100}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1100}))[(1-\pi)E_{LL}(0|\mathbf{1100}) + \pi E_{LH}(0|\mathbf{1100})] \\ +(1-\delta dt - dt\lambda(0)m(\theta(\mathbf{1100})) + dt^2\delta\lambda(0)m(\theta(\mathbf{1100})))E_L(1|\mathbf{1100}) + \delta dtU.$$

After subtracting $E_L(1|\mathbf{1100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_L(1|\mathbf{1100}) \geq w_L(\mathbf{1100}) + \lambda(0)m(\theta(\mathbf{1100}))[(1-\pi)E_{LL}(0|\mathbf{1100}) + \pi E_{LH}(0|\mathbf{1100})] \\ +(-\delta - \lambda(0)m(\theta(\mathbf{1100})))E_L(1|\mathbf{1100}) + \delta U.$$

2. Check Deviations from Search Decisions in γ_H:

2.1 $E_H(1|\mathbf{0100}) \geq E_H(0|\mathbf{0100})$ (no deviation from 'search') if

$$E_H(1|\mathbf{0100})(1+rdt) \geq dtw_H(\mathbf{0100}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{0100}))[(1-\pi)E_{LH}(0|\mathbf{0100}) + \pi E_{HH}] \\ +[1-\delta dt - dt\lambda(0)m(\theta(\mathbf{0100})) + dt^2\delta\lambda(0)m(\theta(\mathbf{0100}))]E_H(1|\mathbf{0100}) + \delta dtU.$$

After subtracting $E_H(1|\mathbf{0100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_H(1|\mathbf{0100}) \geq w_H(\mathbf{0100}) + \lambda(0)m(\theta(\mathbf{0100}))[(1-\pi)E_{LH}(0|\mathbf{0100}) + \pi E_{HH}] \\ +(-\delta - \lambda(0)m(\theta(\mathbf{0100})))E_H(1|\mathbf{0100}) + \delta U.$$

2.2 $E_H(1|\mathbf{1100}) \geq E_H(0|\mathbf{1100})$ (no deviation from 'search') if

$$E_H(1|\mathbf{1100})(1+rdt) \geq dtw_H(\mathbf{1100}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1100}))[(1-\pi)E_{LH}(0|\mathbf{1100}) + \pi E_{HH}] \\ +(1-\delta dt - dt\lambda(0)m(\theta(\mathbf{1100})) + dt^2\delta\lambda(0)m(\theta(\mathbf{1100})))E_H(1|\mathbf{1100}) + \delta dtU.$$

After subtracting $E_H(1|\mathbf{1100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_H(1|\mathbf{1100}) \geq w_H(\mathbf{1100}) + \lambda(0)m(\theta(\mathbf{1100}))[(1-\pi)E_{LH}(0|\mathbf{1100}) + \pi E_{HH}] \\ +(-\delta - \lambda(0)m(\theta(\mathbf{1100})))E_H(1|\mathbf{1100}) + \delta U.$$

3. Check Deviations from Search Decisions in γ_{LL}:

3.1 $E_{LL}(0|\mathbf{0100}) \geq E_{LL}(1|\mathbf{0100})$ (no deviation from 'no search') if

$$E_{LL}(0|\mathbf{0100})(1+rdt) \geq dt(w_{LL} - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{0100}))\pi E_{LH}(0|\mathbf{0100}) \\ +[1-\delta dt - dt\lambda(1)m(\theta(\mathbf{0100}))\pi + dt^2\delta\lambda(1)m(\theta(\mathbf{0100}))\pi]E_{LL}(0|\mathbf{0100}) + \delta dtU.$$

After subtracting $E_{LL}(0|\mathbf{0100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LL}(0|\mathbf{0100}) \geq w_{LL} - pk + \lambda(1)m(\theta(\mathbf{0100}))\pi E_{LH}(0|\mathbf{0100}) + (-\delta - \lambda(1)m(\theta(\mathbf{0100}))\pi)E_{LL}(0|\mathbf{0100}) + \delta U.$$

3.2 $E_{LL}(0|\mathbf{1100}) \geq E_{LL}(1|\mathbf{1100})$ (no deviation from 'no search') if

$$E_{LL}(0|\mathbf{1100})(1+rdt) \geq dt(w_{LL} - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{1100}))\pi E_{LH}(0|\mathbf{1100}) \\ +(1-\delta dt - dt\lambda(1)m(\theta(\mathbf{1100}))\pi + dt^2\delta\lambda(1)m(\theta(\mathbf{1100}))\pi)E_{LL}(0|\mathbf{1100}) + \delta dtU.$$

After subtracting $E_{LL}(0|\mathbf{1100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LL}(0|\mathbf{1100}) \geq w_{LL} - pk + \lambda(1)m(\theta(\mathbf{1100}))\pi E_{LH}(0|\mathbf{1100}) + (-\delta - \lambda(1)m(\theta(\mathbf{1100}))\pi)E_{LL}(0|\mathbf{1100}) + \delta U.$$

4. Check Deviations from Search Decisions in γ_{LH}:

4.1 $E_{LH}(0|\mathbf{0100}) \geq E_{LH}(1|\mathbf{0100})$ (no deviation from 'no search') if

$$E_{LH}(0|\mathbf{0100})(1 + rdt) \geq dt(w_{LH} - pk) + dt\lambda(1)(1 - \delta dt)m(\theta(\mathbf{0100}))\pi E_{HH}$$
$$+[1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{0100}))\pi + dt^2\delta\lambda(1)m(\theta(\mathbf{0100}))\pi]E_{LH}(0|\mathbf{0100}) + \delta dtU.$$

After subtracting $E_{LH}(0|\mathbf{0100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LH}(0|\mathbf{0100}) \geq w_{LH} - pk + \lambda(1)m(\theta(\mathbf{0100}))\pi E_{HH} + (-\delta - \lambda(1)m(\theta(\mathbf{0100}))\pi)E_{LH}(0|\mathbf{0100}) + \delta U.$$

4.2 $E_{LH}(0|\mathbf{1100}) \geq E_{LH}(1|\mathbf{1100})$ (no deviation from 'no search') if

$$E_{LH}(0|\mathbf{1100})(1 + rdt) \geq dt(w_{LH} - pk) + dt\lambda(1)(1 - \delta dt)m(\theta(\mathbf{1100}))\pi E_{HH}$$
$$+(1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{1100}))\pi + dt^2\delta\lambda(1)m(\theta(\mathbf{1100}))\pi)E_{LH}(0|\mathbf{1100}) + \delta dtU.$$

After subtracting $E_{LH}(0|\mathbf{1100})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LH}(0|\mathbf{1100}) \geq w_{LH} - pk + \lambda(1)m(\theta(\mathbf{1100}))\pi E_{HH} + (-\delta - \lambda(1)m(\theta(\mathbf{1100}))\pi)E_{LH}(0|\mathbf{1100}) + \delta U.$$

Analogously to the baseline model, when evaluating these 8 conditions at equalities, we obtain the bounds of the corresponding labor market tightness. We then evaluate the free entry condition (one by one) at these 8 different values of θ to obtain multiplicity bounds on primitive p. This gives us 3 potential candidates for $\overline{p}$ (i.e. it must be that $p < \overline{p}$, otherwise workers would be incentivized to *always* search, that is also in states γ_L in a recession and in γ_{LH} and γ_{LL} in both boom and recession) and 5 candidates for $\underline{p}$ (i.e. $p \geq \underline{p}$, otherwise workers would not want to search in state γ_L in a boom and in state γ_H in both boom and recession). For multiple steady state equilibria to exist it must thus be that $min(\overline{p}) \geq max(\underline{p})$. It is not possible to report an analytical solution for the p-bounds but we show numerically that there exists parameter ranges for which $min(\overline{p}) \geq max(\underline{p})$.

An example of parameter intervals, for which there exist multiple steady state equilibria (i.e. co-existence of the two equilibria **1100** and **0100**) is: $\overline{y} \in [2.16 - \Delta, 2.16 + \Delta]$, $\underline{y} \in [1 - \Delta, 1 + \Delta]1$, $b \in [0.3 - \Delta, 0.3 + \Delta]$, $\lambda_0 \in [.04 - \Delta, .04 + \Delta]$, $\lambda_1 \in [0.45 - \Delta, 0.45 + \Delta]$, $k \in [0.024 - \Delta, 0.024 + \Delta]$, $r \in [0.0235 - \Delta, 0.0235 + \Delta]$, $c \in [1 - \Delta, 1 + \Delta]$, $\alpha \in [0.37 - \Delta, 0.37 + \Delta]$, $\phi \in [1.47 - \Delta, 1.47 + \Delta]$, $\delta \in [0.0468 - \Delta, 0.0468 + \Delta]$, $\pi \in [0.8 - \Delta, 0.8 + \Delta]$ for $\Delta > 0$ small.

The mechanism that generates multiplicity in this general setup is similar to the mechanism in our baseline model (main text). Even though, there are more states (which makes the analysis different), the multiplicity is still driven by a strategic complementarity between search intensity and vacancy posting (with underlying composition externality in the pool of searchers in productive and non-productive jobs).

First of all, firms do not want to meet workers already in a $\overline{y}$ job, i.e., in states γ_H and γ_{LH}, because workers in those states will not move but stay in the incumbent firm and extract all the surplus from the match. Also, existing firms do not like their matched workers to receive offers because that instantaneously depletes their surplus. Firms therefore only generate a positive surplus hiring workers out of unemployment (creating both $\underline{y}$ and $\overline{y}$ jobs) or out of existing $\underline{y}$ jobs, i.e., out of states u, γ_L and γ_{LL} with corresponding values J_L or J_H (when hiring out of unemployment) and J_{LH} (when hiring out of employment). Now what generates the strategic complementarity between the firm's profits from opening a vacancy and the employed workers' search intensity can be read off the firm's free entry condition (43). Intuitively, this is satisfied if the value of hiring a worker out of an existing job πJ_{LH} is larger than the value from hiring out of unemployment, $(1-\pi)J_L + \pi J_H$. A necessary condition for this is that $J_{LH} > J_H$. When is this the case? Using expressions (45) and (46), this is the case whenever

$$\frac{r + \delta + \lambda(\mathbf{\Omega}_H) m(\mathbf{\Omega})\pi}{r + \delta + \lambda(\mathbf{\Omega}_{LH}) m(\mathbf{\Omega})\pi}(\overline{y} - \underline{y}) - (\overline{y} - w_H) > 0 \tag{47}$$

This condition holds whenever the duration of a J_H-job is low enough (because, in the equilibrium we are focusing on, workers search more intensively for outside offers in state γ_H compared to state γ_{LH}, and therefore $\lambda(\mathbf{\Omega}_H) > \lambda(\mathbf{\Omega}_{LH})$). If in addition to (47), $(1-\pi)J_L$ is not too high, i.e., which is the case when the productivity of a $\underline{y}$-job is relatively close to the outside option $b + k$ and/or when π is high enough, then firms prefer hiring employed workers over unemployed ones. Whenever these conditions are satisfied, then the composition of jobs changes in the boom due to endogenous search intensity, i.e. $\frac{\lambda_L \gamma_L}{s} + \frac{\lambda_{LL}\gamma_{LL}}{s}$ grows relative to $\frac{u}{s}$, and the value of opening jobs goes up. Instead, whenever the composition of searchers is more biased towards unemployed workers due to low search intensity in γ_L jobs (recession), then $\frac{\lambda_L \gamma_L}{s} + \frac{\lambda_{LL}\gamma_{LL}}{s}$ is relatively small compared to $\frac{u}{s}$ and the value of opening vacancies falls. This is at the root of the strategic complementarity and generates self-fulfilling prophecies (even though productivity upgrades on the job ladder are *not* deterministic here and workers search an *unrestricted* number of rounds – unlike in the baseline model). Note that an important assumption here is that firms commit to set wages such that outside options of the workers are matched.

