DISCUSSION PAPER SERIES

Nonlinearities and Workers' Heterogeneity in Unemployment Dynamics

Stéphane Adjemian
Frédéric Karamé
Francois Langot

DISCUSSION PAPER SERIES

Nonlinearities and Workers' Heterogeneity in Unemployment Dynamics

Stéphane Adjemian
Le Mans University and CEPREMAP

Frédéric Karamé
Le Mans University and CEPREMAP

Francois Langot
Le Mans University, CEPREMAP, Paris School of Economics and Institut Universitaire de France and IZA

ABSTRACT

Nonlinearities and Workers' Heterogeneity in Unemployment Dynamics*

This study demonstrates that nonlinearities, coupled with worker heterogeneity, make it possible to reconcile the Diamond–Mortensen–Pissarides model with the labor market dynamics observed in the United States. Nonlinearities, induced by firings and downward real wage rigidities, magnify adjustments in quantities, whereas heterogeneity concentrates them on the low-paid workers' submarkets. The model fits the job finding, job separation, and unemployment rates well. It also explains the Beveridge curve's dynamics and the cyclicality of the involuntary component of separations. The estimated dynamics of the aggregate shock that allows generating the US labor market fluctuations has a correlation with unemployment that changes of sign during the 80s. We also show that the differences in adjustment between submarkets predicted by the model are consistent with the data of job flows by educational attainment.

JEL Classification:	C51, E24, E32
Keywords:	search and matching, unemployment dynamics, nonlinearities, particle filter, maximum likelihood estimation

Corresponding author:
François Langot
GAINS, Le Mans Université
Avenue Olivier Messiaen
72085 Le Mans Cedex 9
France
E-mail: flangot@univ-lemans.fr

1 Introduction

This study shows that workers' heterogeneity and nonlinearities are essential for the search and matching (SaM) model to explain the dynamics of job finding, job separation, and unemployment rates in the United States.

In our extension of Diamond (1982)–Mortensen (1982)–Pissarides (1985) model (DMP hereinafter), the economy is a collection of segmented markets, indexed by worker ability, which operates independently of each other and shares only a common aggregate shock. These independent markets are then aggregated to create an economy. To more convincingly model separations and, particularly, generate their involuntary component, we introduce a downward real wage rigidity that could be interpreted as a social norm (see Hall (2005)) and flexible firing costs. With worker heterogeneity, the *fundamental surplus* (Ljungqvist & Sargent (2017)) that determines the labor market elasticity is specific to each submarket. Indeed, the gap between current productivity and the maximum between the opportunity cost of employment and the lower bound of the bargained wages set is specific to each market. For high-ability workers, this gap is large, allowing for large wage adjustments and, thus, the low volatility of quantities. Conversely, this gap is small for low-ability workers, especially as the lower bound of the bargained wages set is higher than the opportunity cost of employment, leading to large quantity adjustments. Indeed, when the wage is at its lowest level, firms react only by adjusting quantities. The responses of the aggregates to macroeconomic shocks depend on the aggregation of these heterogeneous behaviors. We solve our model by considering the occasionally binding constraints and endogenous switching between different regimes of decision rules, which are, at each time, specific to each labor market segment. The four different regimes are as follows: *(i)* it is sometimes optimal for firms to post vacancies (the interior solution of the DMP model); however, in recessions, it may be optimal for them not to post vacancies and *(ii)* not to fire, *(iii)* to fire a fraction of workers, or *(iv)* to close and thus fire all workers.

As the model can generate a large spectrum of elasticities, each specific to one unobserved ability, the model's estimation is a crucial part of the quantitative analysis. Specifically, this analysis must determine the composition of our artificial population (i.e., the "good" weight for each ability) to ensure that the model matches the aggregate worker flows and unemployment. We follow Fernandez-Villaverde et al. (2015) and Herbst & Schorfheide (2015) by estimating the structural parameters of the model using maximum likelihood (ML). The likelihood is approximated using particle filtering coupled with projection methods to solve a model with occasionally binding constraints (Judd (1992) and Christiano & Fischer (2000)). Hence, our quantitative evaluation of the model takes advantage of the nonlinear solution of the model and thus uses information on United States data asymmetries.[1] The

---

[1]The nonlinearities in employment adjustments and asymmetries in the job creation and destruction rates have been discussed in previous works. See Neftci (1984), Sichel (1993), Burgess (1992), Acemoglu & Scott (1994), Caballero & Hammour (1994), and Davis et al. (1998). Keynes (1936) and Burns & Mitchell (1946) discuss the existence of asymmetries in cyclical movements of key macroeconomic aggregates. They point out that recessions are more severe than expansions. Nevertheless, McKay & Reis (2008) and Ferraro (2017)

study's results are based on the most advanced methods in nonlinear estimation to evaluate a SaM model with heterogeneity.

Conditional on our estimate of worker heterogeneity, we find that the model fits well the observed aggregate job finding, job separation, and unemployment rates, which are the data used to estimate the model. This very good fit is obtained by solving the Shimer (2005)'s puzzle. First, whereas Shimer (2005) shows that the size of productivity shocks is too small to match the volatility of unemployment, our results show that it is possible to restrict the model's parameters to match the persistence and variance of the US labor productivity. Moreover, we show that our aggregate shock is positively correlated with US labor productivity before the end of the 1980s and becomes negatively correlated in more recent periods. This finding echoes Barnichon (2010) empirical evidences. Second, our results are obtained under the restriction that the opportunity cost of employment must be in the range of the empirical evidence provided by Hall & Milgrom (2008), and thus significantly lower than the values used by Hagendorn & Manovskii (2008). We show that these empirical successes are based on the nonlinear dynamics of the model: the aggregation of nonlinear individual problems characterized by non-smooth policy functions leads to significant asymmetries in the impulse response functions of aggregates vis-à-vis macroeconomic shocks. Indeed, separations even smoothed by a firing cost function remain a decision not affected by exchange externalities. This is not the case for hiring, for which the competition between firms during recovery generates congestion effects that dampen the intensity of the recovery. These asymmetries allow the model to reproduce the high speed of separations at the beginning of a recession and thus the low speed of adjustments in recovery. We show that these adjustments depend on the size of the recession and the speed of recovery.

Additionally, we show that our model can reproduce the adjustments around the Beveridge curve, thus solving the *Beveridge curve puzzle* (Fujita & Ramey (2012)), that is, the counterfactual result that the DMP model with an endogenous separation rate generates a positive slope of the Beveridge curve. Our model can also reproduce the joint dynamics of unemployment and vacancies before and after an unemployment peak (i.e., a recession). For each NBER recessions, we show that the slopes up to the unemployment peak are steep because they result from abrupt layoffs; in contrast, after the unemployment peak (i.e., during the recovery), the gradient is low because hiring is a time-consuming process. We show that the differences between recessions come from *(i)* the unemployment level before the recession, *(ii)* the size and persistence of the exogenous shock. These results are obtained with a model that can explain a large part of the cyclicality of the involuntary component in the separations (i.e., the layoffs observed in the data).

All these results are based on the fact that submarkets for low-ability workers are more sensitive to the business cycle than the other labor market segments in the model where worker heterogeneity is unobservable. As workers' heterogeneity is assumed to be invariant, it can be linked to educational attainment. We show that our model is consistent with the observed decreasing sensitivity to the business cycle of job finding, separation, and unem-

---

show that only (un)employment adjustments exhibit asymmetries.

3

ployment rates with educational attainment. Therefore, model heterogeneity is consistent with differences across educational attainment in workers' flows.

Our study is related to a large body of literature starting with the Shimer (2005)'s paper. First, we reinforce the view that real wage rigidity is an important ingredient for improving the DMP model fit, as shown previously in Hall (2005) and Hall & Milgrom (2008). We emphasize that it is sufficient to introduce a *downward* real rigidity to improve the model's implications significantly. Second, we also promote the idea that considering worker heterogeneity helps the DMP model to match aggregate dynamics substantially, thus reinforcing the results of Lise & Robin (2017), Ferraro (2017), and Chassamboulli (2013). Finally, our results also favor approaches that consider the rich nonlinearities of the DMP model, as previously discussed in Collard et al. (2002), Hairault et al. (2010), Petrosky-Nadeau et al. (2018), (2020) , and (2017).

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 applies the model to the data. Section 4 discusses the additional implications of the model, such as involuntary separations and the Beveridge curve. Section 5 shows that the estimated heterogeneity that allows the model to fit the aggregate data also allows it to explain disaggregate data, such as workers' flows by diploma. Finally, Section 6 concludes the paper.

2 Matching model with worker heterogeneity

As in Robin (2011), Lise & Robin (2017), and Ferraro (2017), the labor market is segmented by worker ability $\mu \in [1, M]$. To maintain the tractability of the model, there is no skill mobility. ω_μ denotes the mass of the labor force in the μ-type labor market segment, with $\sum_\mu \omega_\mu = 1$. In each labor market segment, the size of the population is normalized to unity. The output per unit of labor is denoted by $a_t y_\mu$, where $y_\mu \in \{y_1, ..., y_M\}$ is the skill-specific productivity for μ-type workers, and a_t is the common aggregate shock. Here, $\log(a_t)$ follows an $AR(1)$ process $\log(a_t) = \rho \log(a_{t-1}) + (1 - \rho) \log(\overline{a}) + \epsilon_t$ with $|\rho| < 1$ and $\epsilon \rightsquigarrow \mathcal{N}(0, \sigma^2)$.

2.1 Matching technology and stock-flow dynamics

To hire workers, firms must open vacancies at a unit cost $\kappa > 0$. μ-type workers and the jobs directed to those workers meet pairwise at the Poisson rate $\mathcal{M}(u_t(\mu), v_t(\mu))$, where $\mathcal{M}(u_t(\mu), v_t(\mu))$ represents the flows of matches, $v_t(\mu)$ is the number of vacancies, and $u_t(\mu)$ is the number of unemployed μ-type workers. Following Den Haan et al. (2000), the matching function is $\mathcal{M}(u_t(\mu), v_t(\mu)) = \varphi \frac{v_t(\mu) u_t(\mu)}{(v_t(\mu)^\nu + u_t(\mu)^\nu)^{1/\nu}}$, with $\nu > 0$, ensuring that the induced probabilities are always between zero and one. If $v_t(\mu) > 0$ in the μ-type market, the job finding ($p(\theta_t(\mu))$) and job filling rates ($q(\theta_t(\mu))$) depend on the ratio of vacancies to unemployed workers ($\theta_t(\mu) = v_t(\mu)/u_t(\mu)$) and are defined as $p(\theta_t(\mu)) = \varphi \frac{\theta_t(\mu)}{(\theta_t(\mu)^\nu + 1)^{1/\nu}}$ and $q(\theta_t(\mu)) = \varphi \frac{1}{(\theta_t(\mu)^\nu + 1)^{1/\nu}}$, respectively. The unemployment dynamics arise from entries to and exits from employment. The former are determined by job findings $p(\theta_t(\mu)) u_t(\mu)$ when the

aggregate shock provides a sufficient incentive to post vacancies, whereas the latter are given by the sum of exogenous separations $s(1-u_t(\mu))$ and endogenous separations $l_t(\mu)(1-u_t(\mu))$, where $l_t(\mu) \in [0, 1-s]$ denotes the endogenous separation rate (see Equation (1)). Hence, we obtain

$$u_{t+1}(\mu) \;\; = \;\; (s + l_t(\mu))(1 - u_t(\mu)) + (1 - p(\theta_t(\mu)))u_t(\mu), \tag{1}$$

implying that $u_{t+1}(\mu) = 1$ if $v_t(\mu) = 0$ and $l_t(\mu) = 1 - s$. In Section 2.4, we discuss why endogenous separations must be restricted to layoffs.

2.2 Worker's utility

We define $U_t(\mu)$ and $W_t(\mu)$ as the state-contingent present values of an unemployed and employed worker, respectively. $z(\mu)$ is the flow value for unemployed workers, who can earn unemployment benefits and home production, which are partially indexed to worker ability μ, and $w_t(\mu)$ is the wage. As discussed below, we assume that the Nash bargaining process is constrained to ensure that voluntary quits are excluded, leading to $W_t(\mu) \geq U_t(\mu)$. The value functions are (the indexes for individuals are omitted because each of them is representative)

$$\begin{aligned} U_t(\mu) &\;\; = \;\; z(\mu) + \beta\mathbb{E}_t[(1 - p(\theta_t(\mu)))U_{t+1}(\mu) + p(\theta_t(\mu))W_{t+1}(\mu)] \\ W_t(\mu) &\;\; = \;\; w_t(\mu) + \beta\mathbb{E}_t\left[(1 - s - l_t(\mu))W_{t+1}(\mu) + (s + l_t(\mu))U_{t+1}(\mu)\right], \end{aligned}$$

implying that $W_t(\mu) = w_t(\mu) + \beta\mathbb{E}_t[U_{t+1}(\mu)]$ if $v_t(\mu) = 0$ and $l_t(\mu) = 1 - s$. When the aggregate shock leads firms to choose $v_t(\mu) = 0$, then $U_t(\mu) = z(\mu) + \beta\mathbb{E}_t[U_{t+1}(\mu)]$ because $p(\theta_t(\mu)) = 0$. This changes the threat point in the bargaining process because employees' alternative options do not depend on labor market tightness.

Moreover, for each μ-type worker, the participation constraint $(W_t(\mu) \geq U_t(\mu))$ must be respected $\forall t$. The Nash bargaining process ensures that this constraint is satisfied. However, this is only the case for μ-type workers with a productivity $y_\mu a_t$ larger than the opportunity cost of employment $z(\mu)$. When the Nash bargaining process is constrained by the lower limit of the wage bargaining set, we must check whether the participation constraint is satisfied. If this is not the case, firms exit the market and lack workers. When $W_t(\mu) < U_t(\mu)$, the market is closed; when this occurs, all workers quit their jobs and, thus, all firms close.

2.3 Firm's behavior

On each submarket, μ, a continuum of firms selling their goods in a perfectly competitive market exists. On each of these submarkets, firms can find the workforce that corresponds to their technology.[2] To account for the reorganization costs and institutional arrangements, such as the experience rating system and temporary layoff agreements, each firm supports

---

[2]The indexes for individual firms are omitted because each of them is representative. Firms can be interpreted as positions that use a homogeneous μ-type of labor. They can also be viewed as establishments where a particular activity of a company is carried out if one activity is associated with one type of labor.

firing costs, which are an increasing function of the number of fired workers, denoted by $F(l_t(\mu)n_t(\mu))$, with $F' > 0$ and $F'' > 0$. Therefore, the value function is nonlinear in employment. In this case, the firm knows that the wage can depend on its firing policy because the marginal profit can depend on the employment level through the endogenous firing costs. Hence, in the general case, we obtain $w_t(\mu) = w_\mu(n_t(\mu), a_t)$, as in Stole et al. (1996) and Cahuc & Wasmer (2001) (see also Bertola & Caballero (1994) and Bertola & Garibaldi (2001)). We define $\lambda_t(\mu)$, $\nu_t(\mu)$, and $\phi_t(\mu)$ as the Lagrange multipliers of each inequality constraint, and the firm's program is

$$
\mathcal{V}_t(\mu) = \max_{v_t(\mu), l_t(\mu), n_{t+1}(\mu)} \left\{ (y_\mu a_t - w_\mu(n_t(\mu), a_t)) n_t(\mu) - \kappa v_t(\mu) - F(l_t(\mu)n_t(\mu)) + \beta \mathbb{E}_t \left[\mathcal{V}_{t+1}(\mu) \right] \right\}
$$

$$
\text{s.c.} \quad
\begin{cases}
n_{t+1}(\mu) = (1 - s - l_t(\mu))n_t(\mu) + q(\theta_t(\mu))v_t(\mu) & \\
v_t(\mu) \geq 0 & (\lambda_t(\mu)) \\
l_t(\mu) \geq 0 & (\nu_t(\mu)) \\
l_t(\mu) \leq (1 - s) & (\phi_t(\mu)).
\end{cases}
$$

By using the notation $J_t(\mu) \equiv \frac{\partial \mathcal{V}_t(\mu)}{\partial n_t(\mu)}$, we find that the optimal behavior is given by

$$
\begin{aligned}
0 &= -\kappa + q(\theta_t(\mu))\beta \mathbb{E}_t \left[J_{t+1}(\mu) \right] + \lambda_t(\mu)q(\theta_t(\mu)) \\
0 &= -F'(l_t(\mu)n_t(\mu)) - \beta \mathbb{E}_t \left[J_{t+1}(\mu) \right] + \nu_t(\mu) - \phi_t(\mu) \\
J_t(\mu) &= y_\mu a_t - w_t(\mu) - \frac{\partial w(n_t(\mu), a_t)}{\partial n_t(\mu)} - l_t(\mu)F'(l_t(\mu)n_t(\mu)) \\
&\quad + (1 - s - l_t(\mu))\beta \mathbb{E}_t \left[J_{t+1}(\mu) \right] + l_t(\mu)\nu_t(\mu) + (1 - s - l_t(\mu))\phi_t(\mu).
\end{aligned}
$$

Given that $v_t(\mu) \geq 0 \Leftrightarrow q(\theta_t(\mu))v_t(\mu) \geq 0$, the Kuhn–Tucker conditions are

$$
\begin{aligned}
0 &= \lambda_t(\mu)q(\theta_t(\mu))v_t(\mu) & (2) \\
0 &= \nu_t(\mu)l_t(\mu)n_t(\mu) & (3) \\
0 &= \phi_t(\mu)((1 - s - l_t(\mu))n_t(\mu). & (4)
\end{aligned}
$$

When $\nu_t(\mu) = 0$, that is, if there are firings, then $\beta \mathbb{E}_t \left[J_{t+1}(\mu) \right] < 0$. In other words, the expected job surplus of the firm can be negative. This outcome is possible only if $J_{t+1}(\mu) < 0$ for some realizations of the technological shock.

2.4 Wages

As is typical in the classical DMP model, we assume that the wage is the solution to the Nash bargaining problem between the firm and its marginal worker at each time point. Nevertheless, following Hall (2005), we introduce wage rigidity in the form of a social norm. The originality of our approach is that this rigidity occurs only when real wages decline; thus, it takes the form of a lower limit of the wage bargaining set, denoted $\underline{w}$. This downward real wage rigidity leads to $w_t(\mu) = \max\{w_t^{\text{Nash}}(\mu), \underline{w}\}$. It also leads us to exclude voluntary quits; for low-paid workers, the lower limit of the wage bargaining set ($\underline{w}$) ensures that

$W_t(\mu) \geq U_t(\mu)$. Hence, in the model, only two types of separations can occur: natural quits, which occur at an exogenous rate s, and layoffs, which occur at an endogenous rate $l_t(\mu)$. Hence, for workers paid at the lower bound of the bargaining set, the firm's marginal surplus may be negative and, in this case, firings or firm exits may occur. If firings occur, that is, if $0 < l_t(\mu) < 1 - s$, implying that $\phi_t(\mu) = \nu_t(\mu) = 0$, we obtain $F'(l_t(\mu)n_t(\mu)) = -\beta\mathbb{E}_t[J_{t+1}(\mu)]$, leading to

$$J_t(\mu) = y_\mu a_t - w_t(\mu) - \frac{\partial w(n_t(\mu), a_t)}{\partial n_t(\mu)} + (1-s)\beta\mathbb{E}_t[J_{t+1}(\mu)]. \tag{5}$$

However, when firings occur, it also follows that $\beta\mathbb{E}_t[J_{t+1}(\mu)] < 0$: in recession (low a_t), the expectations of recovery are insufficient to retain all the firm's workers. If, for a given level of a_t and some μ, we have a negative expected value ($\beta\mathbb{E}_t[J_{t+1}(\mu)] < 0$), then $J_t(\mu) < 0$. Given that the Nash solution would lead to voluntary quits because $W_t(\mu) - U_t(\mu) = \frac{\gamma}{1-\gamma}J_t(\mu) < 0$, the assumption of downward real wage rigidity implies that the wage is bounded by $\underline{w}$.[3]

If, in the same period and for some μ, we have $\beta\mathbb{E}_t[J_{t+1}(\mu)] > 0$ and $J_t(\mu) > 0$. Then, $l_t(\mu) = 0 \Rightarrow F(l_t(\mu)n_t(\mu)) = 0$, leading to the same marginal value of the job as derived previously (equation (5)). Nevertheless, in this regime, the firm has constant returns to scale, leading to $\frac{\partial w(n_t(\mu), a_t)}{\partial n_t(\mu)} = 0$, as in the classical DMP model. Given the assumption that $J_t(\mu) > 0$, we deduce that the Nash solution for the bargained wage $W_t(\mu) - U_t(\mu) = \frac{\gamma}{1-\gamma}J_t(\mu)$ implies that $w_t(\mu) > z(\mu)$ because $W_t(\mu) > U_t(\mu)$.

