ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/130117/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Pizzinelli, Carlo, Theodoridis, Konstantinos and Zanetti, Francesco 2020. State dependence in labor market fluctuations. International Economic Review 61 (3) , pp. 1027-1072. 10.1111/iere.12448

Publishers page: http://dx.doi.org/10.1111/iere.12448

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

State Dependence in Labor Market Fluctuations*

Carlo Pizzinelli
International Monetary Fund

Konstantinos Theodoridis
Cardiff Business School
European Stability Mechanism

Francesco Zanetti
University of Oxford

February 26, 2020

Abstract

This paper documents state dependence in labor market fluctuations. Using a Threshold Vector Autoregression model (TVAR), we establish that the unemployment rate, the job separation rate, and the job finding rate exhibit a larger response to productivity shocks during periods with low aggregate productivity. A Diamond-Mortensen-Pissarides model with endogenous job separation and on-the-job search replicates these empirical regularities well. We calibrate the model to match the standard deviation of the job-transition rates explained by productivity shocks in the TVAR, and show that the model explains 88 percent of the state dependence in the unemployment rate, 76 percent for the separation rate and 36 percent for the job finding rate. The key channel underpinning state dependence in both job separation and job finding rates is the interaction of the firm's reservation productivity level and the distribution of match-specific idiosyncratic productivity. Results are robust across several variations to the baseline model.

JEL Classification: E24, E32, J64, C11.
Keywords: Search and Matching Models, State Dependence in Business Cycles, Threshold Vector Autoregression.

*Please address correspondence to Francesco Zanetti, University of Oxford, Department of Economics, Manor Road, Oxford, OX1 3UQ, UK; Email: Francesco.Zanetti@economics.ox.ac.uk. This study subsumes previous work circulated under the title: "State Dependence in Labor Market Fluctuations: Evidence, Theory and Policy Implications." We would like to thank participants at several seminars and workshops for extremely valuable comments and suggestions. Francesco Zanetti would like to thank the Bank of Japan and its staff for hospitality and support while some of this work was completed. The present project was partially supported by the John Fell Fund, and the British Academy and Leverhulme Research Grants. The views expressed in this paper solely represent those of the authors and should not be interpreted as the views of the International Monetary Fund or its Board, and the European Stability Mechanism.

1 Introduction

Numerous studies, starting with Neftci (1984), show that macroeconomic fluctuations differ across phases of the business cycle. This paper builds on this strand of research and identifies systematic changes in the cyclical properties of labor market variables that are linked to the state of aggregate productivity. A Threshold Vector Autoregression (TVAR) model, which identifies the effect of productivity shocks and allows for two distinct regimes based on aggregate productivity, establishes that the shocks have a significantly larger effect on the unemployment rate, job separation rate, and job finding rate in periods of low aggregate productivity.

To explain these findings, we develop a Diamond-Mortensen-Pissarides (DMP) search model with endogenous job separation and on-the-job search (OJS) where workers have different idiosyncratic productivity. The model embeds two channels that can generate state dependence. The first channel arises from the interaction between the firm's reservation threshold for match-specific productivity and the distribution of workers' individual productivity. The reservation threshold determines whether incumbent workers are dismissed (i.e. job separation) and whether newly matched workers are hired (i.e. job finding). Under standard assumptions, in states of low aggregate productivity, the threshold lies in a region of the match-specific distribution with a high density of workers. An exogenous movement in aggregate productivity that changes the threshold generates large adjustments in the separation rate and the job finding rate and consequently in the unemployment rate. Meanwhile, in a state with high aggregate productivity, the firm sets a low threshold for efficiency matches that is associated with a low density of jobs. An equivalent change in aggregate productivity produces limited movements in the job-transition rates.

The second channel hinges on the nonlinearity in the matching function, which determines the firm's hiring intensity. In states of low aggregate productivity, the surplus of new matches is small and the firm's vacancy postings become more responsive to shocks, thereby increasing the volatility in the job finding rate.

To assess the performance of the theoretical model to replicate the empirical findings and study the contribution of each channel to the state dependence of labor market variables, we

calibrate the system to match the variance of labor market variables explained by productivity shocks in the TVAR. The baseline model replicates closely the observed business cycle fluctuations, and the job separation rates and job finding rates exhibit larger responses to productivity shocks in states of low aggregate productivity. The model replicates 88, 76, and 36 percent of the state dependence in the unemployment rate, job separation rate, and the job finding rate, respectively. The response of the job transition probabilities to shocks imply that the volatility of the unemployment rate is almost twice as large in states of low aggregate productivity. We find that the key mechanism for the state dependence in the job finding rate is related to the changes in the productivity threshold for new matches rather than the nonlinearity in the matching function. Our analysis shows that the state dependence in labor market variables is primarily related to the interaction between the firm's reservation threshold for match-specific productivity and the distribution of workers' individual productivity.

We perform robustness analysis that considers alternative calibrations and specifications of the baseline model that exclude OJS and assumes exogenous job separation, which shows that the calibration of the idiosyncratic productivity distribution is critical to generate plausible aggregate fluctuations and replicate the observed state dependence in labor market variables. Under our baseline calibration, OJS is important to generate significant state dependence in the job finding rate. When OJS is possible, low-productivity workers search for alternative posts that yield higher wages and hence have a high probability of leaving the current firm. Thus, hiring these workers is more costly in presence of OJS since the worker is more likely to separate from the job and the firm must rehire another worker. Therefore, the firm sets a high reservation productivity threshold to compensate for the increase in expected costs. The higher threshold is located in a region of the match-specific productivity distribution with a high density of workers. Therefore, movements in aggregate productivity have a larger impact on the job finding rate. If OJS is not possible, however, the firm's threshold of reservation productivity is low and the degree of state dependence is substantially weakened. In this case, the firm hires workers with lower productivity who remain in the same job and are likely to acquire a higher level of productivity in subsequent periods, which increases the joint surplus of the match.