2. Deterministic Match-Specific Types with Unrestricted Number of Search Periods

In this section, we extend the baseline model in the following sense. We let workers search on the job until they extract the entire output, i.e. there are *two* states of OJS instead of one. Again, we will focus on steady state equilibrium, which is why we drop all the time subscripts.

There are 4 possible states that workers can be in, depending on their history:

1. u: unemployment
2. γ_L: employed out of u in a $\underline{y}$ job
3. γ_{LH}: employed in a $\overline{y}$ job after being employed in an $\underline{y}$ job
4. γ_{HH}: employed in a $\overline{y}$ job after having received an outside offer from another $\overline{y}$ job.

We denote the corresponding values of (un)employment and wages by:

1. U: Value of being unemployed
2. E_L: Value of being employed at $\underline{y}$ out of U, i.e. coming from u into γ_L. Get wage w_L.
3. E_{LH}: Value of being employed at $\overline{y}$ after having matched with at least one $\underline{y}$ and not matched with another $\overline{y}$, i.e. coming from γ_L. Get wage w_{LH}.
4. E_{HH}: Value of being employed at $\overline{y}$ after received at least one outside offer from another $\overline{y}$, i.e. coming from γ_H. Get wage w_{HH}.

As in the baseline model, ties are broken in favor of the incumbent: in case of a tie, assume the worker does not move or equivalently, that there is an ε moving cost. For the values of a filled job, we adopt a similar notation below.

We now make a distinction between the action of the different actively searching workers, depending on which state they are in. Denote by $\mathbf{\Omega}_L$ the search intensity of all workers in low productivity jobs in state γ_L (and the individual search intensity is ω_L). Likewise for $\mathbf{\Omega}_{LH}$ (note that $\mathbf{\Omega}_{HH} = 0$ since at this stage the worker extracts the entire surplus). The overall search intensity that enters the market tightness θ is a vector $\mathbf{\Omega} = (\mathbf{\Omega}_L, \mathbf{\Omega}_{LH})$. Then we will determine the number of searchers as $s = u + (\lambda_0 + \lambda_1 \mathbf{\Omega})\gamma^T$ where γ^T is the (transposed) vector $[\gamma_L, \gamma_{LH}]$. Then $s = u + (\lambda_0 + \lambda_1 \mathbf{\Omega_L})\gamma_L + (\lambda_0 + \lambda_1 \mathbf{\Omega_{LH}})\gamma_{LH}$.

Steady State Equilibrium. The laws of motion in steady state satisfy:

$$\begin{aligned}
1 &= u + \gamma_L + \gamma_{LH} + \gamma_{HH} \\
0 &= um(\theta(\mathbf{\Omega})) - \gamma_L[\delta + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega})] \\
0 &= \gamma_L \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega})) - \gamma_{LH}[\delta + \lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))] \\
0 &= \gamma_{LH}\lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega})) - \gamma_{HH}\delta
\end{aligned}$$

We can write the value functions of the worker in steady state as:

$$\begin{aligned}
rU &= pb + m(\theta(\mathbf{\Omega}))(E_L - U) \\
rE_L &= w_L(\mathbf{\Omega}) - \omega_L pk + \lambda(\omega_L)m(\theta(\mathbf{\Omega}))(E_{LH} - E_L) - \delta(E_L - U) \\
rE_{LH} &= w_{LH}(\mathbf{\Omega}) - \omega_{LH} pk + \lambda(\omega_{LH})m(\theta(\mathbf{\Omega}))(E_{HH} - E_{LH}) - \delta(E_{LH} - U) \\
rE_{HH} &= w_{HH}(\mathbf{\Omega}) - \delta(E_{HH} - U)
\end{aligned}$$

The steady state values of a filled job, high or low productivity (and depending on the workers previous position), are given by:

$$\begin{aligned}
rJ_L &= p\underline{y} - w_L(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))(V - J_L) - \delta(J_L - V) \\
rJ_{LH} &= p\overline{y} - w_{LH}(\mathbf{\Omega}) + \lambda(\Omega_{LH})m(\theta(\mathbf{\Omega}))(J_{HH} - J_{LH}) - \delta(J_{LH} - V) \\
rJ_{HH} &= p\overline{y} - w_{HH}(\mathbf{\Omega}) - \delta(J_{HH} - V)
\end{aligned}$$

The value of a vacancy to the firm is

$$rV = -c + q(\theta(\mathbf{\Omega}))\left[\frac{u}{s(\mathbf{\Omega})}J_L + \frac{\lambda(\mathbf{\Omega}_L)\gamma_L}{s(\mathbf{\Omega})}J_{LH} - V\right]$$

where s denotes the number of searchers $s(\mathbf{\Omega}) = u + \lambda(\mathbf{\Omega}_L)\gamma_L + \lambda(\mathbf{\Omega}_{LH})\gamma_{LH}$. The value of a vacancy V reflects that workers stay with the incumbent firm in case the worker draws the same match-specific productivity. The equilibrium tightness can be written as:

$$\theta(\mathbf{\Omega}) = \frac{v}{s(\mathbf{\Omega})}.$$

We now derive the steady state equilibrium values (where $\omega = \Omega$):

$$
\begin{aligned}
U &= \frac{pb}{r} \\
E_L &= \frac{w_L(\mathbf{\Omega}) - \omega_L pk + \lambda(\omega_L) m(\theta(\mathbf{\Omega})) E_{LH} + \delta U}{r + \lambda(\omega_L) m(\theta(\mathbf{\Omega})) + \delta} \\
E_{LH} &= \frac{w_{LH}(\mathbf{\Omega}) - \omega_{LH} pk + \lambda(\omega_{LH}) m(\theta(\mathbf{\Omega})) E_{HH} + \delta U}{r + \lambda(\omega_{LH}) m(\theta(\mathbf{\Omega})) + \delta} \\
E_{HH} &= \frac{w_{HH}(\mathbf{\Omega}) + \delta U}{r + \delta}
\end{aligned}
$$

The equilibrium wage is set to the maximum amount that the “losing" firm (which is the firm who loses the worker to another firm or that unsuccessfully tries to poach the worker) would be able to pay for the worker. In turn, if the worker is hired out of unemployment than the wage is set such that he is indifferent between remaining unemployed and taking the job. That implies:

$$
\begin{aligned}
w_L(\mathbf{\Omega}): &\quad E_L = U \quad \rightarrow \\
w_L(\mathbf{\Omega}) = &\quad \frac{p\left(b(\delta+r)(\delta k\omega_L + k\omega_L(2(\lambda_0+\lambda_1)m(\Omega)+r) - (\lambda_0+\lambda_1)m(\Omega)((\lambda_0+\lambda_1)m(\Omega)\overline{y} + \underline{y})) + r^2(\delta + (\lambda_0+\lambda_1)m(\Omega) + r\right.}{b(\delta+r)(\delta+(\lambda_0+\lambda_1)m(\Omega)+r)} \\
w_{LH}(\mathbf{\Omega}): &\quad J_L = V \quad \rightarrow \quad w_{LH} = p\underline{y} \\
w_{HH}(\mathbf{\Omega}): &\quad J_{HH} = V \quad \rightarrow \quad w_{HH} = p\overline{y}
\end{aligned}
$$

This implies that the values for the firm (imposing free entry of firms, $V = 0$) are given by:

$$
\begin{aligned}
0 &= -c + q(\theta(\mathbf{\Omega}))\left[\frac{u}{s(\mathbf{\Omega})} J_L + \frac{\lambda(\mathbf{\Omega}_L)\gamma_L}{s(\mathbf{\Omega})} J_{LH}\right] \\
J_L &= \frac{p\underline{y} - w_L(\mathbf{\Omega})}{r + \lambda(\mathbf{\Omega}_L) m(\theta(\mathbf{\Omega})) + \delta} \\
J_{LH} &= \frac{p(\overline{y} - \underline{y})(\mathbf{\Omega})}{r + \lambda(\mathbf{\Omega}_{LH}) m(\theta(\mathbf{\Omega})) + \delta} \\
J_{HH} &= 0
\end{aligned}
$$

Multiple Steady State Equilibria. Now there are two states at which workers make search decisions. Hence, there are $2^2 = 4$ potential equilibria. Multiplicity obtains if any two out of the 4 equilibria coexist for some range of parameters. We can verify numerically that the following two equilibria, that we are particularly interested in, can coexist: one where workers search in *all* two states γ_L, γ_{LH} and one where workers do not search at all (i.e. in neither state). In short we will indicate the equilibrium strategy of in the first equilibrium by 11 and the strategy in the second equilibrium by 00.

We need to verify the two sets of no-deviation conditions, where, as in the baseline model, it suffices to consider *one-shot deviations*. This implies that we need to check the conditions:

1. In state γ_L: $E_L(0|\mathbf{00}) \geq E_L(1|\mathbf{00})$ and $E_L(1|\mathbf{11}) \geq E_L(0|\mathbf{11})$
2. In state γ_{LH}: $E_{LH}(0|\mathbf{00}) \geq E_{LH}(1|\mathbf{00})$ and $E_{LH}(1|\mathbf{11}) \geq E_{LH}(0|\mathbf{11})$

In more detail:

1. Check Deviations from Search Decisions in γ_L:

1.1 $E_L(0|\mathbf{00}) \geq E_L(1|\mathbf{00})$ (no deviation from 'no search') if

$$E_L(0|\mathbf{00})(1+rdt) \geq dt(w_L(\mathbf{00}) - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{00}))E_{LH}(0|\mathbf{00})$$
$$+[1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{00})) + dt^2\delta\lambda(1)m(\theta(\mathbf{00}))]E_L(0|\mathbf{00}) + \delta dtU.$$

After subtracting $E_L(0|\mathbf{00})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_L(0|\mathbf{00}) \geq w_L(\mathbf{00}) - pk + \lambda(1)m(\theta(\mathbf{00}))E_{LH}(0|\mathbf{00})$$
$$+(-\delta - \lambda(1)m(\theta(\mathbf{00})))E_L(0|\mathbf{00}) + \delta U.$$

1.2 $E_L(1|\mathbf{11}) \geq E_L(0|\mathbf{11})$ (no deviation from 'search') if

$$E_L(1|\mathbf{11})(1+rdt) \geq dtw_L(\mathbf{11}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{11}))E_{LH}(1|\mathbf{11})$$
$$+(1 - \delta dt - dt\lambda(0)m(\theta(\mathbf{11})) + dt^2\delta\lambda(0)m(\theta(\mathbf{11})))E_L(1|\mathbf{11}) + \delta dtU.$$

After subtracting $E_L(1|\mathbf{11})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_L(1|\mathbf{11}) \geq w_L(\mathbf{11}) + \lambda(0)m(\theta(\mathbf{11}))E_{LH}(1|\mathbf{11})$$
$$+(-\delta - \lambda(0)m(\theta(\mathbf{11})))E_L(1|\mathbf{11}) + \delta U.$$