Property 1. *The presence of a lower limit of the real wage prevents strategic bargaining.*

Hence, the equilibrium wage is given by

$$w_t(\mu) = \begin{cases} \gamma(y_\mu a_t + \kappa\theta_t(\mu)) + (1-\gamma)z(\mu) & \text{If } J_t(\mu) > 0 \\ \underline{w} & \text{If } J_t(\mu) \leq 0. \end{cases} \tag{6}$$

When $J_t(\mu) > 0$, two cases are possible. First, in the regime in which $\beta\mathbb{E}_t[J_{t+1}(\mu)] > \kappa$, we obtain $\theta_t(\mu) > 0$, leading to the classical wage equation. Second, in the regime in which $0 < \beta\mathbb{E}_t[J_{t+1}(\mu)] \leq \kappa$, we obtain $\theta_t(\mu) = 0$, implying that labor market tightness is absent from the wage equation.

2.5 Equilibrium

By using (1), (2), (3), (4), and (6), we can deduce the equilibrium from

$$\frac{\kappa}{q(\theta_t(\mu))} - \lambda_t(\mu) = \beta\mathbb{E}_t \left\{ \begin{array}{l} y_\mu a_{t+1} - w_{t+1}(\mu) - l_{t+1}(\mu)F'(l_{t+1}(\mu)n_{t+1}(\mu)) \\ + (1 - s - l_{t+1}(\mu))\left(\frac{\kappa}{q(\theta_{t+1}(\mu))} - \lambda_{t+1}(\mu) + \phi_{t+1}(\mu)\right) \\ + \nu_{t+1}(\mu)l_{t+1}(\mu) \end{array} \right\}$$

---

[3] Even if $\beta\mathbb{E}_t[J_{t+1}(\mu)] > 0$, we can have $J_t(\mu) < 0$; in this case, the current profit is negative, but the expectations of recovery imply that firms do not fire workers ($l_t(\mu) = 0$ and thus $F(l_t(\mu)n_t(\mu)) = 0$). However, as previously shown, the solution for the wage is bounded by $\underline{w}$.

Property 2. *The dependence of the marginal cost of firings on the number of firings breaks the block recursivity of this search equilibrium model.*

As it is impossible to fire more workers than the firm's total workforce, the Lagrange multipliers of each constraint depend on each firm's state variable $(n_t(\mu))$. Given that firings and firm exits occur only in regimes where $J_t(\mu) \leq 0$, the block recursivity of the model is preserved for all μ such that $J_t(\mu) > 0$, leading to Property 1. Hence, the level of employment can matter for the expected marginal value of a job. This result breaks the block recursivity of the search equilibrium (see Pissarides (1990)).[4]

Property 3. *Each firm can be in one of the four regimes within the equilibrium distribution. With $\Theta_\mu(a_t, n_t(\mu)) = \frac{\kappa}{q(\theta_t(\mu))} - \lambda_t(\mu)$, these four regimes are*

- **Regime 1.** *When $\lambda_t(\mu) = 0 \Rightarrow v_t(\mu) > 0$ and $l_t(\mu) = 0 \Rightarrow \phi_t(\mu) = 0$, the first-order condition leads to*

$$
\left.
\begin{aligned}
\frac{\kappa}{\Theta_\mu(a_t, n_t(\mu))} &= q(\theta_t(\mu)) \\
\nu_t(\mu) &= \Theta_\mu(a_t, n_t(\mu)) \geq 0
\end{aligned}
\right\} \quad \text{because } \Theta_\mu(a_t, n_t(\mu)) > \kappa > 0
$$

$$
n_{t+1}(\mu) = (1-s)n_t(\mu) + q(\theta_t(\mu))v_t(\mu).
$$

The former equation can be an equilibrium restriction if and only if $q(\theta_t(\mu)) < 1$ because $q(\theta_t(\mu))$ is a probability. Otherwise, $\lambda_t(\mu) \geq 0 \Rightarrow v_t(\mu) = 0$, whereas $q(\theta_t(\mu))$ takes its extreme value, namely $q(\theta_t) = 1$. We deduce that the threshold value of $\Theta_\mu(a_t, n_t(\mu))$ is $\Theta_\mu(\tilde{a}_t(\mu), n_t(\mu)) = \kappa$.

- **Regime 2.** *If $0 < \Theta_\mu(a_t, n_t(\mu)) < \kappa$, $\lambda_t(\mu) \geq 0 \Rightarrow v_t(\mu) = 0$ and $\nu_t(\mu) \geq 0 \Rightarrow l_t(\mu) = 0$, whereas $\phi_t(\mu) = 0$. Hence, we deduce that*

$$
\begin{aligned}
\lambda_t(\mu) &= \kappa - \Theta_\mu(a_t, n_t(\mu)) \geq 0 \quad &&\text{because } \Theta_\mu(a_t, n_t(\mu)) < \kappa \\
\nu_t(\mu) &= \Theta_\mu(a_t, n_t(\mu)) \geq 0 \quad &&\text{because } \Theta_\mu(a_t, n_t(\mu)) > 0 \\
n_{t+1}(\mu) &= (1-s)n_t(\mu).
\end{aligned}
$$

The lower bound of $\Theta_\mu(a_t, n_t(\mu))$ in this regime is zero.

- **Regime 3.** *If $-F((1-s)n_t(\mu)) < \Theta_\mu(a_t, n_t(\mu)) < 0$, we still have $\lambda_t(\mu) \geq 0 \Rightarrow v_t(\mu) = 0$, but $0 < l_t(\mu) < (1-s)$, implying that $\nu_t(\mu) = 0$ and $\phi_t(\mu) = 0$. Hence, it follows that*

$$
\left.
\begin{aligned}
\lambda_t(\mu) &= \kappa - \Theta_\mu(a_t, n_t(\mu)) \geq 0 \\
F'(l_t(\mu)n_t(\mu)) &= -\Theta_\mu(a_t, n_t(\mu)) \geq 0
\end{aligned}
\right\} \quad \text{because } \Theta_\mu(a_t, n_t(\mu)) < 0
$$

$$
n_{t+1}(\mu) = (1 - s - l_t(\mu))n_t(\mu).
$$

---

[4]If the marginal cost of firings does not depend on the number of firings, then the block recursivity of this search equilibrium model is preserved. While this case simplifies the computation of the equilibrium, it reduces the endogenous separations to discrete choices (see Lise & Robin (2017)). It also breaks the orthogonality between labor market tightness and unemployment implied by the DMP model but is rejected by empirical evidence provided by Coles & Kelishomi (2018).

- **Regime 4.** *If $\Theta_\mu(a_t, n_t(\mu)) < -F((1-s)n_t(\mu))$, we still have $\lambda_t(\mu) \geq 0 \Rightarrow v_t(\mu) = 0$, but $l_t(\mu)n_t(\mu) = (1-s)n_t(\mu)$, implying that $\nu_t(\mu) = 0$ and $\phi_t(\mu) \geq 0$. This result leads to*

$$
\left.
\begin{aligned}
\lambda_t(\mu) &= \kappa - \Theta_\mu(a_t, n_t(\mu)) \geq 0 \\
\phi_t(\mu) &= -\Theta_\mu(a_t, n_t(\mu)) \geq 0
\end{aligned}
\right\} \qquad \text{because } \Theta_\mu(a_t, n_t(\mu)) < -F((1-s)n_t(\mu))
$$
$$
n_{t+1}(\mu) = 0.
$$

Although nonlinearities can be found *within* each regime, they are also induced by switches *between* regimes.[5]

Regime 1 is the usual case studied in DMP models. In this regime, the congestion effect already generates some significant nonlinearities, as shown by Hairault et al. (2010), Jung & Kuester (2011) or Iliopulos et al. (2019). Indeed, the low level of unemployment in booms slows hiring, whereas its high level in recessions makes it easier to adjust employment downward. Therefore, this property of the DMP model explains the slow adjustments to employment during recovery.

In **Regime 2**, firms are "inactive," as in Bertola & Garibaldi (2001), Bentolila & Bertola (1990) and Petrosky-Nadeau & Zhang (2017). Hence, the firm's optimal strategy is to hoard its workforce net of exogenous separations (no hiring and no firing). This labor hoarding regime can exist because market opportunities induce the firm to reduce employment, but the exogenous separation rate is sufficient to achieve the firm's objectives. This outcome can occur in periods of small recessions where rapid recovery is anticipated. Moreover, in this regime, the unemployment dynamic is linear because it decreases at a constant rate of exogenous separation. Thus, this regime is on a razor edge, given the low magnitude of exogenous separations.

The novelty of **Regime 3** vis-à-vis the destruction choices described by the model of Mortensen & Pissarides (1994) is that the intensity of the firing rate depends on the firms' employment levels. Indeed, from the second equation, we deduce that the layoff rate is $l_t(\mu) = F'^{-1}(-\Theta_\mu(a_t, n_t(\mu)))/n_t(\mu)$. If the recession is anticipated to be short and shallow, then the decision rules of this regime can be the firm's best option because the losses will be offset quickly.

Finally, **Regime 4** provides the date at which the firm exits the market. This is close to the regime described by Robin (2011), and Lise & Robin (2017) when the quit rate equals one, as the match surplus becomes negative. This regime can occur when a recession is sufficiently persistent and when jobs are highly fragile.

The occurrence and persistence of Regimes 3 and 4 highly depend on the firm's employment history; if employment is low, closing the firm (Regime 4) becomes the best option. Indeed, $-F((1-s)n_t(\mu))$ is close to zero when $n_t(\mu)$ is close to zero. Moreover, the choice

---

[5]For each regime, we must also check that the participation constraint is satisfied: *(i)* when the bargained wage is not constraint by $\underline{w}$, the participation constraint is $z(\mu) + \frac{\gamma}{1-\gamma}\kappa\theta_t(\mu) - (1-s)\frac{1}{1-\gamma}\frac{\kappa}{q(\theta_t(\mu))} < y_\mu a_t$, *(ii)* if the $\underline{w}$ is binding, the participation constraint is $\underline{w} > z(\mu)$. The figures of the decision rules for each regime are in Appendix D.

to fire some proportion of the workforce implies that $-F((1 - s)n_t(\mu))$ is the closest to zero; this result then reduces the distance between Regimes 2 and 4, and thus the probability that Regime 3 occurs. Regime 3 does not exist without firing costs, and firms control their workforce via the hiring process, exogenous separation rate, and closure decisions. In Regimes 3 and 4, the size of the recession (magnitude and persistence) and heterogeneous structural productivity of jobs implies that workers in more fragile jobs are fired; this result is the "cleansing effect of recessions" found by Saint-Paul (1993) and Caballero & Hammour (1994). This instantaneous adjustment of separations can explain the high sensitivity of the job separation rate at the beginning of a recession (see Figure 14 in Appendix A).

To analyze the "complete" dynamics of the DMP, we solve for the following equilibrium:

$$
n_{t+1}(\mu) = \begin{cases} (1 - s)n_t(\mu) + p\left(q^{-1}\left(\frac{\kappa}{\Theta_{t,\mu}}\right)\right)(1 - n_t(\mu)) & \text{If } \Theta_{t,\mu} > \kappa \\ (1 - s)n_t(\mu) & \text{If } 0 < \Theta_{t,\mu} \leq \kappa \\ (1 - s - l_t(\mu))n_t(\mu) & \text{If } -F'_{t,\mu} < \Theta_{t,\mu} \leq 0 \\ 0 & \text{If } \Theta_{t,\mu} \leq -F'_{t,\mu}, \end{cases}
$$

where we use the compact notations $\Theta_{t,\mu} = \Theta_\mu(a_t, n_t(\mu))$ and $F'_{t,\mu} = F'((1 - s)n_t(\mu))$. We compute $\Theta_{t,\mu}, \forall t, \mu$ using a collocation approach with Chebyshev polynomials of order three. In each market, we have two state variables: the efficiency of the production function and employment. Hence, we identify the policy rule by setting the Euler equation residuals to zero on a 2D grid with nine nodes, and we approximate the expectations using the unscented transform (see Appendix B for details on the numerical solution method).

3 Empirical strategy and results

If nonlinearities matter, the evaluation of our model requires a quantitative method that accounts for nonlinearities in the data. We develop a method based on the ML.[6] We use a particle filtering algorithm[7] coupled with a projection method to solve the model with occasionally binding constraints to extract information on the nonlinearities induced by our extension of the DMP model.

The main challenge is to provide an estimation explaining the large increases in the US unemployment rate and its slow drop during recovery. Only ten recessions have occurred since the end of WWII. To base our estimate on the largest number of these experiences, we favor an estimation using the largest possible sample from the beginning of 1951 to the end of 2018. Given the large turnover in the United States labor market, monthly data

---

[6]Using a Real Business Cycle model, Fernandez-Villaverde & Rubio-Ramirez (2005) show that a full information estimation of the nonlinear model (solved with finite elements) provides a better fit than a full information estimation of the linearized model. The authors also find that the parameter estimates are significantly different and translate into noticeable discrepancies between the simulated and estimated moments.

[7]These methods were first used by DeJong et al. (2000) and Schorfheide (2000) to estimate DSGE models. Appendices B and C describe the implemented method.

are also required. These various constraints prevent us from directly using productivity data, which are available over a shorter sample, and only at a quarterly frequency, which would complexify the estimation procedure by mixing data frequencies. These restrictions on the information set used in the estimation can also be seen as opportunities because our aggregate shock is not constrained to US productivity only.

Before discussing the estimation results, we first discuss the identification problem that we must solve. Next, we present the data used for the estimation and our assumptions regarding the functional forms.

3.1   Shock identification

It is well known since Shimer (2005) that the observed variance of labor productivity is too small to allow the DMP model to generate the observed unemployment volatility. Moreover, Barnichon (2010) has shown that the correlation between unemployment and productivity has changed of sign in the mid-1980s (from significantly negative, the correlation became significantly positive). These two points suggest that productivity shocks cannot explain all the labor market dynamics. A solution can be found in Ljungqvist & Sargent (2017) who show that the "fundamental surplus," the variable explaining the labor market dynamics, is driven by productivity shocks, but also by shocks on different costs, such as those linked to credit (Wasmer & Weil (2004)), to delay in bargaining (Hall & Milgrom (2008)), to hiring (Pissarides (2009)), to changes in unemployment benefits, to layoff taxes and to opportunity cost of employment induced by preference shocks (Balleer (2012)). Indeed, with a fundamental surplus defined as[8] $\Xi_t(\mu) = y_\mu \widetilde{a}_t - z(\mu) - \Upsilon_t(\mu)$, it appears that not only the "true" productivity shocks $\widetilde{a}_t$ can change the job value but also the exogenous dynamics of $\Upsilon_t(\mu)$ which are linked to the other shocks. However, the "fundamental surplus" does not allow one to easily identify the structural shocks, particularly the technological shocks. Indeed, if $\Upsilon_t(\mu) = y_\mu v_t$ for simplicity, then we have $\Xi_t(\mu) = y_\mu(\widetilde{a}_t - v_t) - z(\mu)$. Hence, the shock identified by our model is $a_t = \widetilde{a}_t - v_t$. The crucial point is that this unobservable shock a_t is not only linked to changes in labor productivity measured by $\frac{GDP_t}{N_t}$, as it should be the case for $\widetilde{a}_t$. That is the reason we exclude labor productivity from the information set of our estimation. To deal with Shimer's puzzle, understood as the link between the size of labor market fluctuations and the size of the shocks, we control the size of the shocks (a_t) such that the model can generate labor productivity having persistence and variance close to the ones measured in the historical data. With this method, the history of shocks a_t is not restricted to be the same as that of labor productivity. This allows us to reveal the necessary gaps between the identified shock (a_t) and historical productivity $(\widetilde{a}_t)$ that allow the model to match observed workers' flows and unemployment.

To implement this estimation strategy, the objective function is thus a penalized log-

---

[8]In our model, the employment surplus, defined as $S_t(\mu) = J_t(\mu) + W_t(\mu) - U_t(\mu) = y_\mu \widetilde{a}_t - z(\mu) - \max\{0, \frac{\gamma}{1-\gamma}\theta_t(\mu)\} - l_t(\mu)F'(l_t(\mu)n_t(\mu)) + \beta(1 - s - l_t(\mu))\mathbb{E}_t[S_{t+1}(\mu)]$ leads to a "fondamental surplus" $\Xi_t(\mu) \equiv y_\mu \widetilde{a}_t - z(\mu) - \max\{0, \frac{\gamma}{1-\gamma}\theta_t(\mu)\} - l_t(\mu)F'(l_t(\mu)n_t(\mu))$.

likelihood $\mathcal{L}^c = \mathcal{L} - \Gamma \left[(\widehat{\rho}_s - \widehat{\rho}_d)^2 + (10^4 \times (\widehat{\sigma}_s^2 - \widehat{\sigma}_d^2))^2 \right]$, where $\mathcal{L}$ is the model's log-likelihood, $\{\widehat{\rho}_s, \widehat{\sigma}_s\}$ and $\{\widehat{\rho}_d, \widehat{\sigma}_d\}$ are the estimated parameters of the AR(1) processes for the quarterly labor productivity (in log) for the simulated (s) and observed (d). Using the model, the simulated quarterly data for "output" and employment are respectively given by $Y_t^s = \sum_\mu \omega_\mu y_\mu \sum_{\tau=1}^{S} a_{t,\tau} N_{t,\tau}^s(\mu)$ and $N_t^s = \frac{1}{S} \sum_\mu \omega_\mu \sum_{\tau=1}^{S} N_{t,\tau}(\mu)$, where t denotes the quarter, τ the weeks in each quarter and S in the number of weeks in a quarter. This leads to simulated quarterly data for "labor productivity," defined as Y_t^s/N_t^s that we use to estimate $\{\widehat{\rho}_s, \widehat{\sigma}_s\}$. The parameter Γ is the arbitrary weight of the penalty in the objective criterion that we set to 2,000. We multiply the innovation variance by 10^4 in order that ρ and σ^2 have the same relative weight. This approach is close to the one proposed by Acharya et al. (2020), which employs both aggregate time series in the likelihood calculation of a Bayesian estimation and moment restrictions based on microdata. Here, our approach includes information on moments (persistence and variance) from the quarterly labor productivity, whereas the model is estimated on monthly data.

3.2 Data

We extend the unemployment rate (UR), job-finding rate (JFR), and job separation rate (JSR) data of Lise & Robin (2017) to the current period. To construct these data, we use seasonally adjusted data from the BLS covering 1951m1 to 2018m12 and measure employment, unemployment, and the number of individuals unemployed for more than five weeks for all people aged 16 and over (see Appendix A). To stationarize these data, we choose the same smoothing parameter as Lise & Robin (2017) ($\lambda_{HP} = 9^2 \times 2.5 \times 10^5$), which simplifies the comparisons.[9] The simulated data are weekly, and the induced monthly transition rates are calculated from the historical data provided by the BLS (see Appendix A).