Our analysis relates to empirical and theoretical studies on the asymmetry of labor market fluctuations over the business cycle. On the empirical side, the works by Neftci (1984), Altissimo and Violante (2001), Panagiotidis and Pelloni (2007), Barnichon (2012), Abbritti and Fahr (2013), Barattieri et al. (2014), Caggiano et al. (2014), and Benigno et al. (2015) show that unemployment and wages fluctuate differently across phases of the business cycles. Compared to these studies, we show that state dependence in labor market fluctuations is linked to the level of aggregate productivity, and we extend the analysis to job transition rates. A large body of research has identified specific historical periods in which business cycle volatility has changed (e.g. the Great Moderation), attributing those changes to the conduct of monetary policy in the Volcker era or the Zero Lower Bound period (e.g. Liu et al., 2018), or the magnitude of exogenous shocks (e.g. Justiniano and Primiceri, 2008). While focusing on the labor market alone, our empirical analysis uncovers changes in the responses of macroeconomic variables to shocks that occur systematically throughout the business cycle since the post-WWII era rather than specific periods.

On the theoretical side, our work is related to studies that develop structural models to investigate asymmetric dynamics of the labor market. Sedláček (2014) and Kohlbrecher and Merkl (2016) consider the importance of the reservation productivity threshold for the dynamics of the job finding rate. Unlike these studies, we investigate how this mechanism contributes to state dependence in both the job finding and job separation rates, and the TVAR model identifies as important empirical sources of state dependence for the unemployment rate. Ferraro (2018) shows that employment cycles are characterized by large skewness and develops a search model with permanent worker heterogeneity in productivity to explain the finding. Our version of the DMP model hinges on heterogeneity in match-specific productivity, which allows for job-to-job transitions, building on the work of Fujita and Ramey (2012), who show that OJS is critical for the standard DMP model to deliver a realistic performance and generate the Beveridge Curve. Petrosky-Nadeau and Zhang (2017) show that a standard DMP model with exogenous job separation generates state dependence via the inherent nonlinearities of the policy function for market tightness. Although our version allows for this channel, we find it to be less quantitatively important, under our calibration, to explain the state dependence of

labor market variables at business cycle frequency identified by the TVAR model.

The remainder of the paper is structured as follows. Section 2 presents the empirical findings. Sections 3 and 4 outline the benchmark model and discuss the main mechanisms generating state dependence in labor market fluctuations, respectively. Section 5 describes the calibration and presents the main results. Section 6 performs a series of robustness checks on the calibration and alternative specifications of the model. Section 7 concludes.

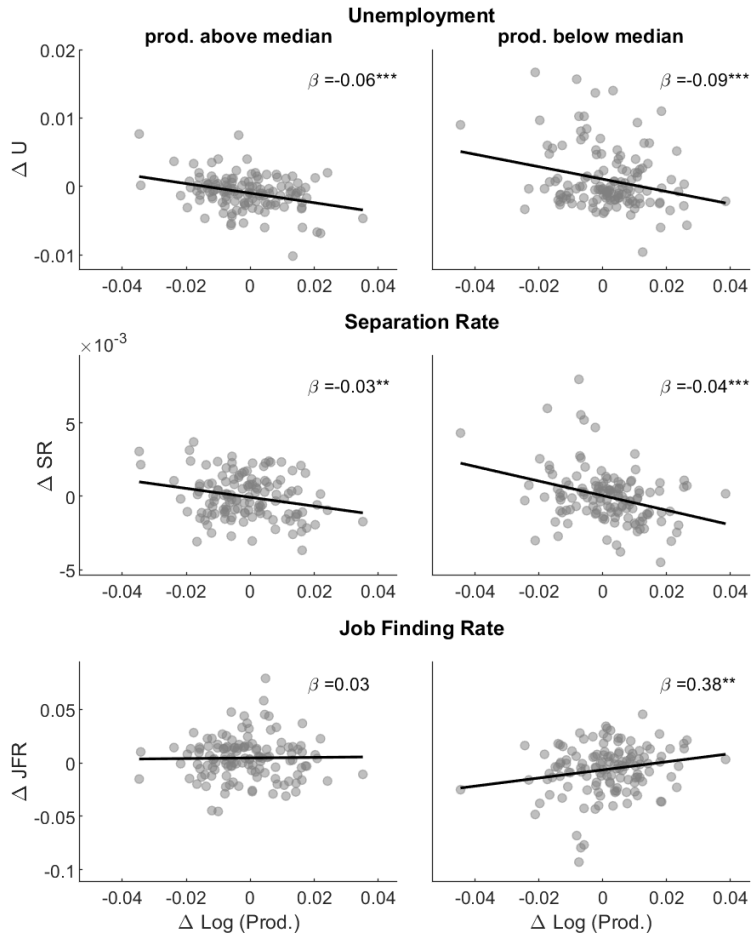
2 Empirical evidence

This section isolates systematic differences in fluctuations of labor market variables linked to the state of aggregate productivity. Comovements of unemployment and job-transition rates with average labor productivity are stronger in periods of low aggregate productivity, resulting in a larger volatility of labor market variables. Before studying these dynamics with a structural TVAR model, we use a simple scatter plot to show the changes in the comovements of these variables over the business cycle.

We use quarterly series for the unemployment rate, the job finding rate, the job separation rate, output, hours, and labor productivity over the period 1950:I-2014:IV. To extract the cyclical component of variables, we use the filtering method in Hamilton (2018).[1] Figure 1 plots quarterly first differences for the unemployment rate, job separation rate, and job finding rate against the quarterly growth rate of labor productivity for periods in which the level of productivity is above (left panels) and below (right panels) the median value. The elasticity coefficients are larger in periods of low productivity, suggesting that the comovement between changes in labor market variables and changes in productivity is stronger in periods of low productivity. In the next subsection, we assess this finding more formally through a TVAR model that isolates the response of variables to shocks across different regimes of productivity.

[1] Appendix A provides details on the data sources.

Figure 1: Labor market fluctuations and labor productivity across phases of the business cycle.