2. Check Deviations from Search Decisions in γ_{LH}:

2.1 $E_{LH}(0|\mathbf{00}) \geq E_{LH}(1|\mathbf{00})$ (no deviation from 'no search') if

$$E_{LH}(0|\mathbf{00})(1+rdt) \geq dt(w_{LH} - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{00}))E_{HH}$$
$$+[1 - \delta dt - dt\lambda(1)m(\theta(\mathbf{00})) + dt^2\delta\lambda(1)m(\theta(\mathbf{00}))E_{LH}(0|\mathbf{00}) + \delta dtU.$$

After subtracting $E_{LH}(0|\mathbf{00})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LH}(0|\mathbf{00}) \geq w_{LH} - pk + \lambda(1)m(\theta(\mathbf{00}))E_{HH} + (-\delta - \lambda(1)m(\theta(\mathbf{00}))\pi)E_{LH}(0|\mathbf{00}) + \delta U.$$

2.2 $E_{LH}(1|\mathbf{11}) \geq E_{LH}(0|\mathbf{11})$ (no deviation from 'search') if

$$E_{LH}(1|\mathbf{11})(1+rdt) \geq dtw_{LH} + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{11}))E_{HH}$$
$$+(1 - \delta dt - dt\lambda(0)m(\theta(\mathbf{11})) + dt^2\delta\lambda(0)m(\theta(\mathbf{11})))E_{LH}(1|\mathbf{11}) + \delta dtU.$$

After subtracting $E_{LH}(1|\mathbf{11})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_{LH}(1|\mathbf{11}) \geq w_{LH} + \lambda(0)m(\theta(\mathbf{11}))E_{HH} + (-\delta - \lambda(0)m(\theta(\mathbf{11})))E_{LH}(1|\mathbf{11}) + \delta U.$$

Analogously to the baseline model, when evaluating these 4 conditions with equalities we obtain the corresponding labor market tightness θ for these 4 bounds. We then evaluate the free entry condition (one by one) at these 4 values of θ to obtain multiplicity bounds on primitive p. This gives us 4 potential p-bounds: two for $\overline{p}$ (i.e. it must be that $p < \overline{p}$, otherwise workers would be incentivized to *always* search, that is also in γ_L, γ_{LH} in the 'no-search' candidate equilibrium) and 2 candidates for $\underline{p}$ (i.e. $p > \underline{p}$, otherwise workers would *never* want to search in γ_L, γ_{LH} in the 'search' candidate equilibrium). Hence, for multiple steady state equilibria to exist it must be that $min(\overline{p}) > max(\underline{p})$. The analytical solution for the p-bounds are very tedious. We show numerically that there exists parameter ranges for which $min(\overline{p}) > max(\underline{p})$. An example of parameter intervals, for which there exist multiple steady state equilibria is: $\overline{y} \in [4, 4.2]$, $\underline{y} \in [1-\Delta, 1+\Delta]1$, $b \in [0.438 - \Delta, 0.438 + \Delta]$, $\lambda_0 \in [0.053 - \Delta, 0.053 + \Delta]$, $\lambda_1 \in [0.217 - \Delta, 0.217 + \Delta]$, $k \in [0.4 - \Delta, 0.4 + \Delta]$, $r \in [0.02 - \Delta, 0.02 + \Delta]$, $c \in [2.79 - \Delta, 2.79 + \Delta]$,

$\alpha \in [8.56 - \Delta, 8.56 + \Delta]$, $\phi \in [8.78 - \Delta, 8.78 + \Delta]$, $\delta \in [0.53 - \Delta, 0.53 + \Delta]$, for $\Delta > 0$ small.

3. Ex-Ante Firm Heterogeneity and Restricted Number of Search Rounds

Now we study a setting with two types of jobs $\overline{y}$ and $\underline{y}$. A firm can open vacancies of either type, where the type is *permanent* until the match is destroyed. Denote vacancies by $v \in \{\underline{v}, \overline{v}\}$. Following Lise and Robin (2017), we model the employment production technology by assuming the the cost of vacancies is increasing in the aggregate number of vacancies of each type: $c(v)$ with $c(0) \geq 0, c' > 0, c'(0) = 0, c'' = 0$.

Suppose that a worker meets a high productivity firm $\overline{y}$ with probability π and a low productivity firm $\underline{y}$ with probability $1-\pi$. Once matched, the worker decides whether to continue to search. The exact history matters for the continuation. Similar to Extension 1 of the baseline model, there are 6 possible states, depending on the worker history (since we focus on steady states we drop time subscripts):

1. u: unemployment
2. γ_L: employed out of u in a $\underline{y}$ job
3. γ_H: employed out of u in a $\overline{y}$ job
4. γ_{LL}: employed in an $\underline{y}$ job after being employed in $\underline{y}$ in a previous period
5. γ_{LH}: employed in a $\overline{y}$ job after being employed in an $\underline{y}$ job or after having received at least once an offer from a $\underline{y}$ job.
6. γ_{HH}: employed in a $\overline{y}$ job after being employed in $\overline{y}$ in a previous period.

The corresponding values of workers are given by:

1. U: Value of being unemployed
2. E_L: Value of being employed at $\underline{y}$ out of U (get wage w_L)
3. E_H: Value of being employed at $\overline{y}$ out of U (get wage w_H)
4. E_{LL}: Value of being employed at $\underline{y}$ after a match with another $\underline{y}$ (wage w_{LL})
5. E_{LH}: Value of being employed at $\overline{y}$ after first having been employed by a $\underline{y}$ or employed at $\overline{y}$ after getting an outside offer from a $\underline{y}$ (get wage w_{LH});
6. E_{HH}: Value of being employed at $\overline{y}$ after matching with another $\overline{y}$ (get wage w_{HH})

As in the baseline version of the model, we assume that search costs are prohibitively high (or the gains are too low) when the wage offer has been matched once (i.e. after one round of OJS), so that no more search occurs to increase the wage further after one round of OJS. Ties are broken in favor of the incumbent: in case of a tie, assume the worker does not move or equivalently, that there is an ε moving cost. For the values of a filled job as well as for the labor market stocks, we adopt a similar notation below.

Steady State Equilibrium. Define as π the equilibrium fraction of high type vacancies (determined below by free entry conditions): $\pi = \frac{\overline{v}}{\overline{v}+\underline{v}}$. We use a similar notation for values and stocks as in Online

Appendix I.1. The value of a filled job, high or low productivity, in steady state is given by:

$$\begin{aligned}
rJ_L &= p\underline{y} - w_L(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))[(1-\pi)(J_{LL}-J_L) + \pi(J_{LH}-J_L)] - \delta(J_L - V)\\
rJ_H &= p\overline{y} - w_H(\mathbf{\Omega}) + \lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))[(1-\pi)(J_{LH}-J_H) + \pi(J_{HH}-J_H)] - \delta(J_H - V)\\
rJ_{LL} &= p\underline{y} - w_{LL}(\mathbf{\Omega}) - \delta(J_{LL} - V)\\
rJ_{LH} &= p\overline{y} - w_{LH}(\mathbf{\Omega}) - \delta(J_{LH} - V)\\
rJ_{HH} &= p\overline{y} - w_{HH}(\mathbf{\Omega}) - \delta(J_{HH} - V)
\end{aligned}$$

We can write the value functions of the worker in steady state as:

$$\begin{aligned}
rU &= pb + m(\theta(\mathbf{\Omega}))((1-\pi)E_L + \pi E_H - U)\\
rE_L &= w_L(\mathbf{\Omega}) - \omega_L pk + \lambda(\omega_L)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LL} + \pi E_{LH} - E_L) - \delta(E_L - U)\\
rE_H &= w_H(\mathbf{\Omega}) - \omega_H pk + \lambda(\omega_H)m(\theta(\mathbf{\Omega}))((1-\pi)E_{LH} + \pi E_{HH} - E_H) - \delta(E_H - U)\\
rE_{LL} &= w_{LL}(\mathbf{\Omega}) - \delta(E_{LL} - U)\\
rE_{LH} &= w_{LH}(\mathbf{\Omega}) - \delta(E_{LH} - U)\\
rE_{HH} &= w_{HH}(\mathbf{\Omega}) - \delta(E_{HH} - U)
\end{aligned}$$

The steady state laws of motions for the labor market stocks are identical to the previous extension with match-specific types (see Online Appendix I.1),

$$\begin{aligned}
1 &= u + \gamma_L + \gamma_H + \gamma_{LL} + \gamma_{LH} + \gamma_{HH}\\
0 &= um(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_L[\delta + \lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega})]\\
0 &= um(\theta(\mathbf{\Omega}))\pi - \gamma_H[\delta + \lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega})]
\end{aligned}$$

$$\begin{aligned}
0 &= \gamma_L\lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_{LL}[\delta + \lambda(\mathbf{\Omega}_{LL})m(\theta(\mathbf{\Omega}))\pi]\\
0 &= \gamma_L\lambda(\mathbf{\Omega}_L)m(\theta(\mathbf{\Omega}))\pi + \gamma_{LL}\lambda(\mathbf{\Omega}_{LL})m(\theta(\mathbf{\Omega}))\pi + \gamma_H\lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))(1-\pi) - \gamma_{LH}[\delta + \lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))\pi]\\
0 &= \gamma_H\lambda(\mathbf{\Omega}_H)m(\theta(\mathbf{\Omega}))\pi + \gamma_{LH}\lambda(\mathbf{\Omega}_{LH})m(\theta(\mathbf{\Omega}))\pi - \gamma_{HH}\delta
\end{aligned}$$

with the only difference that $\pi = \frac{\overline{v}}{\overline{v}+\underline{v}}$ is endogenous. Then the equilibrium tightness can be written as:

$$\theta(\mathbf{\Omega}) = \frac{v}{s(\mathbf{\Omega})} = \frac{v}{u + \lambda(\mathbf{\Omega})\left[\gamma_L + \gamma_H\right]}.$$

The equilibrium wage is set to the maximum amount that the 'losing' firm (which is the firm who loses the worker to another firm or that unsuccessfully tries to poach the worker) would be able to pay for the worker. In turn, if the worker is hired out of unemployment than the wage is set such that he is indifferent between remaining unemployed and taking the job. That implies:

$$
\begin{aligned}
w_L(\mathbf{\Omega}): &\quad E_L = U \quad \to \quad w_L(\mathbf{\Omega}) = pb\left(\frac{r+\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))+\delta}{r+\delta}\right) - \frac{\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))}{r+\delta}p\underline{y} + \mathbf{\Omega}pk \\
w_H(\mathbf{\Omega}): &\quad E_H = U \quad \to \quad w_H(\mathbf{\Omega}) = pb\left(\frac{r+\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))+\delta}{r+\delta}\right) - \frac{\lambda(\mathbf{\Omega})m(\theta(\mathbf{\Omega}))}{r+\delta}p[(1-\pi)\underline{y}+\pi\overline{y}] + \mathbf{\Omega}pk \\
w_{LL}(\mathbf{\Omega}): &\quad J_{LL} = V \quad \to \quad w_{LL} = p\underline{y} \\
w_{LH}(\mathbf{\Omega}): &\quad J_L = V \quad \to \quad w_{LH} = p\underline{y} \\
w_{HH}(\mathbf{\Omega}): &\quad J_{HH} = V \quad \to \quad w_{HH} = p\overline{y}
\end{aligned}
$$

The objective of the vacancy posting firm (as in Lise and Robin (2017), we assume that these are handled by competing intermediaries; in contrast to their setup, our intermediaries operate in a CRS environment and have zero profits, meaning that one firm can post many vacancies) is to maximize the value of vacancies by choosing the measure of either low or high type vacancies. In particular, the values of opening $\underline{v}$ low and $\overline{v}$ high-type vacancies, $V_{\underline{v}}$ and $V_{\overline{v}}$, are given by

$$
\begin{aligned}
V_{\underline{v}} &= -c(\underline{v}) + \underline{v}q(\theta(\mathbf{\Omega}))\frac{u}{s}J_L \\
V_{\overline{v}} &= -c(\overline{v}) + \overline{v}q(\theta(\mathbf{\Omega}))\left[\frac{u}{s}J_H + \frac{\lambda(\mathbf{\Omega_L})\gamma_L}{s}J_{LH}\right]
\end{aligned}
$$

The FOCs indicate that the marginal cost of a vacancy is equal the value of a job of each type:

$$
\begin{aligned}
c'(\underline{v}) &= q(\theta(\mathbf{\Omega}))\frac{u}{s}J_L \equiv q(\theta(\mathbf{\Omega}))J_{\underline{v}} \\
c'(\overline{v}) &= q(\theta(\mathbf{\Omega}))\left[\frac{u}{s}J_H + \frac{\lambda(\mathbf{\Omega_L})\gamma_L}{s}J_{LH}\right] \equiv q(\theta(\mathbf{\Omega}))J_{\overline{v}}
\end{aligned}
$$

In addition, with CRS, profits from opening any vacancy are zero or equivalently, the equilibrium value of opening either vacancy is zero, $q(\theta(\mathbf{\Omega}))vJ_v - c(v) = 0$, where J_v is the value of a low or high vacancy.