3.3 Assumptions on the functional forms

The estimation is conducted under several restrictions. We normalize the scale parameter of the matching function (φ) and the mean of the aggregate productivity ($\bar{a}$) to one. We calibrate the discount factor β such that the real interest rate is 5% per year. For the estimation, we must choose functional forms for the opportunity cost of employment ($z(\mu)$) and firing costs ($F(ln)$). Following Robin (2011), we set the opportunity cost of employment $z(\mu)$ as a proportion ζ of worker ability y_μ as follows: $z(\mu) = z_0 + \zeta(y_\mu - z_0)$. Contrary to Lise & Robin (2017), we do not assume that the income of unemployed workers fluctuates with the business cycle. To test the ability of the model to solve the Shimer's puzzle without extreme values for the opportunity cost of employment, we restrict all solutions to satisfy $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.75$, value in the range of estimates provided by Hall & Milgrom (2008). This value is much lower than the calibrations chosen by Robin (2011) (close to

---

[9]As we use monthly data, we multiply the smoothing parameter of 2.5×10^5, which Lise & Robin (2017) use for quarterly data, by 9^2 (see Ravn & Uhlig (2002) for a discussion of adjustments to the HP filter smoothing parameter).

0.86), Christiano et al. (2016) (0.88), and the extreme calibration proposed by Hagendorn & Manovskii (2008) (close to 0.95). Given the assumption of the functional form for $z(\mu)$ and the distribution of y_μ, this restriction leads to $z_0 = \frac{0.75-\zeta}{1-\zeta}\left(y_1 + \frac{k}{k+\theta}\right)$. The firing cost function is $F(ln) = \frac{\phi}{1+\eta}(ln)^{1+\eta}$. Finally, as in Robin (2011), the weights of abilities in the economy are $\omega_\mu = \mathsf{betapdf}(y_\mu - y_1, k, \theta)$, where y_1 is the lowest level of abilities. The beta distribution is defined over $y_1 + [0,1]$, has a unique mode $y_1 + \frac{k-1}{k+\theta-2}$ and an average equal to $y_1 + \frac{k}{k+\theta}$. The vector of the estimated parameters is

$$\vartheta = \{\rho, \sigma, \nu, s, \kappa, \gamma, \phi, \eta, y_1, k, \theta, \zeta, \underline{w}, \sigma_f, \sigma_s, \sigma_u\}.$$

Our structural approach aims to test the ability of our extended DMP model to fit the aggregate labor market data, considering the occasionally binding constraints, and given that these events depend on the non-stochastic component of productivity heterogeneity and labor market frictions.

3.4   Results

The estimates of the structural parameters in Table 1 provide different values than those usually obtained in the literature. Our estimation provides values for some of the parameters

Table 1: Parameter estimates

Param.	Value	Std.	Interpretation
ρ	0.97874	0.00062	Persistence of technology shock
σ	0.00292	0.00003	Std. innovations in the technology shock
γ	0.64074	0.00266	Bargaining power of workers
ζ	0.65000	0.04903	Indexation of home production
$\underline{w}$	0.92653	0.00040	Lower limit of the bargained wages
s	0.00691	0.00002	Exogenous separation rate
ν	0.68425	0.00391	Matching function
κ	0.43385	0.00401	Cost of a vacancy
k	2.00000	0.00375	Param. 1, Beta dist., abilities
θ	5.50000	0.02612	Param. 2, Beta dist., abilities
y_1	0.89922	0.00044	Lower bound of ability
ϕ	0.55117	0.29744	Scale parameter, firing cost function
η	1.50535	0.13767	Elasticity, firing cost function
σ_f	0.03329	0.00088	Std. measurement error on JFR
σ_s	0.00193	0.00005	Std. measurement error on JSR
σ_u	0.00107	0.00004	Std. measurement error on UR
Log-Lik.	9282.18		
Penalty	11.23		

CPS monthly data, 1951M1–2018M12, 40,000 particles, 30 abilities. Authors' calculations.
See Appendix C for the technical details on the estimation procedure and Appendix E to obtain details about the shape of the likelihood.

that could not be calibrated in previous econometric studies.

Matching function and search costs. Our estimate shows that the parameter ν is lower than that of Den Haan et al. (2000) (1.27) based on a calibration method targeting a selection of second-order moments, whereas it is more in line with Hagendorn & Manovskii (2008)'s calibration (0.407), which targets the first-order moment of the job-finding rate. The cost of posting a vacancy κ is lower than the values used in the SaM literature (see, e.g., Petrosky-Nadeau & Zhang (2017)).

Worker ability distribution. Figure 1 shows the estimated distribution of abilities; we find the same left-skewed distribution as Robin (2011), a characteristic also shared by the wage distribution. The abilities for which $n_t(\mu) = 0$, $\forall t$ are $\mu \in [1; 2]$, representing $7 \times 10^{-3}\%$ of workers.

Opportunity cost of employment and bargaining power. Figure 1 shows that $z(\mu)/y_\mu$ decreases with μ, given that the average value of the opportunity cost of employment is constrained to be equal to $\mathbb{E}_\mu[z(\mu)]/\mathbb{E}_\mu[y_\mu] = 0.75$. The value for z_0 implied by this last constraint is 0.33311. The workers' bargaining power is larger than 0.5, the arbitrary value retained by Den Haan et al. (2000) e.g..

Figure 1: Heterogeneity, opportunity cost of employment, $\underline{w}$, and wages
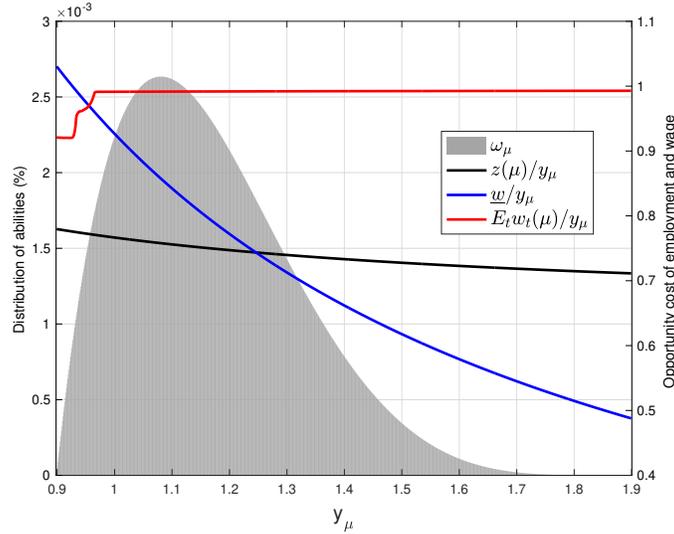
Lower limit of the bargained wages. The estimation leads to a value of $\underline{w} = 0.92653$, which corresponds to 79% of the average productivity ($\underline{w}/\mathbb{E}_\mu[y_\mu] = 0.79$). This is the value of real downward rigidity necessary to allow the model to generate sufficiently large fluctuations, given the persistence and variance of the exogenous shock. Figure 1 shows the ranking between the $\underline{w}$ and opportunity cost of employment and the induced average values of wages for each μ and shows that $\underline{w} > \mathbb{E}_t w_t(\mu)$ for a small set of workers (1.47% of the distribution).

Figure 2: Occurrence of regimes over the business cycle
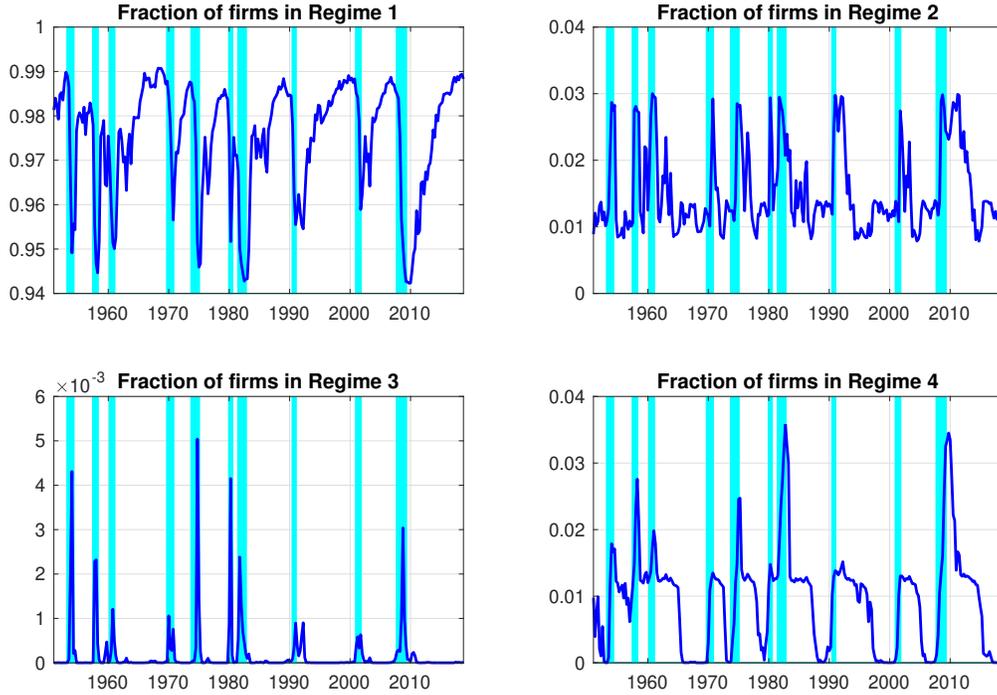
The blue areas represent NBER recession periods

Firing costs. These costs are convex. The standard errors are larger than those of the other parameters because Regime 3 occurs infrequently. As shown in Figure 2, the proportions of firms in each regime for each period of the estimation are 97.51% for Regime 1 and 0.94% for Regime 4, thus leaving 1.55% of the firms in Regimes 2 and 3. Nevertheless, the rich employment dynamics of the labor markets experiencing Regimes 2 and 3, and their non-negligible population sizes, allow us to identify the parameters of the firing cost functions.

The estimation of the parameters of the firing costs function governs the weight of the endogenous separation in unemployment dynamics, a crucial point discussed by Fujita & Ramey (2009) and Elsby et al. (2009). In our estimates, the exogenous separation accounts for 30% of JSR, a lower value than those found by several studies since Den Haan et al. (2000) ($\approx$50%). Beyond the level of separations, the endogenous component, which varies over time, also makes it possible to predict unemployment fluctuations accurately. Without endogenous job separations, the model underestimates the increases in the unemployment rate during recessions (See Appendix G), underlining that the relative contribution of the separation rate to unemployment volatility is higher than that measured by Shimer (2004).

Estimation based on the linearized model The estimation based on the linearized version of the model (see Appendix I) does not lead to the same estimated parameters. Particularly, if we impose that the opportunity cost of employment must be equal to 75%

of the average productivity (in the range of Hall & Milgrom (2008) estimates), the variance of the aggregate shock $(\sigma^2/(1 - \rho^2))$ is 62 times larger than in the estimation based on the nonlinear model. Therefore, the linearized version of the model cannot solve Shimer's puzzle (the possibility of generating the size of the labor market fluctuations with a shock having a size close to the variance of the US labor productivity) because of a too low wage rigidity to generate large adjustments in quantities. Moreover, by omitting the occasionally binding constraints (which are binding in some periods, as shown by the nonlinearized version of the model), the linearized version of the model generates negative values for vacancies (between 0.3% and 2.7% of the realizations of long simulated samples for each submarket), thus cautioning the robustness of these estimates.[10]

3.5 Model fit

Figure 3 displays the predicted series of JFR, JSR, and UR. The R^2 coefficients are 78%, 65%, and 99%, respectively. Contrary to Robin's (2011) model, our model captures the

Figure 3: Model fit (blue line = data, red line = model)

---
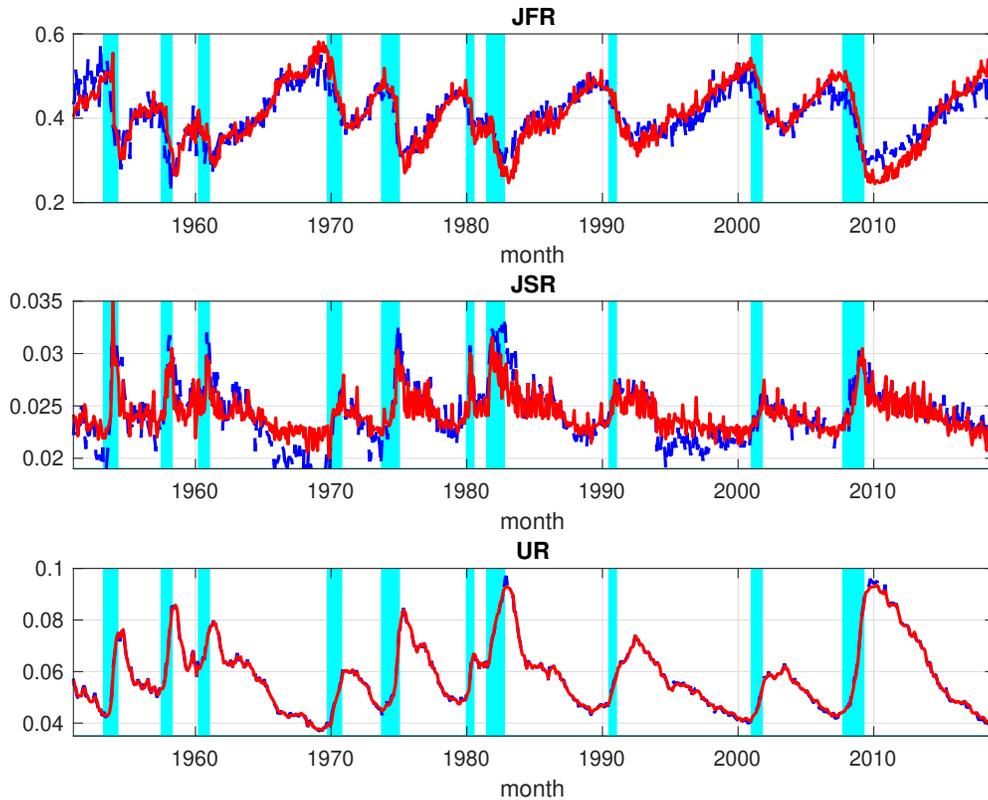
[10]The occurrence of negative or zero vacant jobs increases to reach between 0.9% and 4.6% of realizations in each submarket when the calibration of the opportunity cost of employment is $\mathbb{E}_\mu[z(\mu)]/\mathbb{E}_\mu[y_\mu] = 0.95$, instead of $\mathbb{E}_\mu[z(\mu)]/\mathbb{E}_\mu[y_\mu] = 0.75$, in order to reduce the variance of the aggregate shock.

16

Table 2: Implied moments

		Levels (deepness)				Std.
		Mean	Variance	Skewness	Kurtosis	Log
JFR	simul.	0.41035	5.0878×10^{-3}	-0.17903	2.5239	0.18269
	hist.	0.40889	3.4654×10^{-3}	-0.06588	2.4716	0.14791
JSR	simul.	0.024264	3.3377×10^{-6}	1.3249	5.8007	0.07237
	hist.	0.024182	7.8797×10^{-6}	0.41964	3.59	0.11549
UR	simul.	0.057916	1.7076×10^{-4}	0.82796	3.1125	0.21580
	hist.	0.057972	1.7527×10^{-4}	0.81194	3.1629	0.21909
		Differences (steepness)				
		mean	variance	skewness	kurtosis	
JFR	simul.	2.0317×10^{-4}	4.7890×10^{-4}	-0.75695	7.99068	
	hist.	4.141×10^{-5}	8.9366×10^{-4}	0.24748	4.5962	
JSR	simul.	-8.4964×10^{-7}	2.0056×10^{-6}	0.06678	4.8169	
	hist.	2.121×10^{-6}	1.8127×10^{-6}	0.34898	4.8209	
UR	simul.	-2.1241×10^{-5}	2.1530×10^{-6}	1.1186	8.4804	
	hist.	-2.1565×10^{-5}	3.4597×10^{-6}	0.67442	5.2748	

dynamics of the job separation rate well, which is particularly important for accounting for nonlinear adjustments in the labor market.

Table 2 shows that the worker flows (separations and hirings) and unemployment rate simulated by the model match the first-and second-order moments (mean and variances) as well as the higher-order moments summarizing the nonlinearities detected in the data (i.e., deepness and steepness). Particularly, the separation rate is right-skewed, and it exhibits excess kurtosis (the model reproduces this feature well), which are significant for JFR and UR (see Appendix A.4). This table also shows that the implied moments of the usual statistics reflecting labor market volatilities are reproduced using variables in logs.

Explaining the fit. Our model adds several dimensions to the basic DMP model: *(i)* the downward rigidity of the real wage, which constrains the labor market segments of individuals with low abilities; *(ii)* firing costs, which vary according to the number of dismissed employees; and *(iii)* the heterogeneity of workers' productivities. To evaluate the importance of these three extensions, we re-estimate concurrent models without $\underline{w}$ ($M1$), without firing costs ($M2$), and with a uniform distribution of abilities ($M3$). The log-likelihood ($\mathcal{L}^c$) of the benchmark model is higher than those of models $M2$ and $M3$, which also have an average value of the opportunity cost of employment ($\mathbb{E}_\mu[z(\mu)]/\mathbb{E}_\mu[y_\mu]$) equal to 0.75: $\mathcal{L}^c_{M0} = 9282.18 > \mathcal{L}^c_{M3} = 9145.41 > \mathcal{L}^c_{M2} = 9128.36$. The model without real wage rigidity has a higher log-likelihood than the benchmark model, $\mathcal{L}^c_{M1} = 9470.38 > \mathcal{L}^c_{M0} = 9282.18$.[11] This unsurprising result comes from the fact that the downward rigidity is not μ-specific and thus is less flexible than the opportunity cost depending on μ to fit the data. However, to obtain an estimation of this model $M1$, the average opportunity cost of employment must be set to the unrealistic value of 0.9 (given the empirical results of Hall & Milgrom (2008)),

---

[11]See Appendix F for the estimated parameters for all the models.

putting some doubt on the pertinence of this modeling of the labor market.

Figure 4: Gaps between benchmark model and alternative models
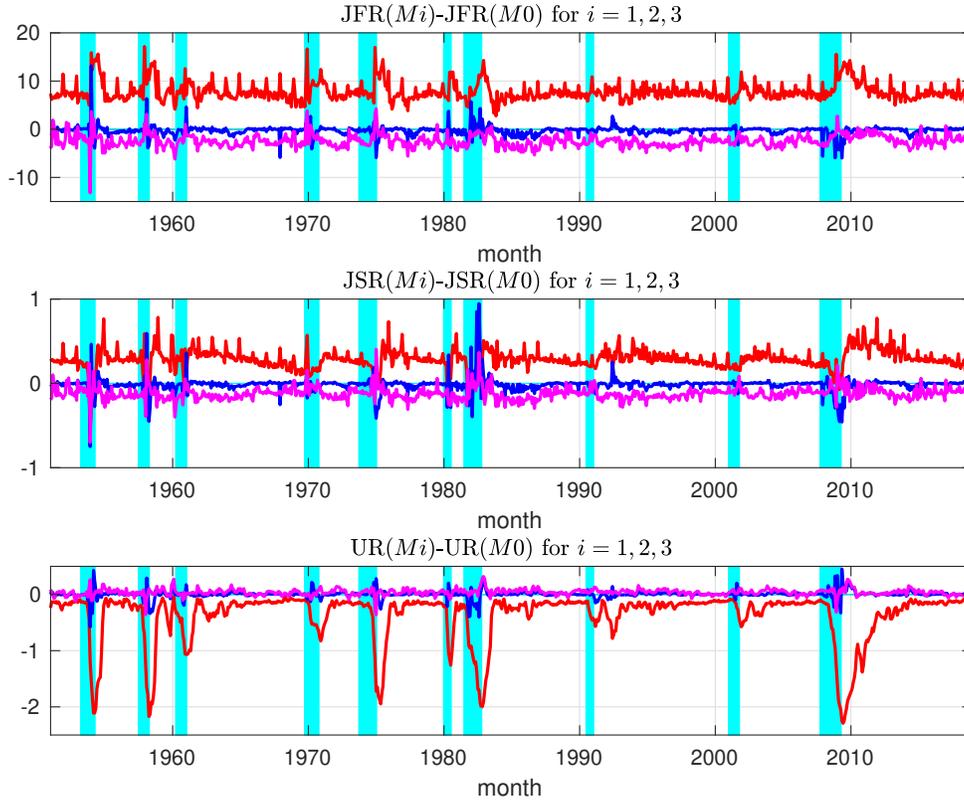
Red line: Model $M1$ - Model $M0$. Blue line: Model $M2$ - Model $M0$. Magenta line: Model $M3$ - Model $M0$. Blue areas: NBER recession periods.