Note. The figures plot quarterly changes of the unemployment rate, the separation rate, and the job finding rate against growth rates of labor productivity (log first differences). The left and right panels consider periods in which the starting level of productivity is respectively above and below the historical median of its cyclical component. The black solid line represents the best fit from a least squares regression, with the slope coefficient reported in each plot. The β is the estimated coefficients from a linear regression, where productivity is the explanatory variable. * p-value<0.1; ** p-value<0.05; *** p-value<0.01.

2.1 The Threshold Vector Autoregression model

The TVAR model, based on the original study by Chen and Lee (1995), allows the VAR parameters to vary across an aggregate state of the economy. The switching mechanism is based on the value of one of the endogenous variables being above or below a threshold, and unlike Markov-Switching models, the parameter change is endogenous to the dynamics of the VAR process. The reduced-form model can be expressed as follows:

$$Z_t = \xi_t \left\{ c_1 + \sum_{k=1}^{K} B_{k,1} Z_{t-k} + \Sigma_1^{1/2} v_t \right\} + (1 - \xi_t) \left\{ c_2 + \sum_{k=1}^{K} B_{k,2} Z_{t-k} + \Sigma_2^{1/2} v_t \right\}, \qquad (1)$$

where Z_t is the vector of N observed variables, c_1 and c_2 are constant coefficients, $B_{k,1}$ and $B_{k,2}$ are coefficients of the VAR, $\Sigma_{k,1}$ and $\Sigma_{k,2}$ are covariance matrices, and v_t is the error term. Switches across regimes are governed by the indicator variable $\xi_t \in \{0, 1\}$, which is equal to 1 if labor productivity in period $t-1$, $\tilde{z}_{t-1}$, is below the threshold z^*, and equal to zero otherwise:

$$\xi_t = 1 \quad \text{if} \quad \tilde{z}_{t-1} < z^*, \quad \text{otherwise} \quad \xi_t = 0. \tag{2}$$

Under conjugate priors for VAR parameters and conditional on the value of the threshold z^*, the posterior distribution of the VAR coefficient vector is a conditional Normal-Wishart distribution, and we obtain draws with the Gibbs sampler. Since the posterior distribution of z^* conditional on the VAR parameters is unknown, we use a Metropolis-Hastings step to obtain the posterior distribution (see Chen and Lee (1995); Chen (1998); Lopes and Salazar (2006)). Appendix B.1 provides details on priors.

The variables of interest are labor productivity, unemployment rate, job separation rate, and job finding rate over the period 1950-2014. To be consistent with the related literature, particularly Balleer (2012) and Canova et al. (2012), we include average hours worked and use Fernald's measure of labor productivity. We use eight lags in the baseline specification (i.e., $k = 8$) and select the one-quarter lag of labor productivity as the variable determining the state ξ_t.[2] Productivity and hours are entered in logs and are detrended using the method proposed by Hamilton (2018). Being stationary, the unemployment, separation, and job finding rates are not detrended.[3] The median of the posterior of the threshold z^* corresponds to the 58th percentile of the unconditional distribution of detrended log productivity, which is close to the median.[4]

To identify productivity shocks, we follow the medium-run maximum variance scheme pro-

[2]We set the number of lags in the benchmark model equal to eight to facilitate comparison of the benchmark identification scheme with the identification that uses long-run restrictions, as in Balleer (2012) and Canova et al. (2012) (results are reported in the online appendix). As established in those studies and further discussed in Erceg et al. (2005) and Ravenna (2007), a sufficiently large number of lags is needed to mitigate identification bias that arises from the (lag) truncation of the VAR. The online appendix shows that results continue to hold with fewer lags.

[3]The results are robust to excluding hours and to de-trending all variables.

[4]Figure B.1 in the Appendix plots prior and posterior distributions for the threshold z^*. Moreover, Figure B.2 reports the time series of the identified regimes using the median of the posterior.

posed by Uhlig (2004) and assume that the productivity shock explains the majority of the forecast error variance of labor productivity at business cycle frequencies (i.e. over the horizon of 0 to 40 quarters). Appendix B.2 provides details on the identification scheme. Unlike Canova et al. (2012), our VAR does not include the relative cost of investment as a variable. Therefore, the shocks we identify are general innovations to labor productivity, making no distinction between neutral technology shocks and investment-specific shocks. This choice is motivated by the need to only use variables that are present in the baseline DMP model discussed below, which does not contain capital.

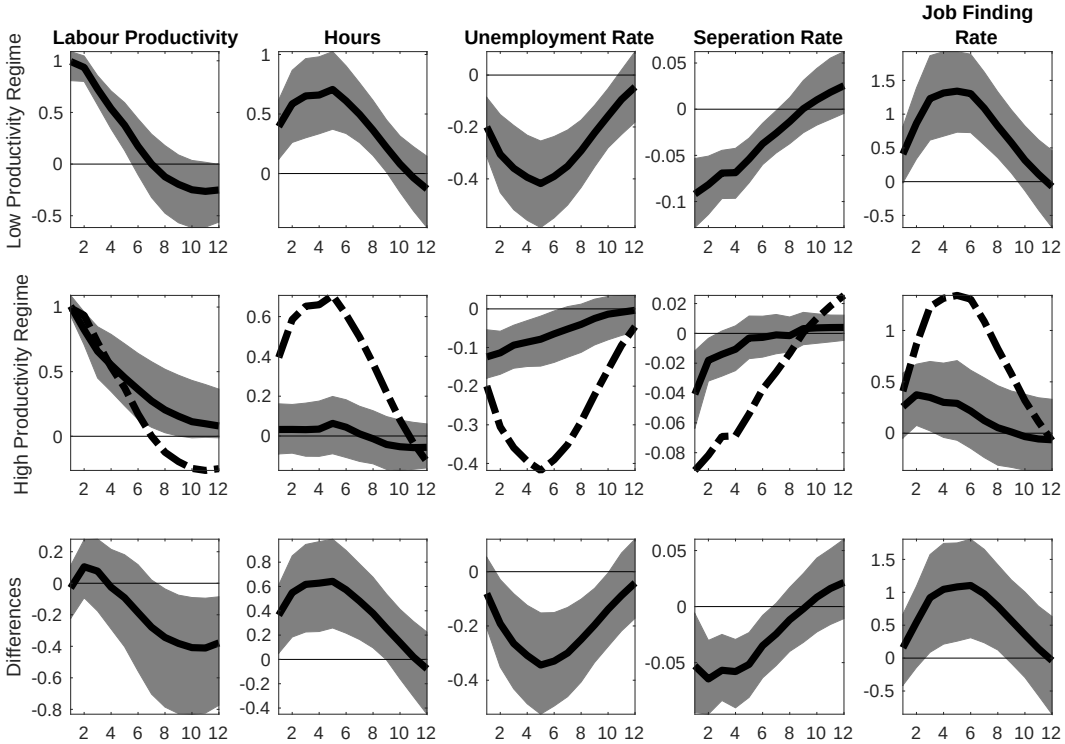
To assess whether responses to productivity shocks are significantly different across states, Figure 2 separately plots the linear Impulse Response Functions (IRFs) for each regime, reporting the median and the 16th and 84th percentiles of the posterior distribution. The responses of the unemployment rate and the two transition rates to a productivity shock when labor productivity is below the threshold (first row) are more than twice as large as in the regime with productivity above the threshold (second row), and the differences are statistically significant (third row). Interestingly, the response of labor productivity is very similar in the two regimes, suggesting that the larger response in labor market variables is not generated by the underlying volatility of the shock but by a stronger response of the variables to exogenous disturbances. Overall, the signs of the IRFs of all variables are in line with those in a broad class of DMP models in response to labor productivity shocks, as in Mumtaz and Zanetti (2015).[5]