Assumption on Vacancy Costs. Let $c(v) = c_0vy$, $y \in \{\underline{y}, \overline{y}\}$. Then the first order conditions for both types of vacancies are given by:

$$
\begin{aligned}
q(\theta(\mathbf{\Omega}))\frac{u}{s}J_L &= c_0\underline{y} \\
q(\theta(\mathbf{\Omega}))\left[\frac{u}{s}J_H + \frac{\lambda(\mathbf{\Omega})\gamma_L}{s}J_{LH}\right] &= c_0\overline{y}
\end{aligned}
$$

Recall that we assume CRS, which is why both types of firms make on average zero profits (i.e. $V_{\underline{v}} = V_{\overline{v}} = 0$), so that firms are indifferent between posting low and high type vacancies. From the FOCs, $\frac{u}{s}J_L\overline{y} = \left[\frac{u}{s}J_H + \frac{\lambda(\mathbf{\Omega})\gamma_L}{s}J_{LH}\right]\underline{y}$. We can then solve this equation for $m(\theta)$ (and, assuming the telegraph matching function, $m(\theta) = \frac{\phi\alpha\theta}{\alpha\theta+1}$, also for θ) as a function of π:

$$
\begin{aligned}
m(\theta(\mathbf{\Omega})) &= -\frac{1}{2\lambda^2(\overline{y}-\underline{y})(b+(\pi-1)(\overline{y}-\underline{y}))} \\
&\times\left[\sqrt{\lambda^2(\overline{y}-\underline{y})^2\left((b(2\delta+r)+\delta k\Omega-(\pi-1)\underline{y}(\delta+r)+\delta\pi\overline{y}-\delta\overline{y}+k\Omega r)^2-4\delta(\delta+r)(b+k\Omega)(b+(\pi-1)(\overline{y}-\underline{y}))\right)}\right. \\
&\left.+\lambda(\overline{y}-\underline{y})(b(2\delta+r)+\delta k\Omega-(\pi-1)\underline{y}(\delta+r)+\delta\pi\overline{y}-\delta\overline{y}+k\Omega r)\right] \qquad (48)
\end{aligned}
$$

where θ can then immediately be computed from inverting the matching function. This condition pins down θ as a function of π.

Multiple Steady State Equilibria. We focus on multiplicity of two equilibria, one where workers always search actively, i.e. in *both* states γ_L and γ_H, with $\omega_L = \omega_H = 1$. And one where workers never search, i.e. $\omega_L = \omega_H = 0$. We thus need to verify *two* no-deviation conditions in two states γ_L and γ_H:

1. No deviation when no one ever searches: $E_L(0|\mathbf{0}) \geq E_L(1|\mathbf{0})$ and $E_H(0|\mathbf{0}) \geq E_H(1|\mathbf{0})$
2. No deviation when all workers in low productivity jobs always search: $E_L(1|\mathbf{1}) \geq E_L(0|\mathbf{1})$ and $E_H(1|\mathbf{1}) \geq E_H(0|\mathbf{1})$

The next proof, adapted from the proof of Lemma 1, shows that the condition for multiplicity is very similar to (but stronger than) the condition from the baseline model, i.e.,

$$\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1\left[\pi\overline{y} + (1-\pi)\underline{y} - b\right]}\right) < m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right) \leq \theta(\mathbf{1}).$$

Proof. **1.1. No deviation in γ_L jobs when no one searches: $E_L(0|\mathbf{0}) \geq E_L(1|\mathbf{0})$.**

In this case, when no one actively searches on-the-job ($\mathbf{\Omega} = \mathbf{0}$), a worker in a low productivity job deviating during an interval dt chooses $\omega = 1$ and gets a payoff

$$E_L(1|\mathbf{0}) = \frac{1}{1+rdt}\left[dt(w_L(\mathbf{0}) - pk) + (1-\delta dt)dt\lambda(1)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LL} + \pi E_{LH}\right] + (1-\delta dt)(1-dt\lambda(1)m(\theta(\mathbf{0})))E_L(0|\mathbf{0}\right.$$

where $E_{LL} = E_{LL}(0|\mathbf{0})$ and $E_{LH} = E_{LH}(0|\mathbf{0})$. There is no deviation provided $E_L(0|\mathbf{0}) \geq E_L(1|\mathbf{0})$ or:

$$\begin{aligned} E_L(0|\mathbf{0})(1+rdt) \quad &\geq \quad dt(w_L(\mathbf{0}) - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LL} + \pi E_{LH}\right] \\ &\quad +(1-\delta dt - dt\lambda(1)m(\theta(\mathbf{0})) + dt^2\delta\lambda(1)m(\theta(\mathbf{0})))E_L(0|\mathbf{0}) + \delta dtU. \end{aligned}$$

After subtracting $E_L(0|\mathbf{0})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_L(0|\mathbf{0}) \geq w_L(\mathbf{0}) - pk + \lambda(1)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LL} + \pi E_{LH}\right] + (-\delta - \lambda(1)m(\theta(\mathbf{0})))E_L(0|\mathbf{0}) + \delta U.$$

Substituting the equilibrium values for $E_L(0|\mathbf{0}), E_{LL}, E_{LH}, U$ and $w_L(\mathbf{0})$ we get:

$$(\underline{y} - b)[\lambda(\mathbf{1}) - \lambda(\mathbf{0})]m(\theta(\mathbf{0})) - k(r+\delta) \leq 0.$$

So there is no deviation provided that:

$$\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b))}\right)$$

1.2. No deviation in γ_H jobs when no one searches: $E_H(0|\mathbf{0}) \geq E_H(1|\mathbf{0})$.

$$\begin{aligned} E_H(1|\mathbf{0}) = \qquad &\frac{1}{1+rdt}\left[dt(w_H(\mathbf{0}) - pk) + (1-\delta dt)dt\lambda(1)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right]\right. \\ &\left.+(1-\delta dt)(1-dt\lambda(1)m(\theta(\mathbf{0})))E_H(0|\mathbf{0}) + \delta dtU\right] \end{aligned}$$

where $E_{LL} = E_{LL}(0|\mathbf{0})$ and $E_{LH} = E_{LH}(0|\mathbf{0})$. There is no deviation provided $E_H(0|\mathbf{0}) \geq E_H(1|\mathbf{0})$ or:

$$\begin{aligned} E_H(0|\mathbf{0})(1+rdt) \quad &\geq \quad dt(w_H(\mathbf{0}) - pk) + dt\lambda(1)(1-\delta dt)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right] \\ &\quad +(1-\delta dt - dt\lambda(1)m(\theta(\mathbf{0})) + dt^2\delta\lambda(1)m(\theta(\mathbf{0})))E_H(0|\mathbf{0}) + \delta dtU. \end{aligned}$$

After subtracting $E_H(0|\mathbf{0})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_H(0|\mathbf{0}) \geq w_H(\mathbf{0}) - pk + \lambda(1)m(\theta(\mathbf{0}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right] + (-\delta - \lambda(1)m(\theta(\mathbf{0})))E_H(0|\mathbf{0}) + \delta U.$$

Substituting the equilibrium values for $E_H(0|\mathbf{0}), E_{LH}, E_{HH}, U$ and $w_H(\mathbf{0})$ we get:

$$(\underline{y} - b)[\lambda(1) - \lambda(\mathbf{0})]m(\theta(\mathbf{0})) - k(r+\delta) + (\lambda(1) - \lambda(\mathbf{0}))m(\theta(\mathbf{0}))\left[\pi(\overline{y} - \underline{y})\right] \leq 0$$

This condition is stronger than the one under 1.1. (that one is implied by this condition) since $(\lambda(1) - \lambda(\mathbf{0}))m(\theta(\mathbf{0}))\left[\pi(\overline{y} - \underline{y})\right] > 0$. Therefore, the requirement for multiplicity is (using the fact that $\lambda(1) - \lambda(\mathbf{0}) = \lambda_1$:

$$(\underline{y} - b)\lambda_1 m(\theta(\mathbf{0})) - k(r+\delta) + \lambda_1 m(\theta(\mathbf{0}))\left[\pi(\overline{y} - \underline{y})\right] \leq 0$$

or

$$\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1\left[\pi\overline{y} + (1-\pi)\underline{y} - b\right]}\right).$$

2.1. No deviation in γ_L job when all workers in low productivity jobs search: $E_L(1|\mathbf{1}) \geq E_L(0|\mathbf{1})$. In this case, when all workers in low productivity jobs actively search on-the-job ($\mathbf{\Omega} = \mathbf{1}$), a worker in a low productivity job deviating for an interval dt chooses $\omega = 0$ and gets a payoff

$$E_L(0|\mathbf{1}) = \frac{1}{1+rdt}\left[dtw_L(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LL} + \pi E_{HH}\right] + (1-\delta dt)(1 - dt\lambda(0)m(\theta(\mathbf{1})))E_L(1|\mathbf{1}) + \delta dtU\right].$$

There is no deviation provided $E_L(1|\mathbf{1}) \geq E_L(0|\mathbf{1})$:

$$\begin{aligned} E_L(1|\mathbf{1})(1+rdt) \quad &\geq \quad dtw_L(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LL} + \pi E_{HH}\right] \\ &\quad + (1 - \delta dt - dt\lambda(0)m(\theta(\mathbf{1})) + dt^2\delta\lambda(0)m(\theta(\mathbf{1})))E_L(1|\mathbf{1}) + \delta dtU. \end{aligned}$$

After subtracting $E_L(1|\mathbf{1})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_L(1|\mathbf{1}) \geq w_L(\mathbf{1}) + \lambda(0)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LL} + \pi E_{HH}\right] + (-\delta - \lambda(0)m(\theta(\mathbf{1})))E_L(1|\mathbf{1}) + \delta U.$$