Another way to analyze the model fit is to use the filter to find the realization of the shock a_t that minimizes the distance between the models and the data each month. Hence, without re-estimation of the model, it is possible to reveal the time series of the "net" labor productivity needed for each model to match the time series $\{JFR_t, JSR_t, UR_t\}$. Figure 4 displays the gaps between our benchmark model and the other three models for these three-time series. The fit is less good for all models Mi, $i = 1, 2, 3$ than for benchmark $M0$. Particularly, it appears that without downward real wage rigidity, the model cannot reach the unemployment rate's peaks and troughs. Moreover, job finding and job separation rates are overestimated on average. These failures are less important for models without firing costs or with a uniform distribution of abilities. This shows that the downward real wage rigidity helps to magnify the impact of the aggregate shock, whereas firing costs and non-uniform distribution of abilities are crucial assumptions to explain the recession episodes.

3.6 Model's shocks

First, Table 3 gives the gaps between the targeted moments on labor productivity and those based on model simulations. These gaps are small enough to show that the size of the exogenous shocks introduced in the model is controlled. The persistence and variance of $\log(a_t)$ are very close to those estimated on the labor productivity series used by Shimer (2005). However, the most interesting result is presented in Figure 5: conditional on an

Table 3: Productivity processes: data and model

	Persistence ρ_p	Variance σ_p^2
Data	0.954945	0.424741×10^{-4}
Model	0.925191	0.493543×10^{-4}

Estimation: CPS monthly data, 1951M1–2018M12, 40,000 particles, 30 abilities. Authors' calculations.
Data: BLS Quarterly data. Model: quarterly data are averages of weekly data. $p = s, d$

Figure 5: Model's structural shock and labor productivity
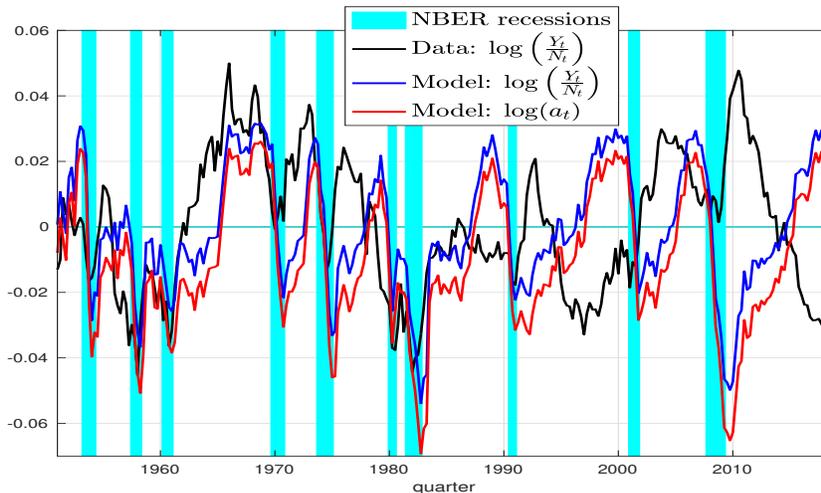
Model: "output" is $Y_t^s = \sum_\mu \omega_\mu y_\mu \sum_{\tau=1}^S a_{t,\tau} N_{t,\tau}(\mu)$ and $N_t^s = \frac{1}{S} \sum_\mu \omega_\mu \sum_{\tau=1}^S N_{t,\tau}(\mu)$ leading to "labor productivity" defined by Y_t^s/N_t^s. Here, t denotes the quarter, τ the weeks in each quarter, and S the number of weeks in a quarter. Data: BLS quarterly data, stationarized using the HP filter with $\lambda_{HP} = 2.5 \times 10^5$,

identical size, the historical chronicle of $\log(a_t)$ is very different from that of observed labor productivity. While these two series are highly synchronized before the mid-1980s, they vary in opposite directions thereafter. This result echoes Barnichon (2010), who shows that the correlation between unemployment and labor productivity has become positive after the mid-1980s, whereas it was negative before. This result shows as Ljungqvist & Sargent (2017) suggest that the "fundamental surplus" is hit by other shocks than those that change labor productivity, these shocks becoming dominant after the mid-1980s.[12] Finally, we can

---

[12]The explanation of the causes of these changes in the structure of shocks affecting labor market dynamics is beyond the scope of this article.

construct by aggregating all labor market segment indicators of the aggregate "output" (Y_t^s) and employment (N_t^s). However, given that a_t is not a TFP shock but can also integrate additional disturbances, this aggregate "output" measure is not a GDP but a resource net of exogenous firm cost fluctuations. This "net" labor productivity $(\frac{Y_t^s}{N_t^s})$ follows closely the historical chronicle of $\log(a_t)$ (see Figure 5), but $a_t \neq \frac{Y_t^s}{N_t^s}$ because there exist entries and exits (the distribution of the active firms changes over time).

Figure 6: Structural shocks for various models
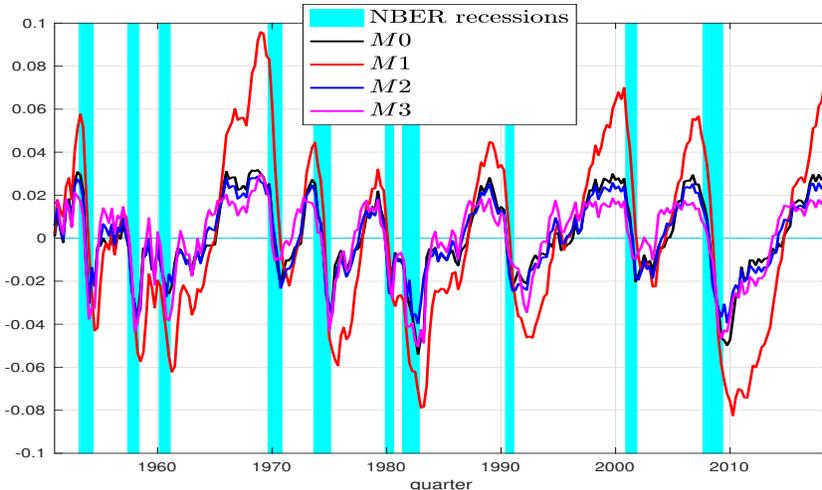
The lack of model elasticity when downward real wage rigidity is omitted (see above) can also be detected when comparing the dynamics of the "net" productivity necessary for the model to match observed data. Figure 6 shows that the magnitude of change in the "net" labor productivity $(\log(a_t))$ generated by the model without downward real wage rigidity $(M1)$ is larger than that of the benchmark model $(M0)$, showing that the elasticity of our benchmark model for the shock is larger than that of the $M1$ model. The models without firing costs $(M2)$ or with a uniform distribution of workers' abilities $(M3)$ do not require a larger magnitude of the shocks, but require a different sequence of a_t, with lower peaks and deeper troughs, allowing these models to compensate for their lower performance in explaining recession episodes.

3.7 Aggregate shock and nonlinearities in unemployment

Figure 7 provides illustrations of the model's nonlinearities by comparing the generalized impulse responses (GIRF) conditional on the size (one vs. five standard deviations) and a sign of the shock as well as on the state of the economy at the time of the shock (recession vs. expansion). The dynamics of the job-finding rate are highly nonlinear, whereas it is not the case with the job separation rate. Particularly, the response of job openings following a positive shock has a greater amplitude than after a negative shock, the adjustment being made in the latter case via an increase in separations. When the shock size is greater, firms

Figure 7: Generalized Impulse Response Functions (GIRF)
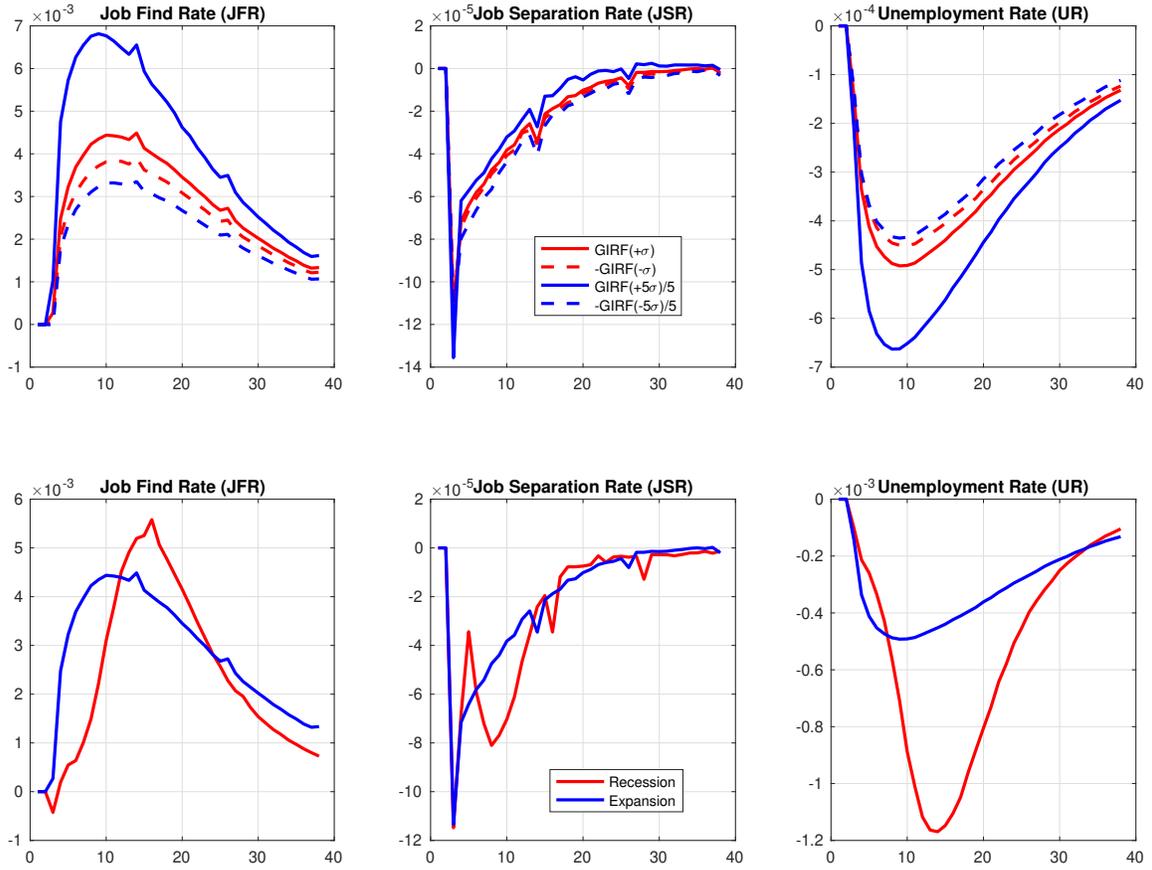
compete strongly for recruitment, which leads them to largely increase their job openings. As this competitive effect is less present in a recession because the number of unemployed is high at the time of the shock, the response of job openings is more moderate at the start, before this initial condition favorable to hiring disappears. Given the nonlinear link between the job-finding rate and the unemployment rate, the nonlinear adjustment is stronger for the unemployment rate than for the job-finding rate.

4 Additional implications for labor market aggregates

This section explores the main implications of our model for the data not used in its estimation because they are only available over a shorter sample (job losers) or are reconstructed from several sources (vacancy rate). In particular, we test whether the model is in accordance with the job losers data, which are an essential component of total separations, and the vacancy data, and thus the cyclical adjustments around the Beveridge curve.

Figure 8: Job losers
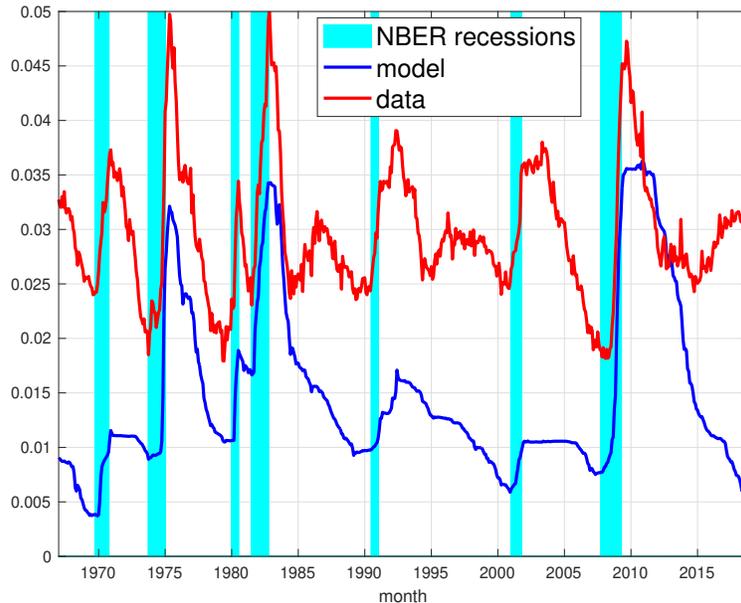
Job losers data (BLS) include both job losers and those who finish temporary jobs. Job losers in the model correspond to endogenous job separations (i.e., firings in Regime 3 and firm closures in Regime 4).

4.1 Job losers over the business cycle

The model predicts that separations are countercyclical, driven by the layoffs of low-skilled workers, whereas high-skilled workers become unemployed at a constant exogenous rate. These layoff inequalities are supported by empirical evidence provided by Davis et al. (1998) and Elsby et al. (2013). We test the model's predictions of layoffs by comparing their implications with the data by decomposing newly unemployed workers into job losers and job leavers. Our model predicts that a large proportion of the cyclical component of flows into unemployment is the result of endogenous job separations (i.e., firings in Regime 3 and firm closures in Regime 4). These separations are involuntary and thus correspond to job losers.[13] The model underestimates the average rate of job losers (0.0145 compared with 0.0308 in the data) because only low-ability workers can become job losers (they frequently experiment with Regimes 3 and 4). However, the cyclical components of the simulated and observed data are close, with a correlation of 0.65 (see Figure 8). As in the data, the model predicts that the most brutal crises, those of 1982 and 2008, are also the two episodes generating higher peaks in job losers.

---

[13]The data on job losers, which include both job losers and those who finish temporary jobs, are provided by the BLS and available at http://www.bls.gov/cps. Series LNS13023621 (1967M1 to 2018M12) is seasonally adjusted, stationarized using the HP filter with $\lambda_{HP} = 2.5 \times 10^5 \times 9^2$, and plotted around its average between 1967M1 and 1974M12 to purge the trend drift.

## 4.2 Dynamics of vacancies and the Beveridge curve

Figure 9: Vacancy rates: Model fit
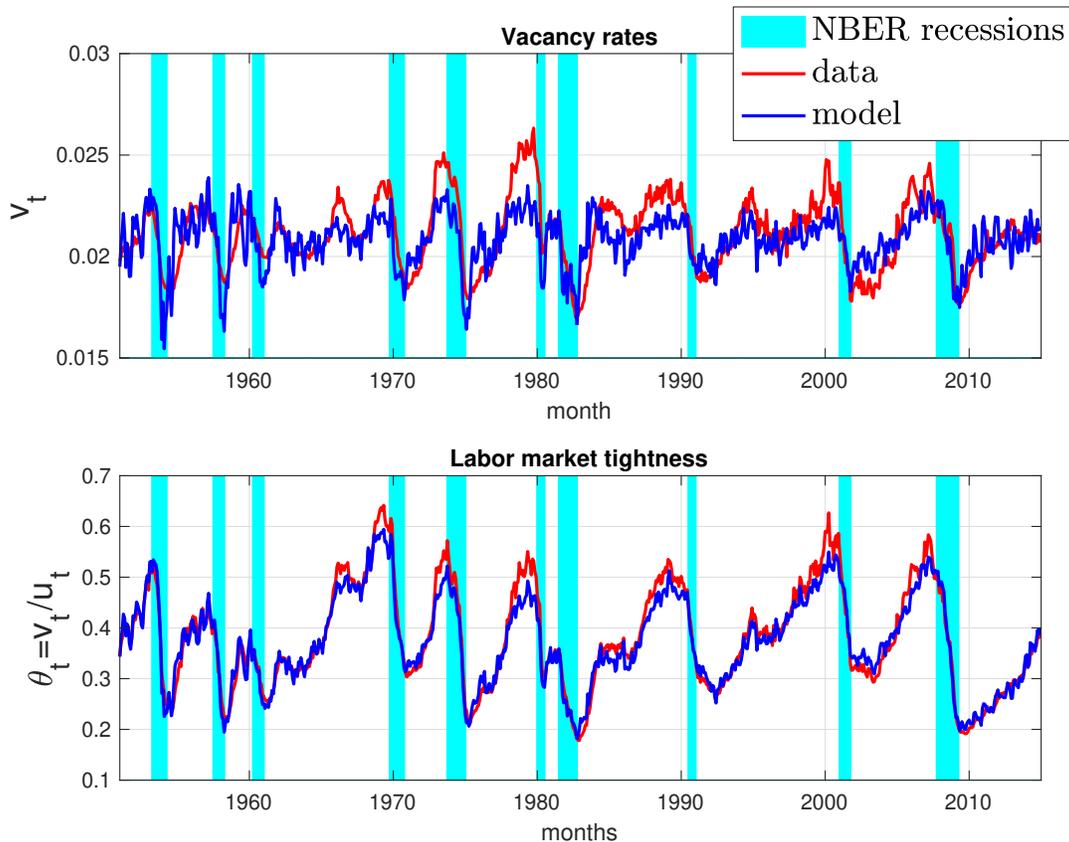
Although the estimation of the model is based on unemployment, job finding, and job separation rates, a crucial time series for the matching model is the vacancy rate. Figure 9 shows the dynamics of the simulated and historical time series (see Appendix A for the data description). The correlation between the two-time series is estimated to be 0.7, underscoring the good fit of the model vis-à-vis this additional dimension. Figure 9 also allows us to compare the historical and predicted labor market tightness dynamics; the gaps between these two-time series are small, with the R^2 coefficient equal to 96%.

The model predicts the Beveridge curve with high precision. The correlation between $\log(V)$ and $\log(U)$ predicted by the model is -0.73, whereas its counterpart in the data is -0.70. Hence, the model fits the observed dynamics of the job separation rate without generating a counterfactual positive correlation between vacancies and unemployment, as the Mortensen & Pissarides (1994) model does (see Fujita & Ramey (2012)). In our model, most labor market segments have a dynamic identical to that of a DMP model with exogenous separation. Therefore, most of the fluctuations in job vacancies are strongly procyclical, as in the data, despite a weaker correlation between vacant jobs and unemployment for low-ability workers' segments.