Table 1 reports the share of the variance of the forecast error explained by productivity shocks at one-, four-, and eight-quarter horizons (Columns (1)-(3)) and at the stationary long-run horizon (i.e., $h \to \infty$, Column (4)). In the low-productivity regime (top table), productivity shocks explain a large fraction of unexpected fluctuations in labor market variables compared to the high-productivity regime (bottom table). To better relate this finding to the historical series, Columns (5) and (6) compare the unconditional standard deviations in the data against the standard deviation explained by the productivity shocks from the TVAR.[6] Column (5)

[5] Our results differ from those of Canova et al. (2012), who, accounting for capital-specific productivity, find that neutral technology shocks fuel job destruction.

[6] To calculate the conditional standard deviation of a variable we multiply its variance in a given regime by the fraction explained by productivity at the stationary horizon and then take the square root.

Figure 2: Benchmark Model: Impulse Responses



Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, and vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The dashed line in the second row is the pointwise median from the low-productivity regime.

shows that the unconditional standard deviations of the variables are generally slightly higher in the low-productivity regime, including productivity itself. Column (6) shows that the conditional volatility of labor market variables is markedly higher in the low-productivity state. The conditional standard deviations of productivity are closer than the unconditional standard deviations for both regimes. Importantly, this result is consistent with the theoretical model we develop in the next section, where the magnitude of shocks is the same across the business cycle.

Figure 3 plots the historical decomposition for each series. The black line reports the historical time series (demeaned), while the grey bars represent the value of the respective variable explained by productivity shocks. The shocks have an asymmetric influence on both transition rates over the business cycle. Except for the 2001 recession, the shocks explain a

Table 1: Forecast error variance decomposition (FEVD) of the baseline TVAR model and standard deviation of the historical series in the low- and high-productivity regimes.

Low Productivity (<58th percentile)						
	TVAR: FEVD				Data: standard deviation	
	(1)	(2)	(3)	(4)	(5)	(6)
Quarters	h=1	h=4	h=8	$h \to \infty$	σ	σ|TVAR
Productivity	78.4%	82.7%	67.2%	64.4%	2.01	1.61
Hours	15.1%	34.9%	41.2%	43.4%	0.67	0.44
Unemployment	27.4%	48.0%	58.2%	55.9%	1.63	1.22
Separation Rate	39.5%	44.6%	42.2%	43.9%	0.60	0.40
Job Finding Rate	5.3%	25.7%	34.6%	38.1%	8.12	5.01

High Productivity (>58th percentile)						
	TVAR: FEVD				Data: standard deviation	
	(1)	(2)	(3)	(4)	(5)	(6)
Quarters	h=1	h=4	h=8	$h \to \infty$	σ	σ|TVAR
Productivity	96.2%	96.0%	93.1%	88.5%	1.33	1.25
Hours	0.9%	2.1%	3.7%	6.8%	0.42	0.11
Unemployment	9.0%	11.5%	13.4%	15.6%	1.51	0.60
Separation Rate	7.0%	8.7%	10.1%	12.7%	0.49	0.17
Job Finding Rate	1.3%	3.6%	5.1%	7.9%	8.88	2.49