Substituting the equilibrium values for $E_L(1|\mathbf{1}), E_{LL}, E_{HH}, U$ and $w_L(\mathbf{1})$ we get:

$$(\underline{y} - b)[\lambda(1) - \lambda(0)]m(\theta(\mathbf{1})) - k(r+\delta) \geq 0.$$

So there is no deviation provided that:

$$\theta(\mathbf{1}) \geq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y} - b))}\right)$$

2.2. No deviation in γ_H job when all workers in low productivity jobs search: $E_H(1|\mathbf{1}) \geq E_H(0|\mathbf{1})$. In this case, when all workers in high productivity jobs actively search on-the-job ($\mathbf{\Omega} = \mathbf{1}$), a worker in a high productivity job deviating for an interval dt chooses $\omega = 0$ and gets a payoff

$$E_H(0|\mathbf{1}) = \frac{1}{1+rdt}\left[dtw_H(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right] + (1-\delta dt)(1 - dt\lambda(0)m(\theta(\mathbf{1})))E_H(1|\mathbf{1}) + \delta dtU\right].$$

There is no deviation provided $E_H(1|\mathbf{1}) \geq E_H(0|\mathbf{1})$:

$$\begin{aligned} E_H(1|\mathbf{1})(1+rdt) \quad &\geq \quad dt w_H(\mathbf{1}) + dt\lambda(0)(1-\delta dt)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right] \\ &\quad +(1-\delta dt - dt\lambda(0)m(\theta(\mathbf{1})) + dt^2\delta\lambda(0)m(\theta(\mathbf{1})))E_H(1|\mathbf{1}) + \delta dt U. \end{aligned}$$

After subtracting $E_H(1|\mathbf{1})$ from both sides and dividing by dt and taking the limit $dt \to 0$, we obtain:

$$rE_H(1|\mathbf{1}) \geq w_H(\mathbf{1}) + \lambda(0)m(\theta(\mathbf{1}))\left[(1-\pi)E_{LH} + \pi E_{HH}\right] + (-\delta - \lambda(0)m(\theta(\mathbf{1})))E_H(1|\mathbf{1}) + \delta U.$$

Substituting the equilibrium values for $E_H(1|\mathbf{1}), E_{LH}, E_{HH}, U$ and $w_H(\mathbf{1})$ we get:

$$(\underline{y} - b)[\lambda(1) - \lambda(0)]m(\theta(\mathbf{1})) - m(\theta(\mathbf{1}))[\lambda(1) - \lambda(0)]\pi(\overline{y} - \underline{y}) - k(r+\delta) \geq 0.$$

And therefore, there is no deviation in this particular point of the tree if:

$$\theta(\mathbf{1}) \geq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1\left[\pi\overline{y} + (1-\pi)\underline{y} - b\right]}\right).$$

Since $\lambda_1(\pi\overline{y} + (1-\pi)\underline{y} - b) > \lambda_1(\underline{y} - b)$ this condition is less strict than the condition under 2.1. As a result, the conditions for no deviation when all workers in high productivity jobs search is:

$$\theta(\mathbf{1}) \geq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1\left[\underline{y} - b\right]}\right).$$

The necessary and sufficient conditions for the existence of multiple steady state equilibria is therefore:

$$\theta(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1\left[\pi\overline{y} + (1-\pi)\underline{y} - b\right]}\right) < m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y} - b))}\right) \leq \theta(\mathbf{1}) \tag{49}$$

□

Notice that this condition is *more stringent* than the one from our baseline model. This is intuitive since workers who obtain the high-productivity match right after unemployment have strong incentives to keep searching in order to obtain another $\overline{y}$ match and extract all rents from matching. Only a *sufficiently* low market tightness prevents them from always wanting to do so.

Notice that these bounds still depend on the endogenous variable π. To obtain the bounds in terms of π that only depend on parameters, we evaluate (49) at equality, using the expression for $\theta(\mathbf{\Omega})$ from

(48) and solve for π. We obtain the following expressions for π:

$$
\begin{aligned}
\pi_h(\mathbf{0}) &= \frac{1}{2\lambda_1(\delta+r)(\overline{y}-\underline{y})(\delta(\overline{y}-\underline{y})(b\lambda_1+k\lambda_0)-k\lambda_0 r\underline{y})}\times \\
&\quad \{[k^2\lambda_0^2(\delta+r)^2\left(\delta^2(\overline{y}-\underline{y})^2(\lambda_1(\overline{y}-b)+k\lambda_0)^2+2\delta r(\overline{y}-\underline{y})(\lambda_1(\overline{y}-b)+k\lambda_0)(\lambda_1(b\overline{y}+2b\underline{y}-\overline{y}\underline{y})+k\lambda_0(\overline{y}-\underline{y}))\right. \\
&\quad \left.+r^2\left(2k\lambda_0\lambda_1(\overline{y}-\underline{y})(b\overline{y}+2b\underline{y}-\overline{y}\underline{y})+\lambda_1^2(b\overline{y}-2b\underline{y}+\overline{y}\underline{y})^2+k^2\lambda_0^2(\overline{y}-\underline{y})^2\right)\right)]^{\frac{1}{2}} \\
&\quad +\delta^2(-(\overline{y}-\underline{y}))\left(k\lambda_0\lambda_1(b-\overline{y}+2\underline{y})+2b\lambda_1^2(\underline{y}-b)+k^2\lambda_0^2\right)+\delta r\left(k\lambda_0\lambda_1\left(b(\underline{y}-2\overline{y})+(\overline{y}-2\underline{y})^2\right)\right. \\
&\quad \left.+2b\lambda_1^2(b-\underline{y})(\overline{y}-\underline{y})+2k^2\lambda_0^2(\underline{y}-\overline{y})\right)-k\lambda_0\lambda_1 r^2(b\overline{y}+\underline{y}(\overline{y}-2\underline{y}))+k^2\lambda_0^2 r^2(\underline{y}-\overline{y})\} \\
\pi_l(\mathbf{1}) &= \frac{1}{k(\lambda_0+\lambda_1)(\delta(\overline{y}-\underline{y})(\lambda_1(\underline{y}-b)+k(\lambda_0+\lambda_1))+r(\lambda_1\underline{y}(b-\underline{y})+k(\lambda_0+\lambda_1)(\overline{y}-\underline{y})))}\times \\
&\quad \{b^3(-\delta)\lambda_1^2+b^2\lambda_1(\delta k(2\lambda_0+\lambda_1)+2\delta\lambda_1\underline{y}+kr(\lambda_0+\lambda_1)) \\
&\quad +b\left(-k(\lambda_0+\lambda_1)(k\lambda_0(\delta+r)+\delta\lambda_1\overline{y})+\delta k\lambda_1\underline{y}(\lambda_1-\lambda_0)-\delta\lambda_1^2\underline{y}^2\right) \\
&\quad +k\left(\delta(k(\lambda_0+\lambda_1)+\lambda_1\underline{y})(\overline{y}(\lambda_0+\lambda_1)-\underline{y}(\lambda_0+2\lambda_1))+r(\lambda_0+\lambda_1)\left(k\overline{y}(\lambda_0+\lambda_1)-k\underline{y}(\lambda_0+2\lambda_1)-\lambda_1\underline{y}^2\right)\right)\}
\end{aligned}
$$

where $\pi_h(\mathbf{0})$ gives the highest value of π that is consistent with an equilibrium where no one searches actively and $\pi_l(\mathbf{1})$ is the lowest π that is consistent with an equilibrium where everyone searches actively.

Finally, to obtain the bounds in terms of primitive p, we use the zero profit conditions (either $V_{\overline{v}}=0$ or $V_{\underline{v}}=0$) for both the equilibrium of active and non-active search and solve for p_l and p_h respectively

$$
\begin{aligned}
p_l &= \frac{c\underline{y}}{\frac{m(\theta(\mathbf{1}))u(\underline{y}-b)}{\theta(\mathbf{1})(\delta+r)(\gamma(\lambda_0+\lambda_1)+u)}-\frac{km(\theta(\mathbf{1}))u}{\theta(\mathbf{1})(\gamma(\lambda_0+\lambda_1)+u)(\delta+m(\theta(\mathbf{1}))(\lambda_0+\lambda_1)+r)}+\frac{m(\theta(\mathbf{1}))^2(1-\pi_l(\mathbf{1}))u(\lambda_0+\lambda_1)(\overline{y}-\underline{y})}{\theta(\mathbf{1})(\delta+r)(\gamma(\lambda_0+\lambda_1)+u)(\delta+m(\theta(\mathbf{1}))(\lambda_0+\lambda_1)+r)}} \\
p_h &= -\frac{c\theta(\mathbf{0})\underline{y}(\delta+r)(\gamma\lambda_0+u)(\delta+\lambda_0 m(\theta(\mathbf{0}))+r)}{m(\theta(\mathbf{0}))u(b\delta+b\lambda_0 m(\theta(\mathbf{0}))+br-\delta\underline{y}+\lambda_0 m(\theta(\mathbf{0}))\pi_h(\mathbf{0})\overline{y}-\lambda_0 m(\theta(\mathbf{0}))\pi_h(\mathbf{0})\underline{y}-\lambda_0 m(\theta(\mathbf{0}))\overline{y}-r\underline{y})}
\end{aligned}
$$

which we evaluate at π_l and π_h as well as $\theta(\mathbf{\Omega})$ (from (48)) and $u, \gamma=\gamma_L+\gamma_H$ from the flow-balance equations to obtain expressions that solely depend on parameters. The expressions are involved but one can show via simulations that there exists a parameter range for which $p_h > p_l$ and $\pi_l(\mathbf{1}) > \pi_h(\mathbf{0})$. We have the following result (exact expressions available upon request).

Proposition 4. *Let $m(\theta)=\phi\frac{\alpha\theta}{\alpha\theta+1}$. Then there are multiple steady states if and only if $p\in[p_l,p_h]$. The set $[p_l,p_h]$ is non-empty for an open set of parameters.*

Also in this set-up with *ex-ante firm heterogeneity*, the strategic complementarity and thus multiplicity survives. When both equilibria coexist, the active OJS equilibrium is characterized by more search effort, larger market tightness and a larger share of high-productivity vacancies $\overline{v}$ (with $\overline{y}$). What gives rise to this strategic complementarity between OJS and (high-type) vacancies? Here workers are incentivized to search actively, not only if tightness is high enough (as before) but also if the fraction of high productivity vacancies is high enough. High type vacancies encourage OJS because it offers workers the chance to obtain a job where they extract the entire surplus (not only after unemployment but, crucially, also after OJS). In turn, firms are encouraged to not only post more vacancies but especially more high type vacancies in the presence of active OJS because OJS biases the pool of searcher towards the employed. This bias implies that high type vacancies *match faster* ($\underline{v}$ cannot attract on-the-job searchers) and the *match duration* with employed searchers is longer than with unemployed ones (due to the restriction to a finite number of search rounds).

4. Dynamic Model With Anticipated Expectation Shocks

In this section, we develop a version of our baseline model in which the agents' expectations about the aggregate OJS follow a two-state Markov switching process. We use this augmented model to show that response to a positive expectation shock is similar to the response to such a shock in the baseline model with *un*anticipated shocks.