Figure 10: Beveridge curve
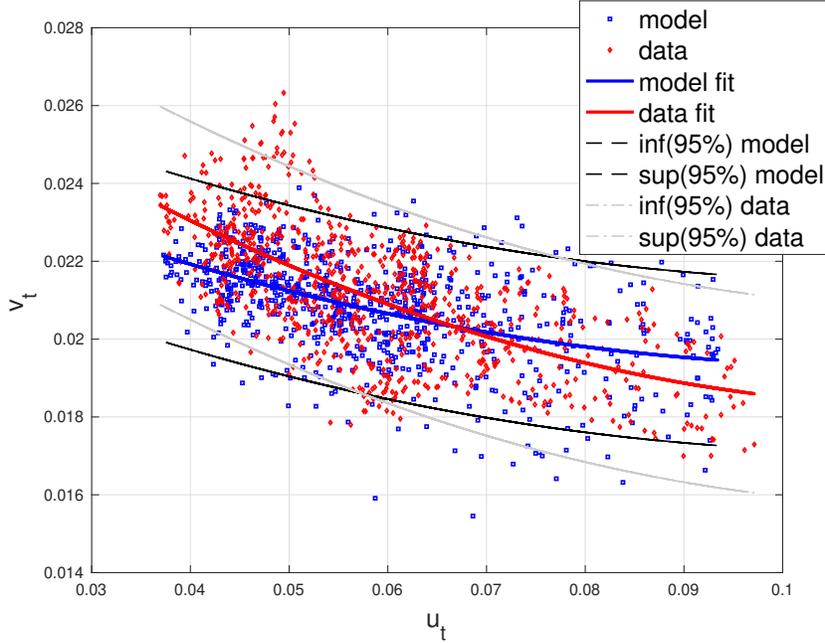
Another way to evaluate the model's fit is to compare its predictions before and after an unemployment peak (i.e., a recession) based on the joint dynamics of the vacancy and unemployment rates. This business cycle analysis is motivated by the debate on the nature of unemployment in the United States since the last recession, during which a notable shift in the Beveridge curve occurred. Indeed, "*[a] more in-depth analysis of the evidence suggests that the apparent shift in the relationship between vacancies and unemployment is neither unusual for a recession nor likely to be persistent*" (Bernanke (2012)). Diamond & Sahin (2015) provide convincing evidence that outward shifts in the Beveridge curve have been common during US recoveries since the 1950s. Figure 11 displays the joint movements of the vacancy and unemployment rates $(v_t; u_t)$ before and after a recession. We retain the same nine recession dates as in Diamond & Sahin (2015).[14] The model's predictions are compared with the $(v_t; u_t)$-adjustments observed in the US economy. For example, for the 2009 crisis, both the model and the data have an unemployment peak at 9.9%. This peak occurs after the period where unemployment increases in both the model and the data from 5.2% in April (M4) 2008 (16 months before) to 9.9%, whereas the vacancy rates (observed and simulated) slightly decline from 2% to 1.8%. This finding suggests a shift of the Beveridge curve. This view is supported by the slowness of the recovery: after this large shift in 16 months, Figure 11 shows that during the 16 months following the unemployment peak of October (M10) 2009, it reduced only slightly to reach 8.5%, with a small rise in vacancy rates from 1.8% to 2% in the same period.

---

[14]These recession dates are those of unemployment peaks: June 1958, March 1961, November 1971, April 1975, June 1980, November 1982, April 1992, March 2002, and October 2009. We use the 16-month periods before and after these unemployment peaks in our analysis.

Figure 11: Shifts in the Beveridge curve: Data vs. model
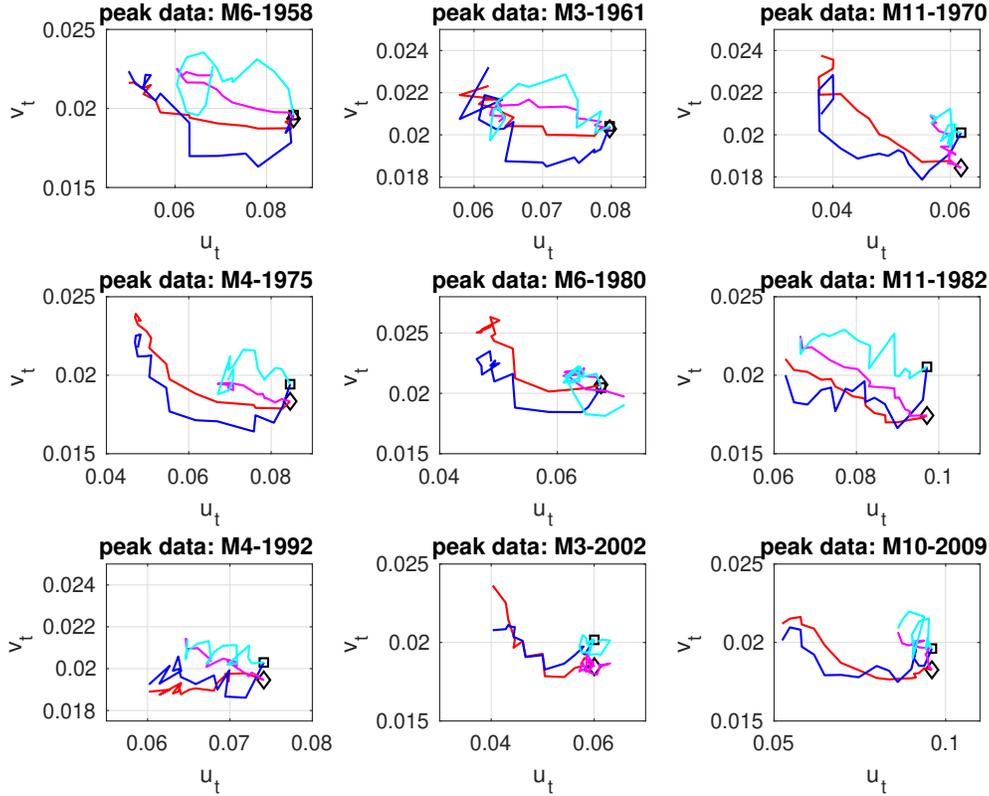
Data. Time series before (red lines) or after (magenta lines) the unemployment peak (diamond)
**Model.** Time series before (blue lines) or after (cyan lines) the unemployment peak (square)

By comparing all the recession episodes, Figure 11 shows that the slopes up to the peak are steep because they are driven by abrupt layoffs; by contrast, after the unemployment peak (i.e., during the recovery), the gradient is low because hiring is a time-consuming process. The model can also predict different adjustments after a recession, as is the case in the data; after the recessions of June 1958, March 1961, and November 1982, the recoveries were faster than those for the other recessions.

The differences between recessions come from two main factors: first, the unemployment level before the recession (see Table 4), and second, the size and persistence of the exogenous shock (see Figure 12). A high unemployment rate before the recession prevents the strong increase driven by the recession; the most fragile jobs are already absent from the distribution. Beyond this initial condition, a large and rapid shock leads to a strong unemployment peak. Moreover, if the shock persists to be at a low level, then it reduces for a long time the incentive to hire and thus brakes the recovery.

Table 4: Unemployment rates (%) 16 months before the unemployment peaks

Peak	M6-1958	M3-1961	M11-1970	M4-1975	M6-1980	M11-1982	M4-1992	M3-2002	M10-2009
UR	4.9870	6.2125	3.8001	4.7635	4.8765	6.2776	6.0166	4.0305	5.2372

Figure 12: Exogenous shocks shifting the Beveridge curve
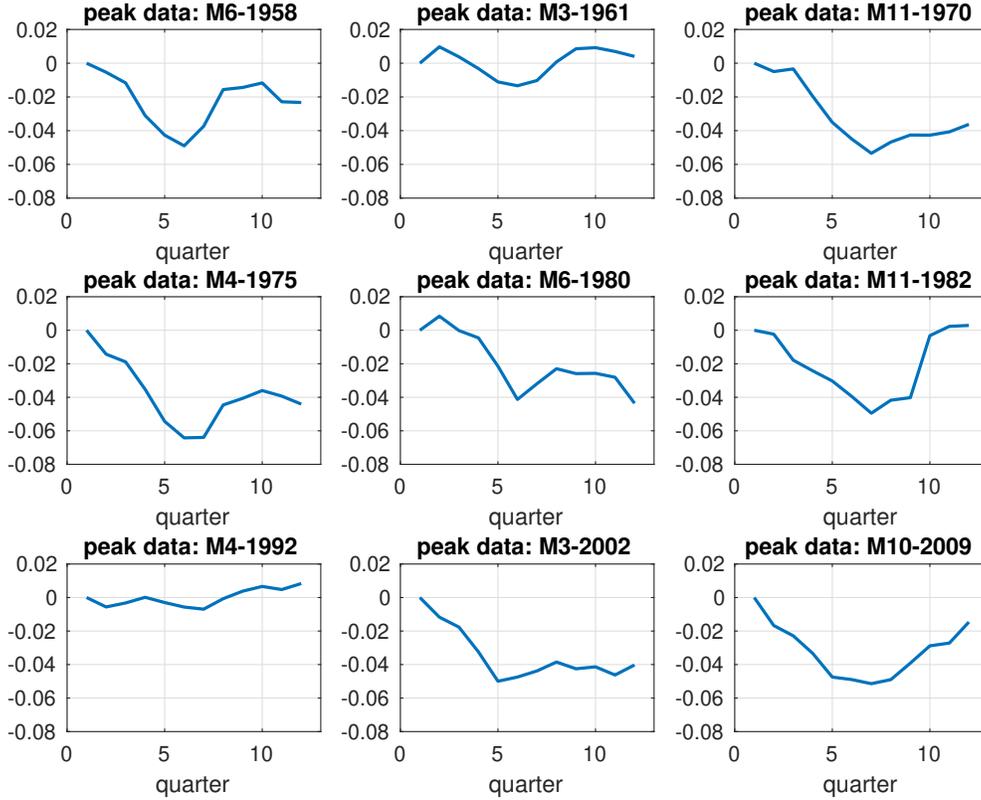
Quarterly data of $\log(a_t) - \log(a_{t0})$ with $t0$ is the date five quarters before the unemployment peaks

5   Estimated and observed heterogeneity

The ability of the model to reproduce US labor market dynamics depends on the estimated heterogeneity among workers. It is therefore important to check whether this heterogeneity accords with observed measures of worker dispersion, such as labor market fluctuations by educational attainment, as in Lise & Robin (2017).

Our model predicts that low-ability labor market segments are more sensitive than the other labor market segments, which explains the model's success accounting for the aggregate dynamics of the US labor market. In the model, unobservable worker heterogeneity corresponds to permanent differences between them. Thus, this heterogeneity could be linked to educational attainment.[15]

If a correspondence exists between our measure of abilities and educational attainment, the data should support the view that low-diploma worker flows and stocks are more sensitive to the business cycle. Figure 13 shows that the sensitivity of UR, JFR, and JSR to the

---

[15]The information on educational attainment (observable characteristic) is an imperfect measure of worker abilities valued by the labor market (unobservable characteristics): for example, in 2018, 1.4% of employees with a Bachelor's or a higher degree were paid at the minimum wage or below, suggesting that a high diploma does not necessarily imply high ability for working.

Figure 13: Differences in sensitivity to the business cycle across heterogeneous worker groups
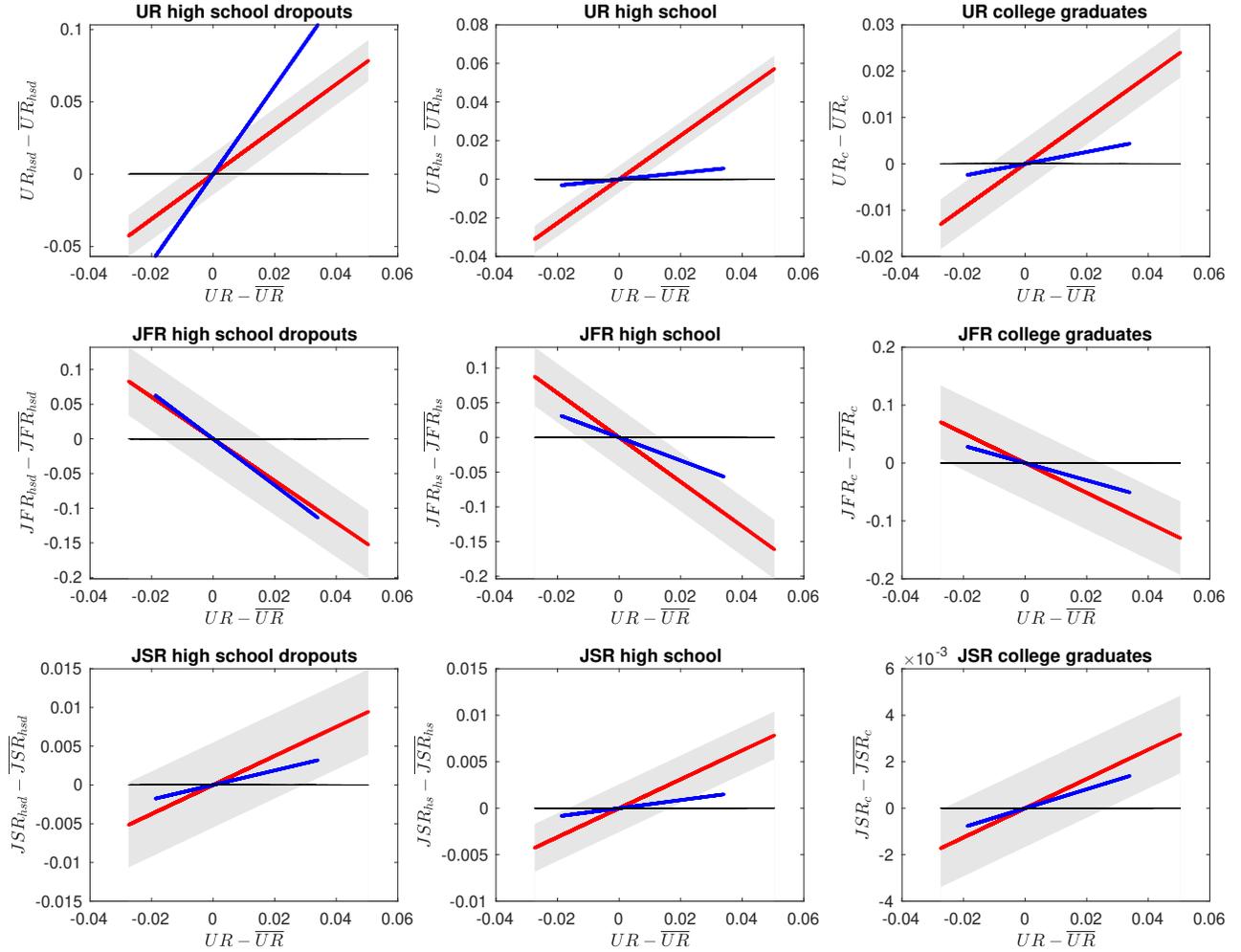
Red lines: observed data. Blue lines: simulated data. Gray areas: confidence interval 68%. Data provided by Lise & Robin (2017). The model assumes that "high school dropouts" (hsd) correspond to the bottom 25% of the ability distribution, "high school graduates" (hs) to 25–60%, and "college graduates" (c) to the top 40%. These weights correspond to the averages of the observed shares of educational attainment in the US population between 1970 and 2018. The x-axis reports the gap between the current aggregate unemployment (UR) and its mean over the sample ($\overline{UR}$): this gives us an indicator of the business cycle. The y-axis reports the gap between the current skill-specific labor market indicators (UR_s, JFR_s, JSR_s, for $s = hsd, hs, c$) and their means over the sample ($\overline{UR_s}, \overline{JSR_s}, \overline{JSR_s}$): this gives us a skill-specific indicator of fluctuations.

business cycle declines with workers' education level. More precisely, a one-point increase in the unemployment rate vis-à-vis its mean corresponds to an increase of 1.55 points in the unemployment rate of high school dropouts. These unemployment rate increases are only 1.13 points for high school graduates and 0.47 points for college graduates. These unequal unemployment risks among workers are largely driven by the larger sensitivity of JSR for low-diploma workers. The JSR of high school dropouts is three times more sensitive to aggregate fluctuations (here measured by the gap between unemployment and its average over the sample) than that of college graduates, whereas the JFR of high school dropouts is only 1.2 times more sensitive than that of college graduates (see Cairo & Cajner (2016)). As shown in Figure 13, the low-ability workers in the model have larger JSR sensitivity to the business cycle of the same order of magnitude as that of high school dropouts in the data. Moreover, the model slightly overestimates the sensitivity of the JFR of these workers, leading it to overestimate UR sensitivity. The model sensitivities to the business cycle for medium- and high-ability workers are not statistically different from those corresponding to high school and college graduates in the data. However, these encouraging results on business cycle sensitivity should not overshadow that the model generates too much churning on average. Indeed, as in Lise & Robin (2017), the averages of JFR and JSR for the three groups of heterogeneous agents in the model are higher than those measured in the data by educational attainment. Nevertheless, our model can replicate the average unemployment rate: for high school dropouts, this is 14.34% in the data and 13.6% in the model, whereas it is only 3.17% for college graduates (a gap of 11.17 points to high school dropouts) and 2.52% in the model (a gap of 11.08 points to low-ability workers).

6   Conclusion

This study presents an extended version of the DMP model to explain aggregate fluctuations in the US labor market. Based on an estimation using full information techniques, we demonstrate that nonlinearities and downward real wage rigidity are crucial for the model fit (i.e., job flows, unemployment rate, and the Beveridge curve) conditional on estimated heterogeneity. We then show that this revealed heterogeneity allows the model to reproduce the observed quantity adjustments conditional on the level of education.

In future work, we aim to plug this partial equilibrium analysis into a DSGE model, such as that of Kaplan et al. (2018), estimated using the same techniques. This approach may provide new evaluations of the effects of demand and supply shocks on labor market aggregates, dynamics of inequalities, and impacts of redistributive policies. We also plan to introduce job-to-job transitions into an unsegmented market and thus contribute to modeling the interactions between heterogeneity and aggregate uncertainty in labor markets.

References

Acemoglu, D. & Scott, A. (1994), 'Asymmetries in the cyclical behavior of UK labour markets', *The Economic Journal* **104**, 1303–1323.

Acharya, S., Chen, W., Del Negro, M., Dogra, K., Matlin, E. & Sarfati, R. (2020), Estimating hank: Macro time series and micro moments, Mimeo, ASSA 2021 Econometric Society Session on "Estimation of Macroeconomic Models".

Adjemian, S., Bastani, H., Juillard, M., Karame, F., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M. & Villemot, S. (2011), Dynare Reference Manual, Version 4, Dynare Working Paper 1, Cepremap.

Bai, J. & Ng, S. (2005), 'Tests for skewness, kurtosis, and normality for time series data', *The Review of Economics and Statistics* **23**, 49–60.

Balleer, A. (2012), 'New evidence, old puzzles: Technology shocks and labor market dynamics', *Quantitative Economics* **3**(3), 363–392.

Barnichon, R. (2010), 'Productivity and unemployment over the business cycle', *Journal of Monetary Economics* **57**(8), 1013–1025.

Bentolila, S. & Bertola, G. (1990), 'Firing costs and labour demand: How bad is eurosclerosis?', *Review of Economic Studies* **57**(3), 381–402.

Bernanke, B. (2012), Recent developments in the labor market, http://www.federalreserve.gov/newsevents/speech/bernanke20120326a.htm, FED.

Bertola, G. & Caballero, R. (1994), 'Cross-sectional efficiency and labour hoarding in a matching model of unemployment', *Review of Economic Studies* **61**, 435–457.

Bertola, G. & Garibaldi, P. (2001), 'Wages and the size of firms in dynamic matching models', *Review of Economic Dynamics* **4**, 335–368.

Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. (2017), 'Julia: A fresh approach to numerical computing', *SIAM Review* **59**, 65–98.

Burgess, S. M. (1992), 'Nonlinear dynamics in a structural model of employment', *Journal of Applied Econometrics* **7**, S101–S118.

Burns, A. & Mitchell, W. (1946), Measuring Business Cycles: The Problem and Its Setting, Nber book series studies in business cycles, NBER.

Caballero, R. & Hammour, M. (1994), 'The cleansing effect of recessions', *The American Economic Review* **84**(5), 1350–1368.

Cahuc, P. & Wasmer, E. (2001), 'Does intrafirm bargaining matter in the large firm's matching model?', *Macroeconomic Dynamics* (5), 742–747.

Cairo, I. & Cajner, T. (2016), 'Human capital and unemployment dynamics: Why more educated workers enjoy greater employment stability', *The Economic Journal* **128**, 652–682.

Chassamboulli, A. (2013), 'Labor-market volatility in a matching model with worker heterogeneity and endogenous separations', *Labour Economics* **24**, 217–229.

Christiano, L., Eichenbaum, M. & Trabandt, M. (2016), 'Unemployment and business cycles', *Econometrica* **84**(4), 1523–1569.

Christiano, L. & Fischer, D. (2000), 'Algorithms for solving dynamic models with occasionally binding constraints', *Journal of Economic Dynamics and Control* **24**, 1179–1232.

Coles, M. & Kelishomi, A. (2018), 'Do job destruction shocks matter in the theory of unemployment?', *American Economic Journal: Macroeconomics* **10**(3), 118–136.

Collard, F., Feve, P., Langot, F. & Perraudin, C. (2002), 'A Structural Model of US Aggregate Job Flows', *Journal of Applied Econometrics* **17**, 197–223.