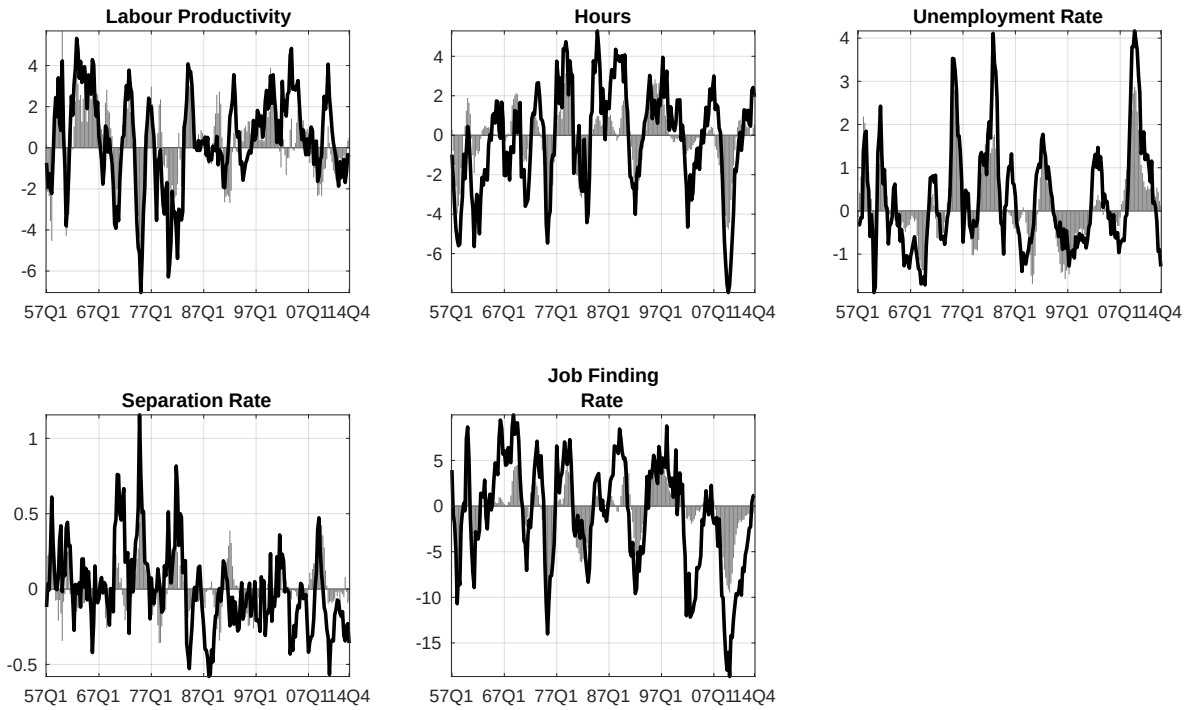
Note. Columns (1)-(4) report the fraction of the forecast error variance of each variable that is explained by the identified productivity shocks at a time horizon of h quarters. Column (5) reports the unconditional standard deviation (σ) of each variable when de-trended log productivity is below (top table) or above (bottom table) the 58th percentile, which is the mean of the posterior distribution of the TVAR threshold. Column (6) reports the standard deviation of the fluctuations of each variable that is explained by the TVAR at the stationary long-run horizon ($h \to \infty$). This is computed through multiplying Column (4) by the square of Column (5) (i.e. the variance) and then taking the square root. Both Columns (5) and (6) are then multiplied by 100 for presentation purposes.

larger fraction of fluctuations in the separation rate (job finding rate) when the historical series is above (below) the mean value. Important for our analysis, the productivity shocks have a symmetric influence on labor productivity itself over expansionary and contractionary phases of the business cycle.

In the Appendix, we examine in more detail the individual role of the job separation and job finding rates for the dynamics of the unemployment rate in the TVAR. In particular, using the steady-state approximation of the unemployment rate proposed by Shimer (2012), Appendix B.5 shows that the separation rate explains roughly one-third of the fluctuations of the unemployment rate driven by productivity shocks over the cycle. An additional question is how much of state dependence in unemployment is driven by each transition rate. To assess this issue in the data, in Appendix B.6, we combine the steady-state approximation with the

IRFs from Figure 2. We show that the separation rate explains a large fraction of the state dependence of the unemployment rate in the initial quarters after the productivity shock, and the relevance of the job finding rate increases after 4 quarters.[7]

Figure 3: Benchmark TVAR model: historical decomposition.



Note. The solid line represents the demeaned time series. The grey bars represent the part of the series explained by the productivity shock.

2.2 Generalized Impulse Response Functions

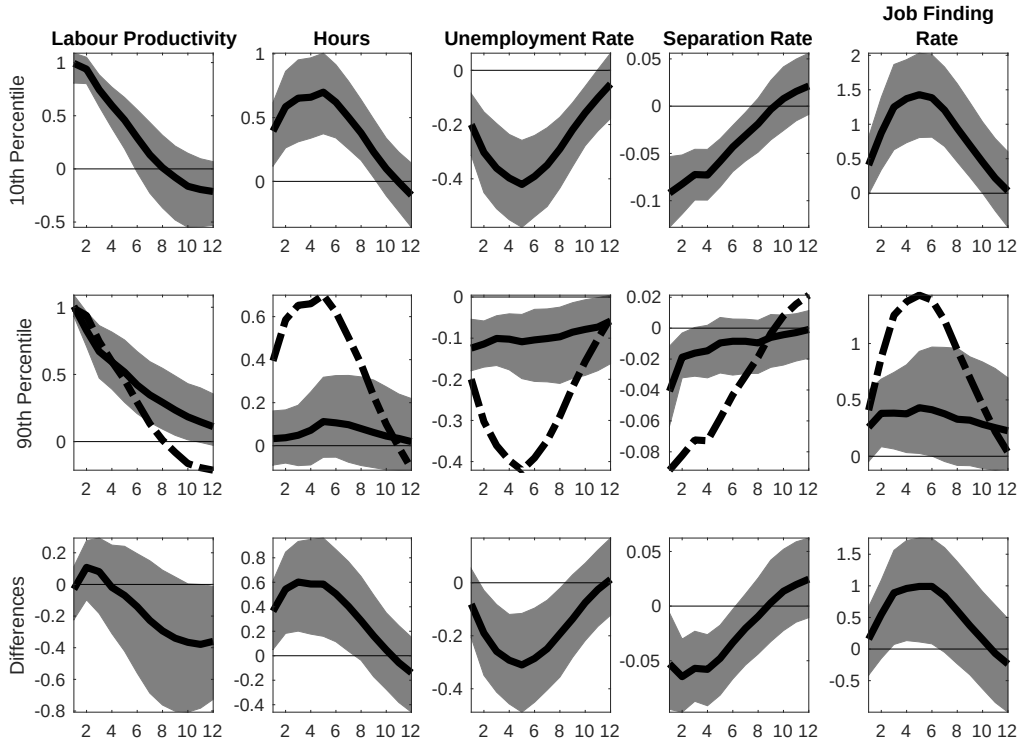
The IRFs in Figure 2 are obtained using regime-specific coefficient matrices, which assume that the system remains in the current regime permanently. To ensure that results are robust to the possibility for the system to switch across regimes of productivity, we compute simulation-based Generalized IRFs (GIRFs), as developed in Koop et al. (1996).[8] The approach computes IRFs at different points of the business cycle, accounting for potential transitions across regimes

[7] In the quarter in which the shock occurs, the differential response of the separation rate in the low-productivity regime relative to the high-productivity regime accounts for 80 percent of the difference in the response of the unemployment rate across states. Four quarters after the shock, however, the job finding rate explains approximately 60 percent of the state dependence in the unemployment rate, showing that the relevance of each transition rate to account for the state dependence in the unemployment rate differs across time horizons.