In the optimistic state, $\mathbf{\Omega}_t = \mathbf{1}$, agents expect the economy to converge to the steady state with active OJS (conditional on remaining in the optimistic state). In the pessimistic state, $\mathbf{\Omega}_t = \mathbf{0}$, agents expect the economy to converge to the steady state with passive OJS (conditional on remaining in the pessimistic state). But agents take into account that their expectations switch from optimistic to pessimistic at the Poisson rate $\pi_{\mathbf{10}}$, in which case the value of a worker to a firm in a low-productivity job jumps by $\Delta J_{\mathbf{10}t} = \underline{J}_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{0})) - \underline{J}_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{1}))$. In turn, at rate $1 - \pi_{\mathbf{10}}$, the agents' beliefs in any instance remain optimistic. Similarly, the agents' expectations switch from pessimistic to optimistic at the Poisson rate $\pi_{\mathbf{01}}$, in which case the value of a worker to a firm in a low-productivity job jumps by $\Delta J_{\mathbf{01}t} = \underline{J}_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{1})) - \underline{J}_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{0}))$. In turn, at rate $1 - \pi_{\mathbf{01}}$, the agents' beliefs in any instance remain pessimistic.

In the optimistic state, the economy evolves according to the following system of equations

$$\begin{aligned}
U_t(\mathbf{\Omega}_t) &= \frac{pb}{r} \\
\underline{E}_t(\mathbf{\Omega}_t) &= \frac{\underline{w}_t(\mathbf{\Omega}_t) - \mathbf{\Omega}_t pk + \lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t)) \overline{E}_t(\mathbf{\Omega}_t)}{r + \lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t))} \\
\overline{E}_t(\mathbf{\Omega}_t) &= \frac{\overline{w}_t(\mathbf{\Omega}_t) + \delta \frac{pb}{r}}{r + \delta} \\
\overline{w}_t(\mathbf{\Omega}_t) &= p\underline{y} \\
\underline{w}_t(\mathbf{\Omega}_t) &= pb \left(\frac{r + \lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t)) + \delta}{r + \delta} \right) - \frac{\lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t))}{r + \delta} p\underline{y} + \mathbf{\Omega}_t pk \\
r\underline{J}_t(\mathbf{\Omega}_t) &= p\underline{y} - \underline{w}_t(\mathbf{\Omega}_t) + [\lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t)) + \delta] \underline{J}_t(\mathbf{\Omega}_t) + \pi_{\mathbf{10}} \Delta J_{\mathbf{10}t} + \dot{\underline{J}}_t(\mathbf{\Omega}_t) \qquad (50) \\
\overline{J}_t(\mathbf{\Omega}_t) &= \frac{p\overline{y} - \overline{w}_t(\mathbf{\Omega}_t)}{r + \delta} \\
0 &= -c + q(\theta_t(\mathbf{\Omega}_t)) \left[\frac{u_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} J_t(\mathbf{\Omega}_t) + \frac{\lambda(\mathbf{\Omega}_t)\gamma_t}{u_t + \lambda(\mathbf{\Omega}_t)\gamma_t} \overline{J}_t(\mathbf{\Omega}_t) \right] \\
\dot{u}_t &= \delta(1 - u_t) - u_t m(\theta_t(\mathbf{\Omega}_t)) \\
\dot{\gamma}_t &= u_t m(\theta_t(\mathbf{\Omega}_t)) - (\delta + \lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t)))\gamma_t
\end{aligned}$$

where we set $\mathbf{\Omega}_t = \mathbf{1}$. Note that value (50) is the only piece that changes compared to the baseline model. This is the case since like in the baseline model, the other values $\underline{E}, \overline{E}, \overline{J}$ are time-invariant under the sequential auctions bargaining.

In the pessimistic state, the economy evolves according to an almost identical system of (differential) equations, except that now $\mathbf{\Omega}_t = \mathbf{0}$ and equation (50) becomes,

$$r\underline{J}_t(\mathbf{\Omega}_t) = p\underline{y} - \underline{w}_t(\mathbf{\Omega}_t) + [\lambda(\mathbf{\Omega}_t) m(\theta_t(\mathbf{\Omega}_t)) + \delta] \underline{J}_t(\mathbf{\Omega}_t) + \pi_{\mathbf{01}} \Delta J_{\mathbf{01}t} + \dot{\underline{J}}_t(\mathbf{\Omega}_t) \qquad (51)$$

reflecting that at rate $\pi_{\mathbf{01}}$ the agents' beliefs become optimistic in which case there is a change in value of a filled low productivity job of size $\Delta J_{\mathbf{01}t}$.

Proceeding similarly to the baseline model, we can reduce these systems of equations to a system of three differential equations, one for the optimistic state (with $\mathbf{\Omega}_t = \mathbf{1}$),

$$\dot{u}_t(\mathbf{\Omega}_t) = \delta(1 - u_t(\mathbf{\Omega}_t)) - u_t(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)) \tag{52}$$

$$\dot{\gamma}_t(\mathbf{\Omega}_t) = u_t(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)) - (\delta + \lambda(\mathbf{\Omega}_t)m(\theta_t(\mathbf{\Omega}_t)))\gamma_t(\mathbf{\Omega}_t) \tag{53}$$

$$\begin{aligned}\dot{\underline{\theta}}_t(\mathbf{\Omega}_t) = & \frac{\frac{bp\left(\frac{\alpha\phi\theta_t\lambda(\mathbf{\Omega}_t)}{\alpha\theta_t+1}+\delta+r\right)-\frac{\alpha p\phi\theta_t\underline{y}\lambda(\mathbf{\Omega}_t)}{\alpha\theta_t+1}+kp(\delta+r)}{\delta+r} + \left(\frac{\alpha\phi\theta_t\lambda(\mathbf{\Omega}_t)}{\alpha\theta_t+1} + \delta + \pi_{\mathbf{10}} + r\right)\left(\frac{c(\alpha\theta_t+1)(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi u_t} - \frac{\gamma_t p\lambda(\mathbf{\Omega}_t)(\overline{y}-\underline{y})}{u_t(\delta+r)}\right)}{\frac{c(\alpha\theta_t+1)(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t} - \frac{c(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t}} \\ & - \frac{\frac{c(\alpha\theta_t+1)\lambda(\mathbf{\Omega}_t)\left(u_t\left(\frac{\alpha\phi\theta_t u_t}{\alpha\theta_t+1}-\gamma_t\left(\frac{\alpha\phi\theta_t\lambda(\mathbf{\Omega}_t)}{\alpha\theta_t+1}+\delta\right)\right)-\gamma_t\left(\delta(1-u_t)-\frac{\alpha\phi\theta_t u_t}{\alpha\theta_t+1}\right)\right)}{\alpha\phi u_t^2}}{\frac{c(\alpha\theta_t+1)(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t} - \frac{c(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t}} \\ & + \frac{\frac{p\lambda(\mathbf{\Omega}_t)(\overline{y}-\underline{y})\left(u_t\left(\frac{\alpha\phi\theta_t u_t}{\alpha\theta_t+1}-\gamma_t\left(\frac{\alpha\phi\theta_t\lambda(\mathbf{\Omega}_t)}{\alpha\theta_t+1}+\delta\right)\right)-\gamma_t\left(\delta(1-u_t)-\frac{\alpha\phi\theta_t u_t}{\alpha\theta_t+1}\right)\right)}{u_t^2(\delta+r)} - \Delta J_{\mathbf{10}t}\pi_{\mathbf{10}} - p\underline{y}}{\frac{c(\alpha\theta_t+1)(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t} - \frac{c(\gamma_t\lambda(\mathbf{\Omega}_t)+u_t)}{\alpha\phi\theta_t u_t}}\end{aligned} \tag{54}$$

and one for the pessimistic state (with $\mathbf{\Omega}_t = \mathbf{0}$),

$$\dot{u}_t(\mathbf{\Omega_t}) = \delta(1 - u_t(\mathbf{\Omega_t})) - u_t(\mathbf{\Omega_t})m(\theta_t(\mathbf{\Omega_t})) \tag{55}$$

$$\dot{\gamma}_t(\mathbf{\Omega_t}) = u_t(\mathbf{\Omega_t})m(\theta_t(\mathbf{\Omega}_t)) - (\delta + \lambda(\mathbf{\Omega_t})m(\theta_t(\mathbf{\Omega}_t)))\gamma_t(\mathbf{\Omega_t}) \tag{56}$$

$$\begin{aligned}\dot{\underline{\theta}}_t(\mathbf{\Omega}_t) = & \frac{\alpha p\phi u_t^2(b(\alpha\theta_t(\delta + \lambda(\mathbf{\Omega}_t)\phi + r) + \delta + r) - \alpha\lambda(\mathbf{\Omega}_t)\phi\theta_t\underline{y}) + cu_t(\alpha\theta_t + 1)(\delta + r)\left(\delta(\alpha\theta_t + 1)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t - 1) + \alpha\phi\theta_t\left(\gamma_t\lambda(\mathbf{\Omega}_t)^2\right.\right.}{\alpha cu_t(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t)} \\ & + \frac{u_t(\alpha\theta_t(\delta + \lambda(\mathbf{\Omega}_t)\phi + \pi_{\mathbf{01}} + r) + \delta + \pi_{\mathbf{01}} + r)(c(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t) + \alpha\gamma_t\lambda(\mathbf{\Omega}_t)p\phi(\underline{y} - \overline{y}))}{\alpha cu_t(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t)} \\ & + \frac{c(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t)(-\delta(u_t - 1)(\alpha\theta_t + 1) - \alpha\phi\theta_t u_t) - \alpha\lambda(\mathbf{\Omega}_t)p\phi u_t(\overline{y} - \underline{y})(\delta(\alpha\gamma_t\theta_t + \gamma_t) + \alpha\phi\theta_t(\gamma_t\lambda(\mathbf{\Omega}_t) - u_t))}{\alpha cu_t(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t)} \\ & - \frac{\alpha\gamma_t\lambda(\mathbf{\Omega}_t)p\phi(\overline{y} - \underline{y})(-\delta(u_t - 1)(\alpha\theta_t + 1) - \alpha\phi\theta_t u_t) - \alpha\Delta J_{\mathbf{01}t}\phi\pi_{\mathbf{01}}u_t^2(\alpha\theta_t + 1)(\delta + r) - \alpha p\phi u_t^2\underline{y}(\alpha\theta_t + 1)(\delta + r)}{\alpha cu_t(\alpha\theta_t + 1)(\delta + r)(\gamma_t\lambda(\mathbf{\Omega}_t) + u_t)}.\end{aligned} \tag{57}$$

where $\Delta J_{\mathbf{10}t} = \underline{J}(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{0})) - \underline{J}(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{1}))$ and $\Delta J_{\mathbf{01}t} = \underline{J}(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{1})) - \underline{J}(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{0}))$ can be computed from the firm's free entry condition. We have,

$$\begin{aligned}\Delta J_{\mathbf{01}t} =& \underline{J}_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{1})) - \underline{J}_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}), \theta_t(\mathbf{0})) \\ =& \left[\frac{c}{q(\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{1}))}\frac{u_t(\mathbf{0}) + \lambda(\mathbf{1})\gamma_t(\mathbf{0})}{u_t(\mathbf{0})} - \frac{\lambda(\mathbf{1})}{u_t(\mathbf{0})}\overline{J}_t\right] \\ & - \left[\frac{c}{q(\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{0}))}\frac{u_t(\mathbf{0}) + \lambda(\mathbf{0})\gamma_t(\mathbf{0})}{u_t(\mathbf{0})} - \frac{\lambda(\mathbf{0})}{u_t(\mathbf{0})}\overline{J}_t\right]\end{aligned} \tag{58}$$

where $\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{0}))$ indicates the market tightness when agents are in the pessimistic state whereas $\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{1})$ indicates the market tightness when agents have *just* switched from the pessimistic to the optimistic state but where the state variables $u_t(\mathbf{0}), \gamma_t(\mathbf{0})$ from the pessimistic state are still in place.