Davis, S., Haltiwanger, J. & Schuh, S. (1998), *Job Creation and Destruction*, The MIT Press, Cambridge.

DeJong, D. N., Ingram, B. F. & Whiteman, C. H. (2000), 'A Bayesian Approach to Dynamic Macroeconomics', *Journal of Econometrics* **98**(2), 203 ï¿½ 223.

Den Haan, W. & Levin, A. (1996), A practitioner's guide to robust covariance matrix estimation, Technical report, NBER.

Den Haan, W., Ramey, G. & Watson, J. (2000), 'Job destruction and propagation of shocks', *American Economic Review* **90**, 482–498.

Diamond, P. (1982), 'Wage determination and efficiency in search equilibrium', *Review of Economic Studies* **49**, 217–227.

Diamond, P. & Sahin, A. (2015), 'Shifts in the Beveridge curve', *Research in Economics* **69**, 18–25.

Douc, R., Cappe, O. & Moulines, E. (2005), 'Comparison of Resampling Schemes for Particle Filtering', *4th International Symposium on Image and Signal Processing and Analysis (ISPA), Zagreb, Croatia* .

Elsby, M., Hobijn, B. & Sahin, A. (2013), 'Unemployment dynamics in the OECD', *Review of Economics and Statistics* **95**(2), 530–548.

Elsby, M., Michaels, R. & Solon, G. (2009), 'The ins and outs of cyclical unemployment', *American Economic Journal: Macroeconomics* **1**(1), 84–110.

Fernandez-Villaverde, J. & Rubio-Ramirez, J. (2005), 'Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood', *Journal of Applied Econometrics* **20**, 891–910.

Fernandez-Villaverde, J., Rubio-Ramirez, J. & Schorfheide, F. (2015), 'Solution and Estimation Methods for DSGE Models', *Handbook of Macroeconomics, John B. Taylor, Harald Uhlig (Editors)* **2**.

Ferraro, D. (2017), The asymmetric cyclical behavior of the U.S. labor market, Working-paper, Arizona State University.

Fujita, S. & Ramey, G. (2009), 'The cyclicality of separation and job finding rates', *International Economic Review* **50**(2), 415–430.

Fujita, S. & Ramey, G. (2012), 'Exogenous versus endogenous separation', *American Economic Journal : Macroeconomics* **4**(4), 68–93.

Gordon, N., Salmond, D. & Smith, A. (1993), 'Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation', *IEEE Proceedings-F* **140**, 107–113.

Hagendorn, M. & Manovskii, I. (2008), 'The cyclical behavior of equilibrium unemployment and vacancies revisited', *American Economic Review* **98**(4), 1692–1706.

Hairault, J.-O., Langot, F. & Osotimehin, S. (2010), 'Matching frictions, unemployment dynamics and the cost of business cycles', *Review of Economic Dynamics* **13**(4), 759–779.

Hall, R. E. (2005), 'Employment fluctuations with equilibrium wage stickiness', *American Economic Review* **95**(1), 50–65.

Hall, R. E. & Milgrom, P. R. (2008), 'The limited influence of unemployment on the wage bargain', *American Economic Review* **98**(4), 1653–1674.

Herbst, E. & Schorfheide, F. (2015), *Bayesian Estimation of DSGE Models*, Princeton University Press.

Iliopulos, E., Langot, F. & Sopraseuth, T. (2019), 'Welfare cost of fluctuations when labor market search interacts with financial frictions', *Journal of Money, Credit and Banking* **51**(8), 2207–2237.

Judd, K. (1992), 'Projection methods for solving aggregate growth models', *Journal of Economic Theory* **58**, 410–452.

Jung, P. & Kuester, K. (2011), 'The (un)importance of unemployment fluctuations for the welfare cost of business cycles', *Journal of Economic Dynamics and Control* **35**(10), 1744–1768.

Kantas, N., Doucet, A., Singh, S., Maciejowski, J. & Chopin, N. (2015), 'On Particle Methods for Parameter Estimation in State-Space Models', *Statistical Sciences* **30**(3), 328–351.

Kaplan, G., Moll, B. & Violante, G. (2018), 'Monetary policy according to HANK', *American Economic Review* .

Keynes, J. (1936), *The General Theory of Employment, Interest and Money*, MacMillan, London.

Lise, J. & Robin, J. (2017), 'The macro-dynamics of sorting between workers and firms', *American Economic Review* **107**(4), 1104–1135.

Ljungqvist, L. & Sargent, T. (2017), 'The fundamental surplus', *American Economic Review* **107**(9), 2630–2665.

McKay, A. & Reis, R. (2008), 'The brevity and violence of contractions and expansions', *Journal of Monetary Economics* **55**, 738–751.

Mortensen, D. (1982), The matching process as a noncooperative bargaining game, *in* J. J. McCall, ed., 'The Economics of Information and Uncertainty', University of Chicago Press.

Mortensen, D. & Pissarides, C. (1994), 'Job creation and Job destruction in the theory of unemployment', *Review of Economic Studies* **61**, 397–415.

Neftci, S. (1984), 'Are Economic Time Series Asymmetric over the Business Cycle?', *Journal of Political Economy* **92**(2), 307–328.

Newey, W. & West, K. (1987), 'A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica* **55**, 703–708.

Petrosky-Nadeau, N. & Zhang, L. (2017), 'Solving the DMP model accurately', *Quantitative Economics* **8**, 611–650.

Petrosky-Nadeau, N. & Zhang, L. (2020), 'Unemployment crises', *Journal of Monetary Economics* .

Petrosky-Nadeau, N., Zhang, L. & Kuehn, L. (2018), 'Endogenous disasters', *American Economic Review* **108**, 2212–2245.

Pissarides, C. (1985), 'Short-run dynamics of unemployment, vacancies, and real wages', *American Economic Review* **75**, 676–690.

Pissarides, C. (1990), *Equilibrium Unemployment Theory*, Basil Blackwell, Oxford.

Pissarides, C. (2009), 'The unemployment volatility puzzle: Is wage stickiness the answer?', *Econometrica* **77**(5), 1339–1369.

Ravn, M. & Uhlig, H. (2002), 'On adjusting the Hodrick-Prescott filter for the frequency of observations', *The Review of Economics and Statistics* **84**, 371–380.

Robin, J. (2011), 'On the dynamics of unemployment and wage distributions', *Econometrica* **79**(5), 1327–1355.

Saint-Paul, G. (1993), 'Productivity growth and the structure of the business cycle', *European Economic Review* **37**(4), 861–883.

Schorfheide, F. (2000), 'Loss Function-based Evaluation of DSGE Models', *Journal of Applied Econometrics* **15**(6), 645–670.

Shimer, R. (2004), Search intensity, Mimeo, University of Chicago.

Shimer, R. (2005), 'The cyclical behavior of equilibrium unemployment and vacancies', *American Economic Review* **95**(1), 25–49.

Shimer, R. (2012), 'Reassessing the ins and outs of unemployment', *Review of Economic Dynamics* **15**, 127–148.

Sichel, E. (1993), 'Business cycle asymmetry : a deeper look', *Economic inquiry* **31**.

Stole, A., L. & Zwiebel, J. (1996), 'Organizational design and technology change under intrafirm bargaining', *American Economic Review* **86**, 195–222.

Wasmer, E. & Weil, P. (2004), 'The macroeconomics of labor and credit market imperfections', *American Economic Review* **94**(4), 944–963.

Appendix: For Online Publication

A   Data

A.1   Observed unemployment and worker flows

We extend the data in Lise & Robin (2017) to the current period; the data are from the BLS and cover 1951m1 to 2018m12.[16] These monthly employment and unemployment levels for all people aged 16 and over are seasonally adjusted. To construct worker flows, we also use the number of individuals unemployed for more than five weeks.

After dividing the unemployment levels in each month by the sum of unemployment and employment, we obtain monthly series for U_m and U_m^5 (m refers to the monthly frequency), which corresponds to the proportion of unemployed individuals and the proportion of individuals unemployed for more than five weeks. The worker flows are given by:

$$JSR_m = \frac{U_{m+1} - U_{m+1}^5}{E_m}, \qquad JFR_m = \frac{U_m - U_{m+1}^5}{U_m}.$$

Figure 14: Worker flows and the unemployment rate in the United States ($HP = 9^2 \times 2.5 \times 10^5$)
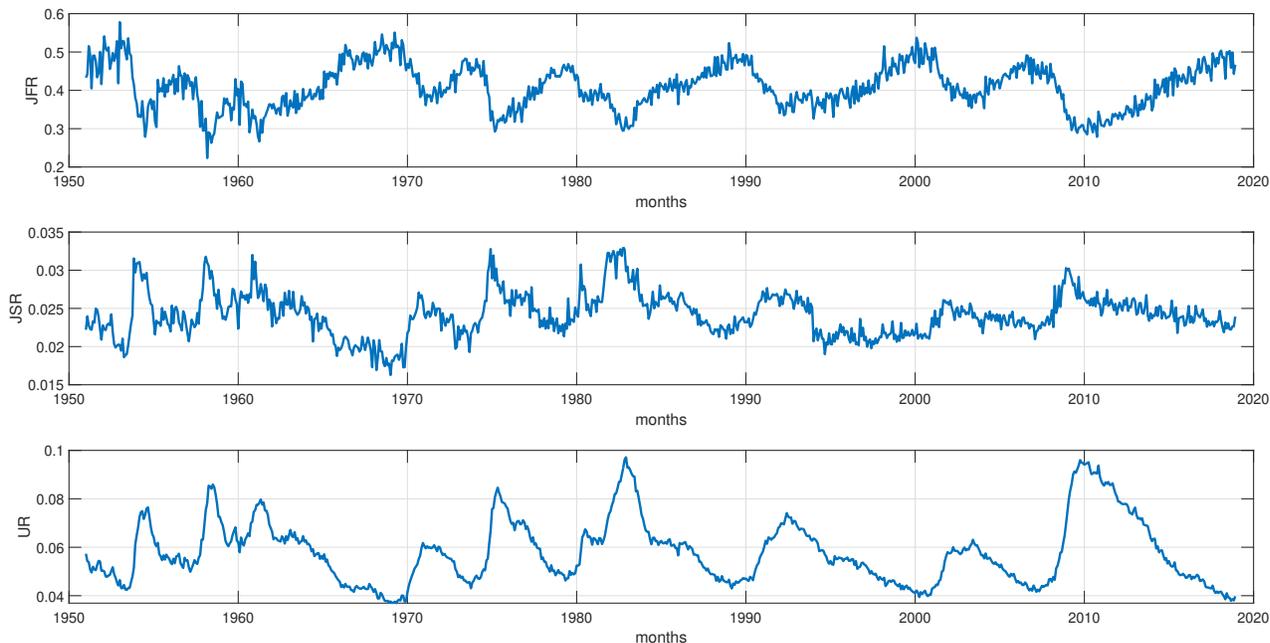
A.2 Simulated data

The model generates weekly data (subscript w denotes weekly data). Thus, we recursively construct the weekly series of the number of workers unemployed for five weeks as

$$u_w^5(\mu) = u_{w-5}(\mu) \prod_{j=0}^{4}(1 - p(\theta_{w-5+j}(\mu))), \qquad u_w^5 = \sum_{\mu} \omega_\mu u_w^5(\mu).$$

The monthly transition rates are calculated as in the historical data. Specifically, for the weeks $w \in \{2, 7, 11, 15, 19, 24, 28, 33, 37, 42, 46, 50\}$, we collect information such as that from the BLS and build $\forall m \in [1, 12]$:

$$JSR_m^{THEO} = \frac{u_{w+1} - u_{w+1}^5}{e_w} \qquad JFR_m^{THEO} = \frac{u_w - u_{w+1}^5}{u_w},$$

with $u_w = \sum_{\mu} \omega_\mu u_w(\mu)$ and $e_w = \sum_{\mu} \omega_\mu(1 - u_w(\mu))$.

A.3 Vacancy rate

This time series is the help-wanted index constructed by Barnichon (2010). It combines the help-wanted advertisements in 51 major newspapers measured by the Conference Board and available over 1951m1–2008m5 and a new measure proposed by the Conference Board since 2005m5 that integrates online advertisements. To transform this index into a vacancy rate comparable to ours, we first detrend the raw data using the HP filter and compute the growth rates for each date. We then apply these growth rates to the initial vacancy rate levels provided by our estimated model to obtain the cyclical component of the vacancy rate used in this study.

A.4 Statistical nonlinearities in the data

Table 5 provides statistical evidence suggesting that labor market adjustments are highly nonlinear. Following Sichel (1993), the first type of asymmetry, "asymmetry in levels" (or deepness), is based on the fact that cyclical troughs are deeper or shallower than peaks are high. This type of asymmetry, therefore, characterizes the difference in the amplitudes of cyclical fluctuations. The second type, described as "asymmetry in slopes" (or steepness), occurs when the variations in the aggregate variables are different in booms than in recessions. Therefore, this type of asymmetry characterizes the difference in the speeds of upward or downward adjustments in a series. To check the robustness of our results, we report the results of the test proposed by Bai & Ng (2005), in which statistical tests are computed using two types of optimal weight matrices.

Concerning deepness, the distributions of JSR and UR are right-skewed (with positive and significant skewness at the 10% level). The masses of these distributions are concentrated on the left and exhibit a long right tail, suggesting that recessions may greatly impact the levels of separation and unemployment.

Table 5: Third-order moments—Data in levels and in differences

Method of computing the optimal weight matrix	JFR		JSR		UR	
	NW	HAC	NW	HAC	NW	HAC
Deepness						
Skewness (data in levels)	-0.367	-0.242	1.474	1.418	1.568	1.830
P-value (Bai & Ng (2005) test)	0.356	0.404	0.070	0.077	0.058	0.033
Steepness						
Skewness (data in first differences)	1.570	1.607	1.642	1.588	2.400	2.105
P-value (Bai & Ng (2005) test)	0.058	0.053	0.050	0.056	0.008	0.017

CPS monthly data, 1951M1–2018M12 ($\lambda_{HP} = 9^2 \times 2.5 \times 10^5$). Authors' calculations.
Statistics use either Newey & West (1987)'s (NW) or Den Haan & Levin (1996)'s (HAC) method.

Table 5 also shows that these time series in first differences are not normally distributed (steepness); instead, the large positive variations generate a long right tail and more outliers. Hence, from the table and a visual inspection of the data (see Figure 14 in Appendix A), it is clear that nonlinearities and asymmetries are stylized facts of the US labor market.

B Solution using the projection method

Table 6 reports the four regimes of the model.

<div align="center">Table 6: Regimes as functions of $\Theta_\mu(a_t, n_t(\mu))$</div>

Regime 4	Regime 3
$\Theta_\mu(a_t, n_t(\mu)) \leq -F'((1-s)n_t(\mu))$	$-F'((1-s)n_t(\mu)) \leq \Theta_\mu(a_t, n_t(\mu)) \leq 0$
$n_{t+1}(\mu) = 0$	$n_{t+1}(\mu) = (1 - s - l_t(\mu))n_t(\mu)$
$w_t(\mu) = \underline{w}$	$w_t(\mu) = \underline{w}$
Regime 2	**Regime 1**
$0 \leq \Theta_\mu(a_t, n_t(\mu)) \leq \kappa$	$\Theta_\mu(a_t, n_t(\mu)) > \kappa$
$n_{t+1}(\mu) = (1-s)n_t(\mu)$	$n_{t+1}(\mu) = (1-s)n_t(\mu) + q(\theta_t(\mu))v_t(\mu)$
$w_t(\mu) = \gamma y_\mu a_t + (1-\gamma)z(\mu)$	$w_t(\mu) = \gamma(y_\mu a_t + \kappa\theta_t(\mu)) + (1-\gamma)z(\mu)$

Numerical solution. The approximation method is based on Judd (1992). Suppose that there exists an approximation grid of the states $\{n_t(\mu)^{(i)}\}_{i=1}^{n_1} \times \{a_t^{(j)}\}_{j=1}^{n_2}$ for a given μ. We approximate

$$\frac{\kappa}{q(\theta_t(\mu))} - \lambda_t(\mu) = \Theta_\mu(n_t(\mu), a_t) \approx \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} c_{(i,j)}(\mu)\Gamma_{(i-1)}(\varphi_n(n_t(\mu)))\Gamma_{(j-1)}(\varphi_a(a_t)),$$

namely, a linear combination of Chebyshev polynomials $\left\{\Gamma_{(i-1)}(.)\right\}_{i=1:n}$ of maximal power $n-1$ in $n_t(\mu)$ or a_t and $\varphi_X(.)$, a function mapping $X \in [X_{min}, X_{max}]$ into $[-1, 1]$.

The set of parameters $c(\mu)$ depends on the model's structural parameters, and these are defined such that the model equations are verified on each node of the approximation grid (orthogonal collocation) with high accuracy. The goal is to implement the approximation function $\Theta_\mu(.)$ for any values of $n_t(\mu)$ and a_t.

To find the set of parameters $c(\mu)$, we use a guess-and-verify resolution approach.

Suppose that $n_t(\mu)^{(i)}$ and $a_t^{(j)}$ are particular nodes on the approximation grid. We guess $\lambda_t^{(i,j)} = 0$ and calculate the implied $q(\theta_t(\mu))^{(i,j)} = \frac{\kappa}{\Theta_\mu(n_t(\mu)^{(i)}, a_t^{(j)})}$. As this value is a probability, finding $q(\theta_t(\mu))^{(i,j)} > 1$ implies that the guess is wrong. We then correct it by setting $q(\theta_t(\mu))^{(i,j)} = 1$ and $\lambda_t^{(i,j)} = \kappa - \Theta_\mu(n_t(\mu)^{(i)}, a_t^{(j)})$. Then, the model equations can be implemented, and we deduce the associated values of the variables and multipliers for the current date.

We then calculate the states for the next period using the equations from Section 2.5 and approximating the Gaussian innovation ε_t with a sparse grid $\left\{\varepsilon_t^{(k)}, w_k\right\}_{k=1:n_3}$. Here, we use the scaled unscented transform proposed by Jullier and Ulhmann (1997). Hence, $n_3 = 3$ in this case. We then repeat the previous steps of the guess-and-verify approach at $t+1$ and

the approximate

$$\frac{\kappa}{q(\theta_t(\mu))} - \lambda_t(\mu) - \beta \mathbb{E}_t \left\{ \begin{array}{l} y_\mu a_{t+1} - w_{t+1}(\mu) \\ -l_{t+1}(\mu) F'(l_{t+1}(\mu) n_{t+1}(\mu)) \\ +(1 - s - l_{t+1}(\mu)) \left(\frac{\kappa}{q(\theta_{t+1}(\mu))} - \lambda_{t+1}(\mu) + \phi_{t+1}(\mu) \right) \\ +\nu_{t+1}(\mu) l_{t+1}(\mu) \end{array} \right\}$$

with

$$\frac{\kappa}{q(\theta_t(\mu))^{(i,j)}} - \lambda_t(\mu)^{(i,j)} - \beta \sum_{k=1}^{n_3} w_k \left\{ \begin{array}{l} y_\mu a_{t+1}^{(j,k)} - w_{t+1}(\mu)^{(i,j,k)} \\ -l_{t+1}(\mu)^{(i,j,k)} F'(l_{t+1}(\mu)^{(i,j,k)} n_{t+1}^{(i,k)}(\mu)) \\ +(1 - s - l_{t+1}(\mu)^{(i,j,k)}) \left(\frac{\kappa}{q(\theta_{t+1}(\mu))^{(i,j,k)}} - \lambda_{t+1}(\mu)^{(i,j,k)} + \phi_{t+1}(\mu)^{(i,j,k)} \right) \\ +\nu_{t+1}(\mu)^{(i,j,k)} l_{t+1}(\mu)^{(i,j,k)} \end{array} \right\}$$

The parameters $c(\mu)$ of the approximation rule are obtained when this expression is close to 0 with high accuracy for each node (i, j) of the grid. To solve this nonlinear system, we use Tensolve, a fast and efficient numerical method proposed by Bouaricha and Schnabel (1997).