[8] Appendix B.3 describes the computation of GIRFs.

that may influence the dynamic response of variables to the productivity shock.

Figure 4: Generalized Impulse Responses: Low- versus High-Productivity Starting Point



Note. The solid line represents the pointwise median GIRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. The horizontal axes are in quarters; the vertical axes are in percentage deviations from the trend. The first row reports the GIRFs from a starting point in which productivity is below the 10th percentile of its unconditional distribution. The second row reports the GIRFs from a starting point in which productivity is above the 90th percentile of its unconditional distribution. The third row displays the posterior distribution of the difference between the two. The dashed line in the second row is the pointwise median from the first row.

Figure 4 compares GIRFs at two different starting points of the business cycle: a low-productivity initial state with productivity below the 10th percentile of the productivity distribution, and a high-productivity one with productivity above the 90th percentile of the productivity distribution. In contrast to the responses in Figure 2, the reaction of variables to productivity shocks in Figure 4 accounts for the possibility that the current regime may change depending on the size and sign of the shock. At the peak of the response (roughly four to five quarters), the reaction of the unemployment rate is more than twice as large as in the low-productivity regime compared to the high-productivity regime. Similarly, peak responses of the job separation rate and job finding rate are also at least twice as large in the former case.

Overall, these findings establish that fluctuations in the unemployment rate, the job separation rate, and the job finding rate exhibit statistically different responses to technology shocks across states of aggregate productivity.

2.3 Robustness analysis

This section shows that the results continue to hold across several variations to the benchmark model. In particular, we consider the following changes: (i) looser priors, (ii) shorter number of lags (i.e. 4 instead of 8), and (iii) using a series of labor productivity not adjusted for capital utilization. We also check that state dependence is robust to two specific sub-periods: (i) the pre-Great Recession period 1950-2007 and (ii) the post-Volker period 1980-2014. Appendix B reports these robustness results.

We also investigate to what extent results hold if we use an alternative econometric approach and estimate a Markov-Switching VAR (MS-VAR), using the methodology in Sims and Zha (2006). The key difference between the TVAR and the MS-VAR is the formulation of transitions across regimes. The TVAR implies that, based on perfect observation of the Data Generating Process, these transitions are systematic and deterministic for a given value of the threshold variable. The MS-VAR implies that these transitions are stochastic and occur according to a Markov process, allowing flexibility in the identification of the two states. The MS-VAR constitutes an agnostic approach on the likelihood of different economic regimes and therefore provides a valuable alternative assessment on the state dependence of labor market variables. Appendix C presents the specification and estimation of the MS-VAR. We find that the MS-VAR corroborates the results of the TVAR, and the responses of labor market variables are stronger in states of low aggregate productivity. The responses of the variables to productivity shocks in the high-productivity regime in the MS-VAR model entail a larger degree of noise than the responses in the TVAR model, revealing greater uncertainty in the relation between productivity and labor market variables in this state of the economy. Meanwhile, the IRFs from the low-productivity regime have a tight posterior distribution, suggesting a stronger relationship between productivity and unemployment fluctuations.

Figure C.1 shows that the posterior probability of the low-productivity state is closely

aligned to the low-productivity periods estimated in the TVAR. A key difference in the results of the TVAR is that the response of productivity to the shock is also different across regimes, being smaller in the high-productivity regime. This difference results from the fact that in the MS-VAR, the regime transitions and the identification of the shocks to productivity are not as intrinsically related as in the TVAR.[9]

Finally, we assess the robustness of the results to our identification scheme for the productivity shock. In the Online Appendix, we identify productivity shocks combining long-run (LR) restrictions on the variance-covariance matrix of the forecast errors with sign restrictions on the response of labor market variables that are consistent with the baseline DMP model. We find that this alternative identification scheme generates larger responses of labor market variables in the low-productivity regime, as in the benchmark model.

3 The model

This section lays out a DMP search and matching model with endogenous job separation and OJS. The main features of the model are similar to those in Mortensen and Pissarides (1994), Merz (1999), Krause and Lubik (2007), Thomas and Zanetti (2009), Fujita and Ramey (2012), Sedláček (2014), and Mueller (2017).

Economic environment and timing. A continuum of households of mass one and a continuum of firms operate in a discrete time environment. Households supply labor to firms inelastically. Matching frictions in the labor market prevent full employment. Firms pay a fixed cost for each vacancy posted to recruit new workers. However, in a given period, neither are all vacancies filled nor are all job-seekers hired. Employed workers produce a single consumption good, whose price is normalized to one, and they may search for a new job while employed. In every period t, production by a single worker depends on aggregate labor productivity, a_t,

[9] The switch in the productivity regime is a random event in the MS-VAR and therefore is not necessarily the result of a large shock. Conversely, in the TVAR the switch from the high- to the low-productivity regime occurs when productivity falls below a threshold, which by construction has to be driven by a shock to productivity itself. As a result, the regime transitions in the TVAR are associated with "large" productivity fluctuations that the estimation interprets as responses to exogenous shocks to productivity. Consequently, the TVAR identifies shocks that cause large productivity fluctuations in both regimes.

and match-specific idiosyncratic productivity, x. For each a_t, there is a reservation level of individual productivity $x^r(a_t)$, below which jobs are not mutually efficient and are dismissed. Similarly, there is a level of individual productivity $x^s(a_t)$, below which employed workers find it efficient to pay a fixed cost k^s to search for other jobs.[10]

Within each period t, the timing of events is as follows. At the start of period, firms post vacancies that are matched with job seekers by the end of the period. Employed workers produce and may search for a new job within the period. At the end of the period, a fraction of employed workers is exogenously separated, and another fraction of employed workers obtains a new draw of individual productivity. At the beginning of the next period $t+1$, aggregate productivity a_{t+1} and the individual productivity x of each worker are observed. Each firm converts profitable matches into jobs, and each worker that searches on the job decides whether to move to a new firm or remain in the current job.