Similarly, we have,

$$\begin{aligned}\Delta J_{\mathbf{10}t} =& \underline{J}_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{0})) - \underline{J}_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}), \theta_t(\mathbf{1})) \\ =& \left[\frac{c}{q(\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{0}))}\frac{u_t(\mathbf{1}) + \lambda(\mathbf{0})\gamma_t(\mathbf{1})}{u_t(\mathbf{1})} - \frac{\lambda(\mathbf{0})}{u_t(\mathbf{1})}\overline{J}_t\right] \\ & - \left[\frac{c}{q(\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{1}))}\frac{u_t(\mathbf{1}) + \lambda(\mathbf{1})\gamma_t(\mathbf{1})}{u_t(\mathbf{1})} - \frac{\lambda(\mathbf{1})}{u_t(\mathbf{1})}\overline{J}_t\right]\end{aligned} \tag{59}$$

where $\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{1}))$ indicates the market tightness when agents are in the optimistic state whereas

$(\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{0}))$ indicates the market tightness when agents have *just* switched from the optimistic to the pessimistic state but where the state variables $u_t(\mathbf{1}), \gamma_t(\mathbf{1})$ from the optimistic state are still in place.

We will focus on a *rational expectations Markov switching equilibrium*, which is defined as follows: For an initial belief $\mathbf{\Omega}_0 = \{\mathbf{0}, \mathbf{1}\}$, the path $\{U_t, \underline{E}_t, \overline{E}_t, V_t, \underline{J}_t, \overline{J}_t, \theta_t, u_t, \gamma_t, \xi_t, \underline{w}_t, \overline{w}_t, \omega_t, \Delta J_{\mathbf{01}t}, \Delta J_{\mathbf{10}t}\}_{t\geq 0}$ is such that for all $t \in [0, \infty)$:

1. Given $\mathbf{\Omega}_t = \mathbf{0}$, system (55)- (57) holds where $\Delta J_{\mathbf{01}t}$ is given by (59). Given $\mathbf{\Omega}_t = \mathbf{1}$, system (52)-(54) holds where $\Delta J_{\mathbf{01}t}$ is given by (58).
2. Given $\mathbf{\Omega}_t = \mathbf{0}$, $\omega_t = 0$ maximizes $\underline{E}_t(\mathbf{0})$. And given $\mathbf{\Omega}_t = \mathbf{1}$, $\omega_t = 1$ maximizes $\underline{E}_t(\mathbf{1})$.
3. After a belief switch – the value of a low productivity job lands on the stable manifold associated with the new steady state. Since we focus on a system of differential equations where we replaced $\underline{J}_t$ by θ_t, this means that tightness θ_t must jump onto the stable manifold associated with the new steady state: First, when the economy switches from the optimistic to the pessimistic state, market tightness $\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{0})$ must be on the stable manifold associated with the steady state of passive job search. Thus, if the economy remains in the pessimistic state forever after, it will converge to the steady state of passive job search. Second, when the economy switches from the pessimistic to the optimistic state, market tightness $\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{1})$ must be on the stable manifold associated with the steady state of active job search. Thus, if the economy remains in the optimistic state forever, it will converge to the steady state of active job search.
4. The initial market tightness must be on the stable manifold associated with active job search steady state if the initial state of the economy is optimistic, and it must on the stable manifold of the passive search steady state if the initial state of the economy is pessimistic.

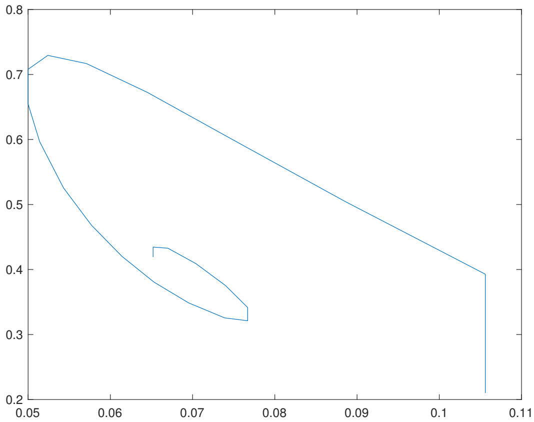
Requirements 1. and 2. are of the same nature as requirements 1.-5. in Definition 1, only that we collapsed the system of equilibrium equation to 3 differential equations in both states. Requirement 2. in both definitions are identical and satisfied if there is no one-shot deviation from the current aggregate search strategy $\mathbf{\Omega}_t$. Lemma 2 applies here since the expectation shocks in the firms' value function (and hence in the differential equation for θ_t) affect workers' values only indirectly through time-varying tightness but not directly. This again stems from the sequential auction bargaining, under which the workers' values $\underline{E}, \overline{E}$ and U are all time invariant and not directly impacted by expectation shocks. The following condition, which has to hold for all $t \in [0, \infty)$ thus guarantees that there are no profitable deviations from either path (i.e. for any $\mathbf{\Omega}_t = \{\mathbf{0}, \mathbf{1}\}$)

$$\theta_t(\mathbf{0}) \leq m^{-1}\left(\frac{k(r+\delta)}{\lambda_1(\underline{y}-b)}\right) \leq \theta_t(\mathbf{1}).$$

Requirement 4. replaces transversality condition 6. from Definition 1 and states that we need to consider initial conditions that, depending on the workers' beliefs, puts the economy on one of the two stable manifolds. From then on, by the property of stable manifolds and how we specified the switching process, the economy will be contained in the two manifolds forever, meaning θ_t is finite.

Satisfying requirement 3. is complicated since the stable manifolds in our numerical solution have no perfect overlap in the (u, γ)-space. Also, since we are dealing with two state variables u_t and γ_t, characterizing ΔJ as a function of u_t, γ_t is in general complicated. We therefore proceed using the following approximation: We fix $\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{0})$ to be the *average* tightness on the stable manifold associated with the steady state of passive search. And similarly, we fix $\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{1})$ to be the *average* tightness on the stable manifold associated with the steady state of active search. Given $\theta_t(u_t(\mathbf{1}), \gamma_t(\mathbf{1}); \mathbf{0})$ and $\theta_t(u_t(\mathbf{0}), \gamma_t(\mathbf{0}); \mathbf{1})$, we can compute $\Delta J_{\mathbf{01}}$ and $\Delta J_{\mathbf{10}}$.

Here is how we proceed in practice to construct this equilibrium. We first guess that in this setting with anticipated expectation shocks, the two steady states under consideration admit the same stability



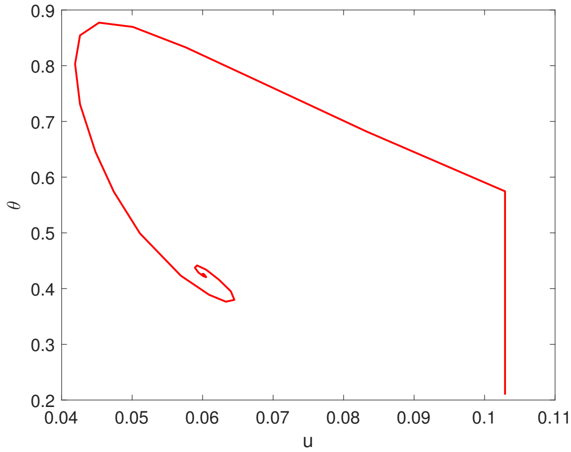


Figure 15: A. Beveridge Curve in $\theta - u$ space (Model with Anticipated Expectation Shocks). B Beveridge Curve in $\theta - u$ space (Model with Unanticipated Expectation Shocks)

properties as in our baseline model with unanticipated expectation shocks (something which we will need to verify). We then use the approximation described above to compute the jump variables $\Delta J_{\mathbf{01}}$ and $\Delta J_{\mathbf{10}}$. We use the same calibration as in the baseline model and set $\pi_{\mathbf{01}}$ and $\pi_{\mathbf{10}}$ such that we match the average duration of boom (optimistic state) and recession (pessimistic state) in the data. We then compute the stable manifolds based on systems (55)- (57) and (52)- (54). Last, we fix our starting value at the *recession steady state* in which agents have pessimistic beliefs, $(u_t^*(\mathbf{0}), \gamma_t^*(\mathbf{0}), \theta_t^*(\mathbf{0}))$. Given that this steady state is covered by the stable manifold under optimistic beliefs in the (u, γ)-space, a belief switch from pessimistic to optimistic will bring the economy onto the stable manifold of the boom. Last, just as in the baseline model, we track the economy during its transition path during recovery and towards the steady state with optimistic beliefs.

Bearing in mind that we have used an approximation for the jump variables ΔJ, we find that the dynamics of this equilibrium with anticipated expectation shocks are remarkably similar to those from our baseline model. For brevity, we do not repeat all the exercises here, but simply show the shift of the Beveridge Curve during the economy's recovery (i.e. after switching from pessimistic to optimistic) in Figure 15a, which is very similar to Figure 15b that was constructed using our baseline model with unanticipated expectation shocks and which, for the ease of comparison, we report here again.

Online Appendix II. Firm Deviation to Back-Loaded Wage Contract

One concern of our analysis is that the assumption of fixed wages drives the multiplicity result. While the fixed wage assumption is common in this literature, it is well-known that it is not necessarily the optimal contract. A firm may find it optimal to offer *time-varying wages* to discourage workers' OJS in the equilibrium with active OJS. Here we make a modest attempt to address this issue. We extend the contract space to a two-part wage with back-loading, and ask whether firms would want to deviate and offer a wage different from the constant wage. In particular, we allow a firm to deviate from the current contract with constant wages and to post a relatively low wage for T periods (where T is optimally chosen by the firm) which incentivizes OJS, followed by a relatively high wage from $T+1$ onward that discourages active search. We find that for the relevant parameter values, a firm is worse off when deviating and posting the time-varying wages compared to the equilibrium contract with stable wages. In this case, the value of a filled low productivity job under the deviating contract approaches the equilibrium value of a filled low productivity job in the limit for $T \to \infty$ and is strictly below for finite T, implying that the firm would *not* want to deviate from fixed wages and pay a higher wages to discourage search. Only if discounting is (unnaturally) high, such that workers do not value much the benefits of search, and if at the same time search costs are high, then it is profitable for the firm to deviate from the equilibrium wage contract because discouraging search is cheap. Of course, this result does not prove that there exists no profitable firm deviation through some more complicated contract. But it does demonstrate that allowing for a natural class of wage contracts does not induce profitable deviations by the firms that would destroy the equilibrium with OJS (under reasonable parameter restrictions).

Here, we sketch the analysis if we allow firms to commit to the above mentioned wage contract, given that other firms offer fixed wages. We do not aim to provide a full analytical characterization at this stage but rather to give the conceptual framework and the intuition for the results. For convenience, we do the analysis in discrete time. There is only one stage in which this deviating contract may be profitable to the firm, and that is when employing a worker in a low productivity job (this is the only stage at which the worker searches). We therefore focus on a deviation by a single firm regarding the contract in the low productivity job in the equilibrium with active OJS .