In practice, we choose $n_1 = n_2 = 3$. The results do not change if we use higher values. The global method requires us to choose an interval of resolution for each state variable. For the productivity shock, we choose a grid centered on its ergodic mean ± 3 times its ergodic standard deviation, as is the standard. For employment, we choose the grid $[0, 1]$.

The resolution and estimation are implemented in Julia v1.6.2 (see Bezanson et al. (2017)).

C   Non-linear filtering and likelihood approximation

C.1   Non-linear state-space model

Once the model is solved using the projection method, it can be written in a nonlinear and non-Gaussian state-space form as follows:

$$x_t = f(x_{t-1}, \xi_t; \vartheta)$$
$$y_t = g(x_t; \vartheta) + e_t,$$

where y_t is the set of observed (measurement) variables; x_t is the set of unobserved state variables; e_t and ξ_t are the measurement and state innovations, respectively, which are assumed to be independently and identically Gaussian (with zero mean), and ϑ is a vector of unknown parameters. Here, $x_t = (a_t, n_t(1), \ldots, n_t(\mu), \ldots, n_t(M))'$ and $y_t = (JFR_t, JSR_t, UR_t)'$; $\xi_t = \varepsilon_t$ and $e_t = (e_t^f, e_t^s, e_t^u)'$.

From this model and the data, we extract two quantities of interest: the density of the state variables at time t conditional on the sample up to time t, that is, $p(x_t|y_{1:t}; \vartheta)$, and the one-step-ahead predictive density, $p(y_t|y_{t-1}; \vartheta)$, which is required to evaluate the likelihood. In principle, this extraction can be performed recursively. Given ϑ and $p(x_{t-1}|y_{1:t-1}; \vartheta)$, the prediction density is

$$p(x_t\,|y_{1:t-1}; \vartheta) = \int p(x_t\,|x_{t-1}; \vartheta)\, p(x_{t-1}\,|y_{1:t-1}; \vartheta)\, dx_{t-1},$$

where the first conditional density within the integral is deduced from the following state equation. The prediction is then updated using the observation y_t and Bayes' rule:

$$p(x_t\,|y_{1:t}; \vartheta) = \frac{p(y_t\,|x_t; \vartheta)\, p(x_t\,|y_{1:t-1}; \vartheta)}{p(y_t\,|y_{1:t-1}; \vartheta)},$$

where the evidence is defined as $p(y_t\,|y_{1:t-1}; \vartheta) = \int p(y_t\,|x_t; \vartheta)\, p(x_t\,|y_{1:t-1}; \vartheta)\, dx_t$. In nonlinear models, approximations are required to compute these densities recursively.

C.2   Monte Carlo approximations and sequential importance sampling

We consider a sequential importance-sampling algorithm to recursively approximate the aforementioned conditional densities. The method consists of choosing an easy-to-sample proposed distribution $q(x_t|y_{1:t}; \vartheta)$, which is informative about the target distribution $p(x_t|y_{1:t}; \vartheta)$ such that $q(x_t\,|y_{1:t}; \vartheta) = q(x_t|x_{t-1}, y_t; \vartheta) \times q(x_{t-1}\,|y_{1:t-1}; \vartheta)$. The approximation can then be implemented sequentially by randomly drawing the proposed particles in $q(x_t|x_{t-1}, y_t; \vartheta)$ at each date, given the approximation of the previous state distribution with a swarm of particles $\left\{x_{t-1}^{(i)}, w_{t-1}^{(i)}\right\}_{i=1:N}$. A swarm of particles $\left\{\tilde{x}_t^{(i)}\right\}_{i=1:N}$ is drawn in the proposed distribution $q(x_t|x_{t-1}^{(i)}, y_t; \vartheta)$ to approximate the distribution of the current state variables. Their

respective (unnormalized) weights are given by

$$\hat{w}_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(y_t|\tilde{x}_t^{(i)};\vartheta)p(\tilde{x}_t^{(i)}|x_{t-1}^{(i)};\vartheta)}{q(\tilde{x}_t^{(i)}|x_{t-1}^{(i)},y_t;\vartheta)}.$$

The conditional density $p(x_t|y_{1:t};\vartheta)$ is then approximated by $\left\{\tilde{x}_t^{(i)}, \tilde{w}_t^{(i)} = \frac{\hat{w}_t^{(i)}}{\sum_{i=1}^{N}\hat{w}_t^{(i)}}\right\}_{i=1:N}$, where $\tilde{w}_t^{(i)}$ is the normalized weight. Systematic resampling[17] is used to avoid the well-known issue of weight degeneracy. This process consists of randomly drawing particles from their empirical distribution, approximated by $\left\{\tilde{x}_t^{(i)}, \tilde{w}_t^{(i)}\right\}_{i=1:N}$, with replacement. It discards particles with low weights and replicates particles with high weights to focus on the interesting areas of the distribution using a constant number of particles. Consequently, all new particles have the same weight afterward, that is, $\left\{x_t^{(i)}, w_t^{(i)} = \frac{1}{N}\right\}_{i=1:N}$.

For simplicity, the usual choice for proposal $q(.)$ is $q(x_t|x_{t-1}, y_t;\vartheta) = p(x_t|x_{t-1};\vartheta)$. The weight expression simplifies to $\hat{w}_t^{(i)} \propto w_{t-1}^{(i)}p(y_t|\tilde{x}_t^{(i)};\vartheta)$, which is easy to write because

$$p(y_t|\tilde{x}_t^{(i)};\vartheta) = (2\pi)^{-\frac{\dim(y_t)}{2}} |P_e|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left[y_t - g(\tilde{x}^{(i)};\vartheta)\right]' P_e^{-1}\left[y_t - g(\tilde{x}^{(i)};\vartheta)\right]\right\},$$

where $P_e \equiv \mathbb{V}(e_t) = \text{diag}(\sigma_f^2, \sigma_s^2, \sigma_u^2)$ is the covariance matrix of the measurement errors. The sample likelihood is then obtained as follows.

$$p(y_{1:T}|\vartheta) = p(y_1|x_0;\vartheta)p(x_0|\vartheta)\prod_{t=2}^{T}p(y_t|y_{1:t-1};\vartheta),$$

where $p(y_t|y_{1:t-1};\vartheta) \approx \sum_{i=1}^{N}\hat{w}_t^{(i)} = \frac{1}{N}\sum_{i=1}^{N}p(y_t|\tilde{x}_t^{(i)};\vartheta)$ in the case of systematic resampling. Here, we choose $N = 40,000$ particles.

C.3 Implementation

From weekly to monthly simulated data. As the model generates weekly series and the measures are taken on a monthly basis (we assume a year is always 52 weeks), the resampling step only occurs when the simulated monthly series can be constructed and compared with its counterpart in the data, namely, in weeks $w \in \{2, 7, 11, 15, 19, 24, 28, 33, 37, 42, 46, 50\}$.[18]

Discretization of the ability distribution. In the estimation step, we do not use a uniform grid when discretizing the segment $[y_1, y_M]$, as doing so would require too many points and thus too much computational effort. We prefer to focus on the lower part of the

---

[17]This method was initially proposed by Gordon et al. (1993) and has been implemented for many algorithms (see Douc et al. (2005) for a comparative study).

[18]More details on the data are provided in Appendix A.

distribution (roughly the first decile), in which nonlinearities are important, and firms visit all or most of the regimes. Heterogeneity matters in this area. For the remaining 90% of the distribution, firms are always in Regime 1, which means they behave similarly in response to the aggregate shock. Using a fine grid on this part of the support would be a waste of computational time. Therefore, we build a uniform grid of 20 nodes on the first decile of the distribution and ten nodes for the remaining 90%. This approximation has a low cost, as we do not have to fit the distribution characteristics built from a fine grid of the distribution. During the simulation steps, and given the estimated parameters for the ability distribution, we use a uniform grid of 1,000 nodes to compute accurate statistics for all distributions.

Initialization of the filter   To avoid the usual adjustment period of the state variables and a poor fit of the measurement variables at the beginning of the sample owing to the approximative initial conditions on the states, we elicit an initial distribution of the state variables consistent with the first measure. We initialize the particle swarm using draws from the ergodic AR(1) Gaussian distribution for the technological shocks and draws from a uniform distribution on $[0, 1]$ for the employment sectors. To avoid a degenerate distribution at the initial point, we add a "learning period" of three years before the first sample observation and run the filter on a weekly basis. During this learning period, only the first observation of the sample is used for the resampling step. In other words, the states have 156 time periods to determine an initial particle swarm that most closely matches the first measurement observation before exploring the data. This initialization period is not incorporated in the likelihood calculation.

Maximization of the likelihood   The most important limitation of resampling is that it complicates the maximum log-likelihood inference using standard (gradient-based) numerical maximization techniques.[19] We then use non-gradient-based methods to maximize the model likelihood. As of the non-smoothness of the likelihood, the estimation delivers a "plausible" area for the parameters. After the estimation, polynomials are estimated to smooth the contour of the log-likelihood function in the neighborhood of each estimated value, keeping all the other parameters equal (see Figure 15 in Appendix E). Therefore, the standard errors of the parameters are obtained by taking the second derivatives of the estimated polynomial. A standard R^2 measure is presented for each parameter to assess the accuracy of the contour fit. This measure can be viewed as a confidence index for standard error evaluation (and, consequently, for the parameter *per se*).

---

[19]Even when the seed for the random draws is fixed across simulations, the traditional likelihood estimator depends on both resampled particles and unknown parameters. A small change in the parameter value causes a small change in the importance weights, which may generate a different set of resampled particles. Consequently, the likelihood function is non-differentiable, which explains why the applied approaches depart from the usual likelihood maximization of standard numerical techniques. For a recent survey of estimations in state-space models, see Kantas et al. (2015).

D   Decision rules

D.1   Regime (1/4): the "regime of hirings"
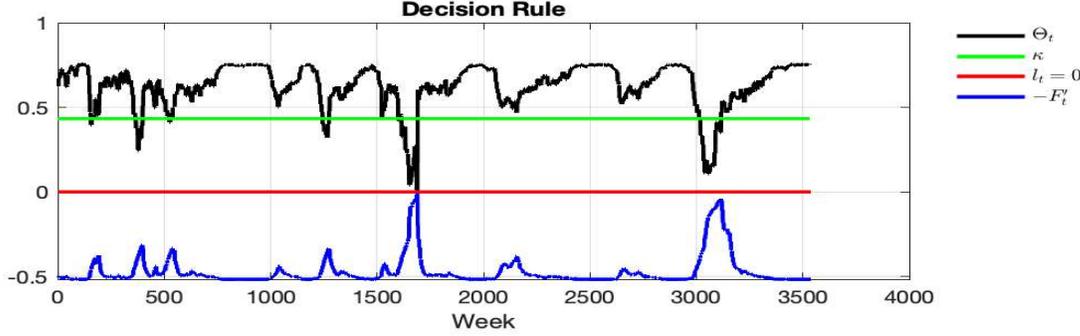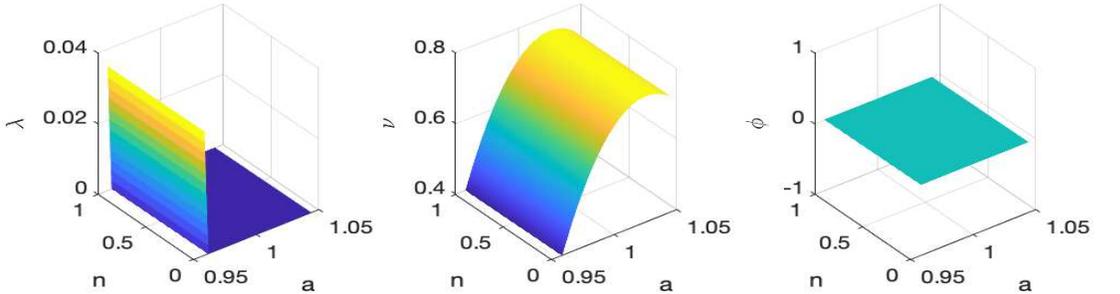
$\lambda_t(\mu) = 0 \Rightarrow v_t(\mu) > 0$ and $l_t(\mu) = 0 \Rightarrow \phi_t(\mu) = 0$.
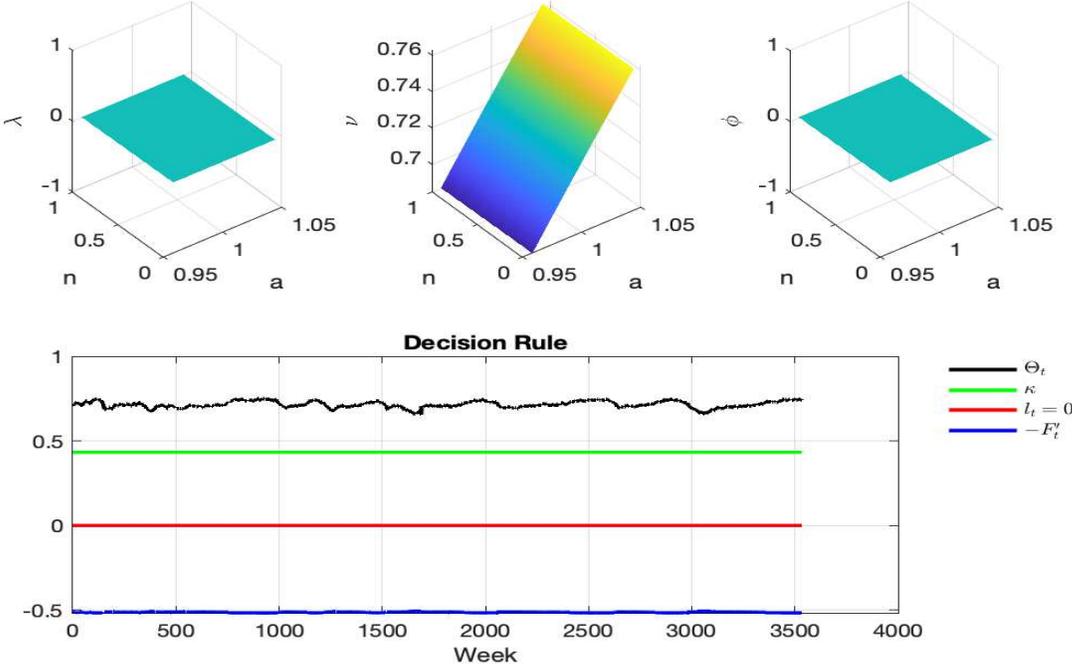
D.2 Regime (2/4): the "regime of labor hoarding"

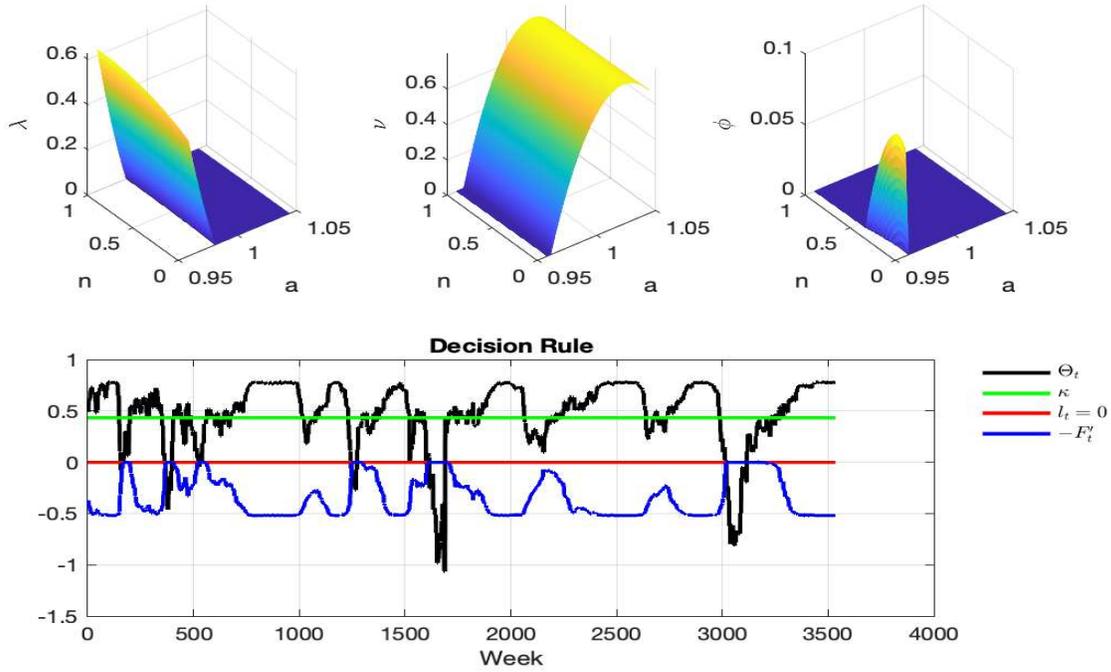If $0 < \Theta_\mu(a_t, n_t(\mu)) < \kappa$, then

$$\begin{cases} \lambda_t(\mu) \geq 0 & \Rightarrow & v_t(\mu) = 0 \\ \nu_t(\mu) \geq 0 & \Rightarrow & l_t(\mu) = 0 \\ \phi_t(\mu) = 0 & \Rightarrow & l_t(\mu) < 1 - s. \end{cases}$$

D.3 Regimes (3/4 and 4/4): "the regimes of firings and closures"

There are two cases:

$$
\begin{array}{c}
-F((1-s)n_t(\mu)) < \Theta_\mu(a_t, n_t(\mu)) < 0 \\
\lambda_t(\mu) \geq 0 \quad \Rightarrow \quad v_t(\mu) = 0 \\
\left. \begin{array}{rcl} \nu_t(\mu) &=& 0 \\ \phi_t(\mu) &=& 0 \end{array} \right\} \quad \Rightarrow \quad 0 < l_t(\mu) < (1-s)
\end{array}
\quad \middle| \quad
\begin{array}{c}
\Theta_\mu(a_t, n_t(\mu)) < 0 \\
\lambda_t(\mu) \geq 0 \quad \Rightarrow \quad v_t(\mu) = 0 \\
\left. \begin{array}{rcl} \nu_t(\mu) &=& 0 \\ \phi_t(\mu) &\geq& 0 \end{array} \right\} \quad \Rightarrow \quad l_t(\mu) = (1-s)
\end{array}
$$

E Smoothed cross-sections of the log-likelihood function

Figure 15 shows the original log-likelihood function in the neighborhood of each estimated value keeping constant the other parameters. As this function is not smooth, the variances cannot be calculated with a numerical evaluation of the Hessian matrix. We, therefore, replace the original log-likelihood function with its polynomial fit (in red).