The matching function. A matching function encapsulates search frictions in the labor market. In each period t, the constant-returns-to-scale matching function establishes the number of matches between job seekers and vacancies:

$$m_t = m(u_t + \psi_t, v_t) = \gamma(u_t + \psi_t)^{1-\eta} v_t^{\eta}, \tag{3}$$

where u_t is unemployment, ψ_t is the mass of OJS workers, v_t are vacancies, and $0 < \eta < 1$. The sum of employed OJS workers and unemployed workers forms the number of job searchers. The probability for a job seeker to match a vacancy and for a vacancy to be filled can be expressed in terms of the "labor market tightness," defined as the ratio of vacancies to job seekers, $\theta_t = v_t/(u_t + \psi_t)$. The probability of a job seeker finding a suitable vacancy is $p(\theta_t) = m(u_t + \psi_t, v_t)/(u_t + \psi_t) = m(1, v_t/(u_t + \psi_t))$, and the probability for the firm to find a suitable worker is $q(\theta_t) = m((u_t + \psi_t), v_t)/v_t = m((u_t + \psi_t)/v_t, 1)$.[11]

[10] Below, we use the following notation: when x has a time subscript, it refers to an aggregate variable (e.g. x_t^r is the individual productivity threshold that applies to all firms given an aggregate productivity level a_t). Without a subscript, it refers to any individual productivity level for a match independent of aggregate states.

[11] As we discuss below, due to individual productivity shocks, $p(\theta_t)$ and $q(\theta_t)$ cannot be interpreted as the job finding and job-filling probabilities, respectively. We therefore refer to them as the "contact" probabilities for workers and firms, respectively. Also note that labor market tightness includes OJS workers and hence differs from the empirically observable vacancy/unemployment ratio.

Production and matched workers. Each firm manufactures a unique final good by hiring labor. Each hired worker produces $a_t x$ units of output. Aggregate productivity a_t follows the auto regressive process:

$$\ln a_{t+1} = \rho \ln a_t + \epsilon_{t+1}, \tag{4}$$

where $\epsilon \sim N(0, \sigma^2)$ and $\|\rho\| < 1$. During each period $t+1$, an existing worker maintains the previous individual productivity level with probability $(1-\lambda)$, and with probability λ, the worker receives a new productivity drawn from the constant distribution $F(x)$, which is assumed to be a log-normal distribution with standard deviation σ_x^2 and mean $-\frac{\sigma_x^2}{2}$ truncated above at x^H. Job seekers matched in period t also receive a productivity value from the same distribution in the beginning of period $t+1$.

Job separation and job creation. During each period t, total job separations comprise exogenous and endogenous terminations. Existing workers are separated from their jobs with an exogenous probability $s < 1$. Given aggregate productivity a_t, the firm establishes a threshold of individual productivity $x^r(a_t)$, below which existing matches are mutually inefficient. All workers whose individual productivity satisfies $x < x^r(a_t)$ are dismissed whereas if $x \geq x^r(a_t)$ the job relation continues.

On-the-job search. A worker may search for a new job at the cost k^s. An employed job searcher is matched to a firm from the same pool as the unemployed job seekers and therefore is subject to the same matching frictions. Once matched, the worker receives an idiosyncratic x from the distribution $F(x)$, as any other newly-matched job seeker. If the draw of individual productivity is below the reservation threshold, the match is discontinued and the employed job searcher stays in the original job. Also, as any existing worker, the job searcher who remains with her current firm draws a new individual productivity with probability λ and faces exogenous job separation. Each firm applies the same separation threshold to employed and unemployed job seekers. As a result, a fraction of job-to-job transition may entail lower wages for the workers. This simplifying assumption abstracts from the fact that the actual outside option for employed job seekers is their current employment contract rather than unemployment. This

simplification avoids the issue of heterogeneity in wage bargaining and hence the fact that new wages depend on the value of x for the current contract and the value of x from the previous employer. These dynamics would substantially complicate the aggregation for the solution of the model because the entire distribution of x over employed workers, which is history-dependent, would become a relevant state variable for firms' decisions.[12] Instead, under the simplifying assumption, OJS workers are perfectly substitutable for unemployed ones from the viewpoint of hiring firms. The main reason for making this assumption is that it allows to have a highly tractable model with both OJS and firms' reservation productivity for new matches from unemployment, one of our channels of interest for state dependence.

Recursive formulation. Four value functions solve the model: the value of unemployment (U), the value of a vacancy (V), the *joint* value of a match (M), and the *joint* surplus of a match (S). The joint surplus of a match is split in constant proportions through Nash bargaining for wages, assigning the fraction ϕ of the joint match surplus to the worker and the fraction $1-\phi$ to the firm. The value of unemployment is:

$$U(a_t) = b + \beta \mathbb{E}_t \Big[U(a_{t+1}) + p(\theta_t)\phi \int S(a_{t+1}, x')dF(x') \Big]. \tag{5}$$

Equation (5) shows that the value of unemployment is equal to the opportunity cost of working (i.e., the flow value of unemployment b) and the expected benefits that finding a job brings in the next period. In period $t+1$, the prospective worker encounters a suitable vacancy with probability $p(\theta_t)$, and if the match is mutually profitable, the worker gains a fraction ϕ of the total surplus on top of the value of staying unemployed. Otherwise the job seeker remains unemployed, gaining the continuation value $U(a_{t+1})$.

The value of an open vacancy is:

$$V(a_t) = -k + \beta \mathbb{E}_t \Big[V(a_{t+1}) + q(\theta_t)(1-\phi) \int S(a_{t+1}, x')dF(x') \Big]. \tag{6}$$

[12] Within the microeconomic literature, the details of wage bargaining from on-the-job search have been considered by Postel-Vinay and Robin (2004) and Shimer (2006), among others, in models without aggregate uncertainty. See also Gottfries (2018) for a recent contribution.