Denote by $\underline{E}_T$ the value of a low productivity job to a worker, in which he will receive the low wage $\underline{w}_1$ for T periods and the high wage $\underline{w}_2$ from $T+1$ onward. This implies that $\underline{E}_T$ is the value of a job from the perspective of an unemployed worker. Denote by $\underline{E}_0$ the value of a job to a worker from period $T+1$ onward. We adopt the same notation for the firm, i.e. $\underline{J}_T$ is the value of a filled low productivity job to a firm when paying the worker a low wage for T periods; $\underline{J}_0$ is the value of that job when starting to pay the worker a higher wage from $T+1$ on.

In steady state, these values (we spell out this model in discrete time) are given by

$$\begin{aligned}
\underline{J}_T &= \frac{1-(\beta(1-\delta)(1-\lambda(\mathbf{1})m))^T}{\beta(1-\delta)(1-\lambda(\mathbf{1})m)}(p\underline{y}-\underline{w}_1)+(\beta(1-\delta)(1-\lambda(\mathbf{1})m))^T\underline{J}_0 \\
\underline{J}_0 &= \frac{p\underline{y}-\underline{w}_2}{\beta(1-\delta)(1-\lambda(\mathbf{0})m)} \\
\underline{E}_T &= \frac{1-(\beta(1-\delta)(1-\lambda(1)m))^T}{\beta(1-\delta)(1-\lambda(1)m)}(\underline{w}_1-pk)+(\beta(1-\delta)(1-\lambda(1)m))^T\underline{E}_0+\beta\lambda(1)m\frac{1-(\beta(1-\delta)(1-\lambda(1)m))^T}{\beta(1-\delta)(1-\lambda(1)m)}\overline{E} \\
&\quad +\delta\beta\frac{1-(\beta(1-\delta)(1-\lambda(1)m))^T}{\beta(1-\delta)(1-\lambda(1)m)}U \\
\underline{E}_0 &= \frac{w_2+\beta\lambda(0)m\overline{E}+\beta\delta U}{1-\beta(1-\delta)(1-\lambda(0)m)} \\
\overline{E} &= \frac{\overline{w}+\beta\delta U}{1-\beta(1-\delta)} \\
U &= \frac{b}{1-\beta}
\end{aligned}$$

which take into account that $V=0$ due to free entry.

The firm's objective is to choose a triple $(T, \underline{w}_1, \underline{w}_2)$ to maximize the value of a low productivity job $\underline{J}_T$ subject to three constraints:

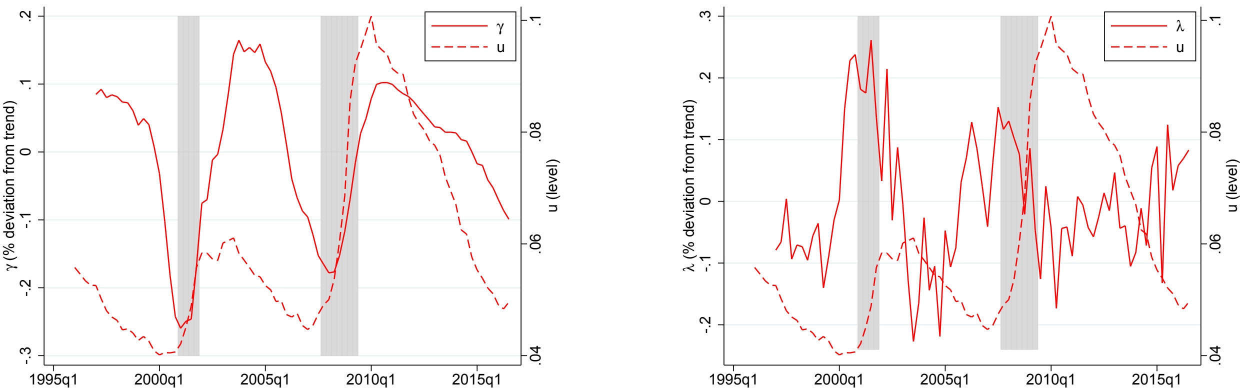
$$\begin{aligned}
& \max_{T,\underline{w}_1,\underline{w}_2} \underline{J}_T \\
s.t. \quad & \underline{E}_T \geq U \\
& \underline{E}_0(0) \geq E_0(1) \\
& \underline{E}_{T-1}(1) \geq E_{T-1}(0)
\end{aligned}$$

The first constraint states that the wages must be such that the worker is at least as well off taking the job as in unemployment; the second states that after T periods, wages must be such that the worker weakly prefers not to search; the third constraint ensures that the worker does not want to deviate from the strategy 'search until period T (and not thereafter)': If the worker prefers to search in period $T-1$, he also prefers to search in all periods $t < T-1$ since in period $T-1$ it is most tempting to not search due to the soon-to-be-expected wage increase. The first two constraints will bind, otherwise the firm would forgo profits. We recover $\underline{w}_2$ from constraint 2 and, given $\underline{w}_2$, we recover $\underline{w}_1$ from constraint 1. Last, we verify that constraint 3 holds for any parameter constellation; it is slack.

We then evaluate the objective function $\underline{J}_T$ at the wages and check its properties: Our simulations (available upon request) reveal that it is either monotonic increasing or decreasing, i.e. T^* is at a corner. For most parameter ranges, $\underline{J}_T$ is increasing, always weakly below the value of the on-the-job search equilibrium $\underline{J}$ with $\lim_{T\to\infty} \underline{J}_T = \underline{J}$. This implies that the deviation is *not* profitable; firms do not seek to discourage search by backloading wages. For some parameter constellations (in particular, for unnaturally low β and high k), $\underline{J}_T$ can be decreasing with its maximum at $T^* = 0$. In this case, in which OJS is costly and workers do not value much future benefits, it is worth it for the firm to discourage search, and the firm would do so immediately after hiring.

Online Appendix III. Additional Empirical Results

We repeat the derivation of γ_t (based on (17)) and of λ_t from the CPS data under two variations, different from those in the main text. First, instead of using the data on separations, we impose a constant separation rate δ, which is consistent with the model (Figure 16). Second, we construct the separation rate that adjust for time-aggregation as in Shimer (2012) (Figure 17).



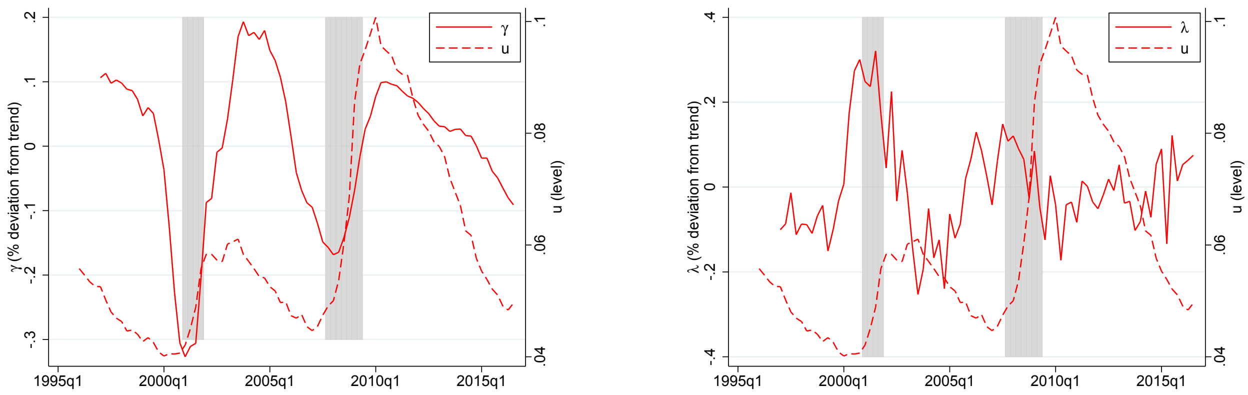
(a) γ

(b) λ

Figure 16: Active On-The-Job Searchers and Search Intensity (Constant δ).

The results are similar, except for a level difference. To see why we obtain similar results, consider the EU rates under the different specifications in Figure 18. The time-series of the different EU rates are very similar only that when adjusting for time-aggregation as in Shimer (2012) the EU rate has a higher level (denoted by 'eu time aggregated'). For comparison, we also plot in that figure the EU rate from the BLS (denoted by 'eu BLS') as well as the one from Fallick and Fleischman (2004) (denoted by 'eu ff 2004'), which are both very similar to the one we used here.

In Figure 19 we report the equivalent of Figure 9, expressed in terms of the *percentage point* deviation instead of the percentage deviation from trend.



(a) γ (b) λ

Figure 17: Active On-The-Job Searchers and Search Intensity (δ adjusting for time aggregation).

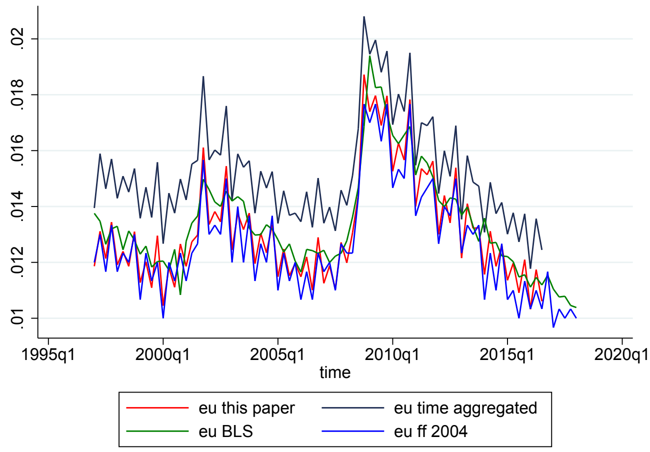
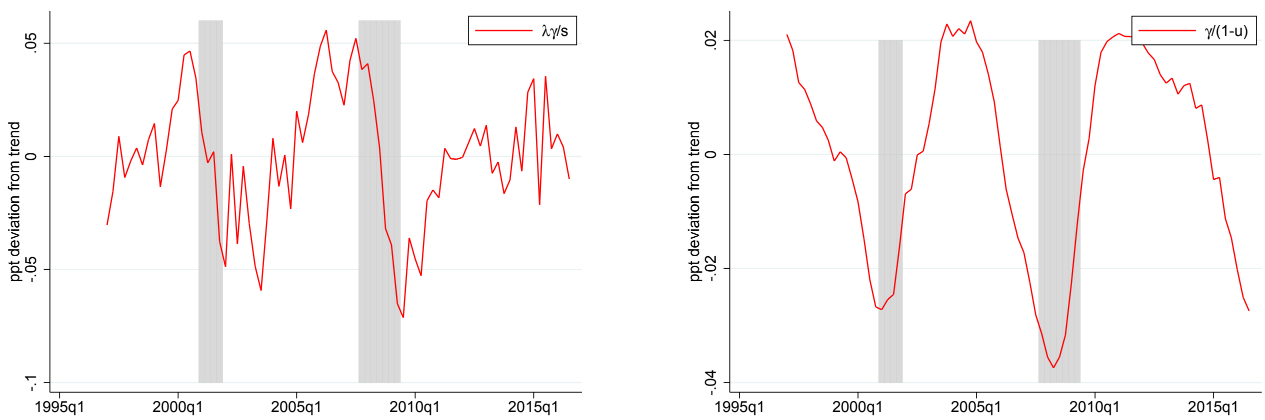


Figure 18: Various Separation (EU) Rates



(a) Composition of Searchers (b) Composition of Employed Workers

Figure 19: Composition of Searchers and Employed Workers (Percentage Point Deviation from Trend)