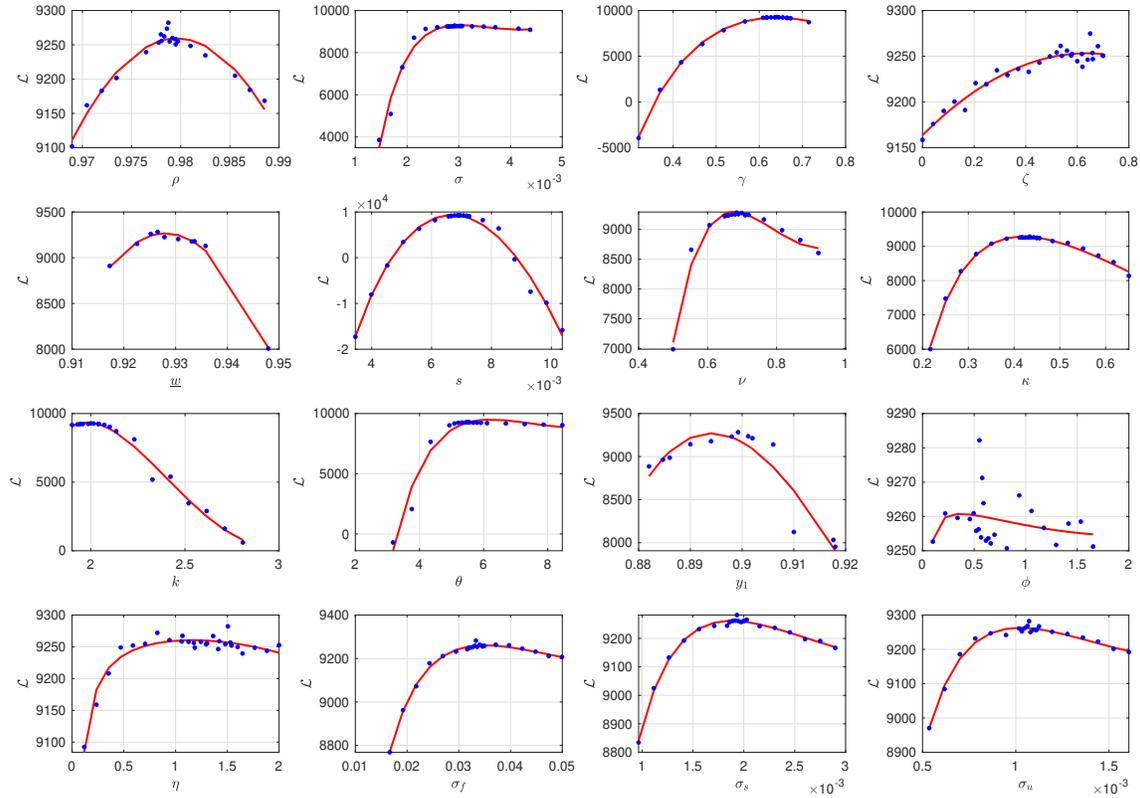Figure 15: Smoothed contour of the log-likelihood function

Table 7: Estimators and R^2 of the polynomial approximation of the likelihood

Parameter	Value	Std.	R^2	Parameter	Value	Std.	R^2
ρ	0.97874	0.00062	0.95519	k	2.00000	0.00375	0.98711
σ	0.00292	0.00003	0.97638	θ	5.50000	0.02612	0.95961
γ	0.64075	0.00266	0.99907	y_1	0.89923	0.00044	0.87029
ζ	0.65000	0.04903	0.91353	ϕ	0.55118	0.29744	0.09394
$\underline{w}$	0.92653	0.00040	0.99388	η	1.50536	0.13767	0.92491
s	0.00691	0.00002	0.98910	σ_f	0.03330	0.00088	0.99517
ν	0.68425	0.00391	0.97874	σ_s	0.00193	0.00005	0.99424
κ	0.43385	0.00401	0.99595	σ_u	0.00107	0.00004	0.98399

45

F Estimations of alternative models

F.1 Estimated parameters

Table 8: Parameter estimates

Param.	$M0$	$M1$	$M2$	$M3$	Interpretation
ρ	0.97874	0.97868	0.98867	0.98100	Persistence of the technology shock
σ	0.00292	0.00292	0.00298	0.00281	Std. innovations in the technology shock
γ	0.64074	0.64536	0.64075	0.69866	Bargaining power of workers
z_0	0.33311	0.78073	0.24832	0.25257	Opportunity cost of employment
ζ	0.65000	0.68250	0.68250	0.69500	Indexation of home production
$\underline{w}$	0.92653	-	0.92521	0.91547	Lower limit of bargained wages
s	0.00691	0.00597	0.00686	0.00691	Exogenous separation rate
ν	0.68425	0.83893	0.68913	0.69891	CES matching function
κ	0.43385	0.43385	0.44088	0.43695	Cost of a vacancy
k	2.00000	2.16154	2.00000	-	Param. 1, Beta dist., abilities
θ	5.50000	5.65763	5.44107	-	Param. 2, Beta dist., abilities
y_1	0.89922	0.90643	0.89923	0.90060	Lower bound of ability
ϕ	0.55117	-	-	1.29503	Scale parameter, firing cost function
η	1.50535	-	-	0.99916	Elasticity, firing cost function
σ_f	0.03329	0.03366	0.03306	0.03330	Std. measurement error on JFR
σ_s	0.00193	0.00177	0.00216	0.00193	Std. measurement error on JSR
σ_u	0.00107	0.00095	0.00119	0.00107	Std. measurement error on U
Log-Lik.	9282.18	9470.38	9128.36	9145.41	
Penalty	11.23	0.32	22.53	37.26	Losses due to constraint on the Log-Likelihood
$\mathbb{E}z/\mathbb{E}y$	0.75	0.90	0.76	0.75	Hall & Milgrom (2008)'s constraint

CPS monthly data, 1951M1–2018M12, 40,000 particles, and 30 abilities. Authors' calculations.
$M0$: benchmark model; $M1$: model without downward real wage rigidity; $M2$: model without firing costs; $M3$: model with a uniform distribution of abilities. The solution for the objective function used for the estimation is the log-likelihood minus the penalty (distances between data and models for the AR(1) process of the labor productivity).

Table 9: Productivity processes: data and models

	Data	$M0$	$M1$	$M2$	$M3$
Persistence ρ_p	0.954945	0.925191	0.947115	0.930345	0.917632
Variance σ_p^2	0.4247×10^{-4}	0.4935×10^{-4}	0.4349×10^{-4}	0.5279×10^{-4}	0.5560×10^{-4}

$M0$: benchmark model; $M1$: model without downward real wage rigidity; $M2$: model without firing costs; $M3$: model with a uniform distribution of abilities. $p = s, d$.
Data: BLS Quarterly data. Model: quarterly data are averages of the monthly data.

For models $M0$, $M2$, and $M3$, the value of z_0 is deduced from the constraint $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.75$. In the model without downward real wage rigidity, the constraint $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.75$ is not imposed (with this constraint, the values for $\log(a_t)$ are always negative), We then relax this restriction and estimate z_0. The estimation provides a larger value for z_0 and then $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.9$.

G Dynamics of separations

Before distinguishing between layoffs and quits, it is necessary to analyze the properties of the job separation rate generated by the model. Figure 16 shows that the aggregate separation rate spikes at the onset of economic downturns, which is consistent with the evidence found by Fujita & Ramey (2009) and Elsby et al. (2009). Figure 16 also provides a measure of the weight of endogenous separations in aggregate unemployment.

Figure 16: Job separation rate and unemployment
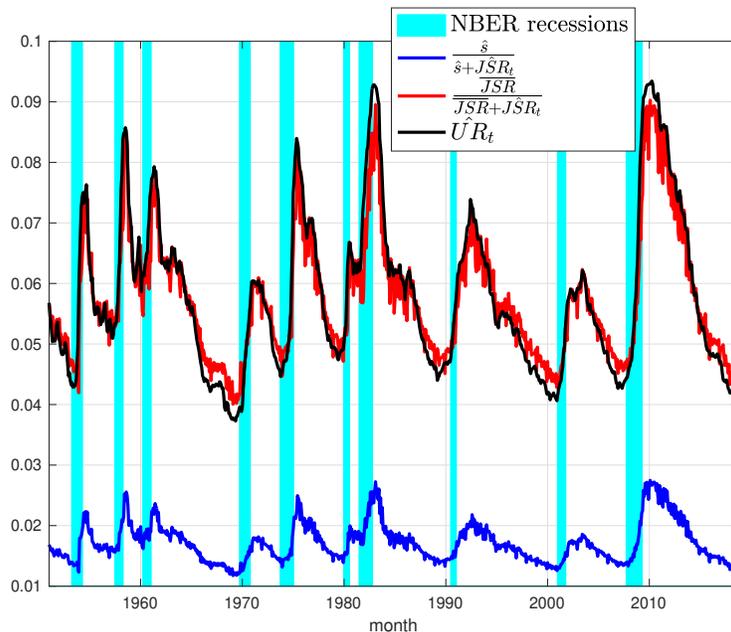
To this end, we construct the unemployment rate following Shimer (2012). In other words, its stationary expression uses the dynamics of worker flows (i.e., JFR_t and JSR_t lead to $UR_t = \frac{JSR_t}{JSR_t + JFR_t}$). We plot two counterfactual scenarios: (i), where only the estimated exogenous separation rate is considered (in this case, the unemployment rate is $\frac{\widehat{s}}{\widehat{s} + \widehat{JFR_t}}$) and (ii), where the separation rate is constant over the business cycle and equals the mean value of the model estimate (in this case, the unemployment rate is $\frac{\overline{JSR}}{\overline{JSR} + \widehat{JFR_t}}$). In case (i), the greatest proportion of the unemployment gap comes from a level effect. Here, the exogenous separation estimated by the model is a small proportion of the total separations. Thus, by omitting the separations driven by the business cycle, we underestimate unemployment by four percentage points on average. Case (ii) shows that even if we control for the level effect, we underestimate the increases in the unemployment rate during recessions by setting a constant separation rate fixed to the average of its estimated levels. Hence, the dynamics of the separation rate are non-negligible, confirming the results of Fujita & Ramey (2009) and Elsby et al. (2009).

H   Dynamics of the employment rates by ability

Figure 17 shows the employment rate dynamics in each labor market segment for $\mu = 35, 40, 45, 50, 55, 70, 150, 300, 500$. These employment rates by submarket are given as weekly time series. These panels provide counterparts in terms of the employment of the above-mentioned decision rules. Hence, for the lowest-ability workers, the probability of being employed is very low. By contrast, for abilities higher than $\mu = 55$, events that correspond to firm closures do not exist. These panels also illustrate the large decrease in the magnitude of unemployment fluctuations as workers' abilities increase.

Figure 17: Model implications for the employment rates of μ-type workers
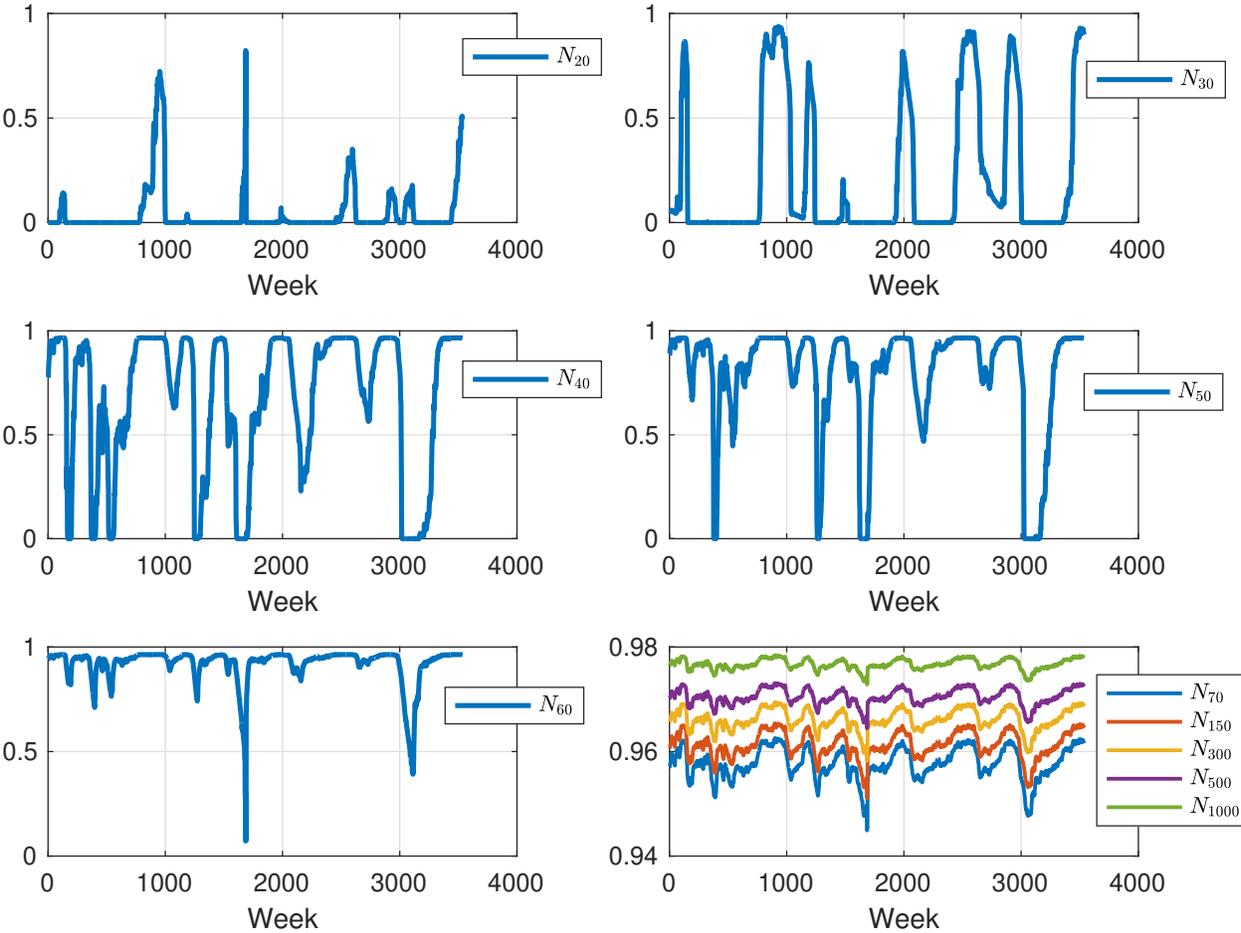
I Linearized version of the model

I.1 Model

We consider the following simplification of our benchmark model:

For $\mu \in [1, ..., S]$ with $\omega_\mu \sim \texttt{betapdf}(y_\mu - y_1, k, \theta)$

$$\frac{\kappa}{q(\theta_t(\mu))} = \beta\mathbb{E}\left[(1-\gamma)(y_\mu a_{t+1} - z_0 - \zeta(y_\mu - z_0)) - \gamma\kappa\theta_{t+1}(\mu) + (1-s_\mu)\frac{\kappa}{q(\theta_{t+1}(\mu))}\right]$$

$$u_{t+1}(\mu) = s_\mu(1 - u_t(\mu)) + [1 - \theta_t(\mu)q(\theta_t(\mu))]u_t(\mu)$$

$$q(\theta_t(\mu)) = \varphi\frac{1}{(\theta_t(\mu)^\nu + 1)^{1/\nu}}$$

Aggregates

$$\log(a_{t+1}) = \rho\log(a_t) + \epsilon_t \quad \text{with } \epsilon_t \sim N(0, \sigma^2)$$

$$UR_t = \sum_\mu \omega_\mu u_t(\mu)$$

$$JFR_t = \frac{\sum_\mu \omega_\mu\theta_t(\mu)q(\theta_t(\mu))u_t(\mu)}{\sum_\mu \omega_\mu u_t(\mu)}$$

$$JSR_t = \sum_\mu \omega_\mu s_\mu$$

where the job separation rates (s_μ) are assumed to be constant across abilities because they would not be identified (we have only one time series for separations, it is possible to identify only one separation rate $s = s_\mu, \forall\mu$).

Therefore, additional constraints must be introduced. First, we restrict the parameters to satisfy $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.75$, as in our benchmark model and in the range of empirical evidences provided by Hall & Milgrom (2008), implying that $z_0 = \frac{0.75-\zeta}{1-\zeta}\left(y_1 + \frac{k}{k+\theta}\right)$. Second, given that the linearized version of the model must have an interior solution, we must have $y_1 > z_0$, leading to $y_1 > \frac{0.75-\zeta}{1-0.75}\frac{k}{k+\theta}$, given the previous expression of z_0. Therefore, we impose $y_1 = \varpi + \frac{0.75-\zeta}{1-0.75}\frac{k}{k+\theta}$ with $\varpi = 1.075$. The value of ϖ is chosen such that y_1 value is approximately the same in the nonlinear and linear estimations. Finally, we calibrate the distribution of abilities as in the nonlinear model, setting $k = 2$ and $\theta = 5.5$. The estimated parameters are

$$\Theta = \{\kappa, \nu, \gamma, \zeta, s, \rho, \sigma, \sigma_u, \sigma_f\} \quad \text{with } dim(\Theta) = 9$$

given that $\varphi = 1$ and $\beta = 1/((1 + 0.00415)^{12})$ because $(1 + 0.00415)^{12} \approx 1.05$. For the nonlinear model, the data are $Y_t = \{UR_t, JFR_t, JSR_t\}$. We use Dynare (see Adjemian et al. (2011)) to solve (First-order perturbation method) and estimate the model by maximum likelihood.

I.2 Results

The estimated parameters are shown in the first two columns of Table 10. The likelihood is

Table 10: Estimated parameters: linearized model

| | $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.75$ | | $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.95$ | |
	mode	s.d.	mode	s.d.
ρ	0.9648	0.00523	0.9648	0.00472
σ	0.0296	0.00948	0.0069	0.00281
γ	0.3408	0.21586	0.4216	0.12469
ζ	0.9189	0.11036	0.9829	0.03012
s	0.0250	0.00008	0.0250	0.00009
ν	0.7462	0.10621	0.6146	0.06826
κ	0.5648	0.23017	0.0561	0.02633
σ_u	0.0065	0.00020	0.0065	0.00017
σ_s	0.0029	0.00008	0.0029	0.00008
Log-Lik	8178.02		8177.99	

CPS monthly data, 1951M1–2018M12.
Maximum likelihood estimation with Dynare and 10 abilities

noticeably lower than in the estimation of the nonlinear model, but this was expected as the linearized model cannot account for the variability of the job separation rate. With respect to the nonlinear model, the parameters of the matching function (ν) and the wage bargaining power (γ) are not very different, whereas the vacancy costs (κ) are significantly greater. Given that it is not possible to introduce a downward real wage rigidity in a linearized model, the opportunity cost of employment is different because it must encompass all wage rigidities. However, the important result is that the estimation based on the linearized version of the model requires an aggregate shock having a variance 62 times larger than in the estimation based on the nonlinear model ($\sigma^2/(1 - \rho^2)$). Therefore, the linearized version of the model cannot solve Shimer's puzzle (the possibility of generating the size of the labor market fluctuations with a shock having a size close to the variance of the US labor productivity). By imposing an opportunity cost of employment equal to 75% of the average productivity, wage rigidity is too low to generate large adjustments in quantities.

Table 11: Estimated parameters: linearized model

Moments	Linear model with heterogeneity	Linear model with a single market
$\mathbb{E}[UR]$	0.065756	0.065695
$\mathbb{E}[JFR]$	0.354979	0.355330
$\mathbb{E}[JSR]$	0.024985	0.024985
$\mathbb{V}[UR]$	2.636×10^{-4}	2.663×10^{-4}
$\mathbb{V}[JFR]$	9.100×10^{-3}	9.229×10^{-3}

The hyper-parameters of the distribution of abilities are calibrated using the estimates from the nonlinear model ($M0$). The attempts to estimate the distribution of abilities in

this approximated model all ended up into degenerate distributions (zero variance), suggesting that heterogeneity does not matter significantly (w.r.t. the three observed variables we consider) if we abstract from the nonlinearities. To illustrate this, Table 11 compares the theoretical moments of the model with heterogeneity and the corresponding representative agent model. All first-and second-order moments are very close, suggesting that heterogeneity can not be identified without the nonlinearities of the model.

Another important shortcoming of the model evaluation based on its linearized version comes from the following result: in the absence of a constraint on positive vacancies, the stochastic simulations of the linearized model for vacancies display negative values with a probability ranging from 0.3% to 2.7% across sectors (with a median of 1%), the assumption that an interior solution is not necessarily satisfied at each period.[20] Given that this constraint is binding in the nonlinear version of the model, this result is not surprising but casts some doubt on the robustness of evaluations based on the linearized versions of DMP model.

Alternative calibration.   As is often proposed in the literature on the Shimer's puzzle, it is possible to reduce the size of the exogenous shock by calibrating the opportunity cost of employment to a higher value than that proposed in Hall & Milgrom (2008), which can be viewed as the Hagendorn & Manovskii (2008)'s solution. The last two columns of Table 10 present the re-estimated values for the parameters when $\mathbb{E}[z(\mu)]/\mathbb{E}[y_\mu] = 0.95$. As expected, the variance of the macro shock decreased sharply because a lower fundamental surplus implies a higher elasticity of the model. However, the steady-state solution of the model is also closer to the frontier $v_t(\mu) \geq 0$. Therefore, a simulation over a large number of periods indicate that the possibility of observing a negative vacancy rate ranges between 0.9% and 4.6% (with a median of 2.05%). This result reinforces the importance of considering nonlinearities when evaluating the DMP model.

---

[20]Simulations are based on a model with a finer grid of abilities, over 100 levels