Equation (6) shows that the present value of an open vacancy is equal to the fixed cost of posting the vacancy k and the expected benefits that the vacancy brings in the next period. In period $t+1$, the firm finds a prospective worker with probability $q(\theta_t)$, and if the match is profitable, the firm gains a fraction $(1-\phi)$ of the total surplus. Otherwise, the vacancy remains open, giving the firm a continuation value $V(a_{t+1})$. In equilibrium, the free-entry condition leads firms to post vacancies until their expected value is equal to zero in each period (i.e. $V(a_t) = 0$, for all t). This equilibrium condition applied to equation (6) yields the job-creation condition:

$$\frac{k}{q(\theta_t)} = (1-\phi)\beta\mathbb{E}_t\Big[\int S(a_{t+1}, x')dF(x')\Big]. \tag{7}$$

Equation (7) shows that the expected cost of a match (left-hand side) is equal to the expected benefit that the match brings into the firm if the job is established (right-hand side). With this formulation, the problem can be recast in terms of choosing a given market tightness $\theta(a_t)$ for a level of aggregate productivity.

For each given vector of (a_t, x), an employment relationship is established if the match is mutually efficient. Therefore, the joint value of the job relation is greater than the outside options (i.e., the individual values from separation). Thus, the joint value of a match is

$$M(a_t, x) = \max\Big\{M^{n,c}(a_t, x), M^{s,c}(a_t, x), U(a_t) + V(a_t)\Big\}, \tag{8}$$

where $M^{n,c}(a_t, x)$ is the joint value of a continued match without OJS, $M^{s,c}(a_t, x)$ is the joint value of the continued match with OJS, and $U(a_t) + V(a_t)$ is the joint value of the outside option. The joint value of a continued match without OJS is

$$\begin{aligned} M^{n,c}(a_t, x) = a_t x + \beta\mathbb{E}_t\Big\{&U(a_{t+1}) + V(a_{t+1}) \\ &+ (1-s)\Big[(1-\lambda)S(a_t, x) + \lambda\int S(a_{t+1}, x')dF(x')\Big]\Big\}. \end{aligned} \tag{9}$$

Equation (9) shows that the joint value of a continued match is equal to production plus the expected continuation value of the work relationship. Meanwhile the joint value of a continued

match while searching on the job is

$$M^{s,c}(a_t,x) = a_t x - k^s + \beta\mathbb{E}_t\Big\{U(a_{t+1}) + V(a_{t+1}) + \Big[1 - p(\theta_t)\overline{F(x^r_{t+1})}\Big](1-s)\Big[(1-\lambda)S(a_{t+1},x) + \lambda\int S(a_{t+1},x')dF(x')\Big] + p(\theta_t)\phi\int S(a_{t+1},x')dF(x')\Big\}. \quad (10)$$

where $\overline{F(x^r_{t+1})} = \Big(1 - F(x^r_{t+1})\Big)$.

The last term uses the fact that $\int_{x^r(a_{t+1})}^{x^H} S(a_{t+1},x')dF(x') = \int^{x^H} S(a_{t+1},x')dF(x')$, which represents the expected surplus that may accrue to the worker if she is matched with another firm and the match is continued. This event materializes with probability $p(\theta_t)\overline{F(x^r_{t+1})}$ and encompasses all of the values of x above the reservation threshold x^r_{t+1}.

The joint surplus of a match equals the value of a match M, net the outside option for the worker U, and the firm V (i.e. $S{=}M - U - V$). Thus, the value function for the joint surplus of a continuing match is:

$$S(a_t,x) = \max\{S^{n,c}(a_t,x), S^{s,c}(a_t,x), 0\}, \quad (11)$$

where $S^{n,c}(a,x)$ is surplus of the match when the job relation continues without OJS, and $S^{n,c}(a,x)$ is the surplus of a continued match with OJS. The surpluses are defined as follows:

$$S^{n,c}(a_t,x) = a_t x - b + \beta\mathbb{E}_t\Big\{(1-s)\Big[(1-\lambda)S(a_{t+1},x) + \lambda\int S(a_{t+1},x')dF(x')\Big] - p(\theta_t)\phi\int S(a_{t+1},x')dF(x')\Big\}, \quad (12)$$

$$S^{s,c}(a_t,x) = a_t x - k^s - b + \beta\mathbb{E}_t\Big\{\Big[1 - p(\theta_t)\overline{F(x^r_{t+1})}\Big](1-s)\Big[(1-\lambda)S(a_{t+1},x) + \lambda\int S(a_{t+1},x')dF(x')\Big]\Big\}. \quad (13)$$

A worker searches while on the job if $S^{s,c}(a_t,x) \geq S^{n,c}(a_t,x)$. The presence of OJS introduces a threshold $x^S(a_t)$, below which it is efficient to search on the job. For values of the thresh-

old such that $x^r(a_t) < x^S(a_t)$, it is efficient to incur in the search costs for all $x \in (x^r(a_t), x^S(a_t)]$. Substituting equations (12) and (13) into the condition $S^{s,c}(a_t, x) \geq S^{n,c}(a_t, x)$ yields

$$k^s \leq \beta \mathbb{E}_t \Big\{ -p(\theta_t)\overline{F(x^r_{t+1})}(1-s)\Big[(1-\lambda)S(a_{t+1}, x) + \lambda \int S(a_{t+1}, x')dF(x')\Big] + p(\theta_t)\phi \int S(a_{t+1}, x')dF(x') \Big\}. \quad (14)$$

With equality, equation (14) determines the efficient threshold under which workers engage in OJS. Intuitively, the cost of searching has to be smaller than the increase in the continuation value coming from the possibility of finding a new match.

Finally, for a given aggregate state a_t, the individual productivity threshold for endogenous separations is the value of x that makes the joint surplus of continuing a match equal to zero, such that:[13]

$$S^{s,c}(a_t, x^r(a_t)) = 0. \quad (15)$$

Labor flows depend on the distribution of x across employed matches. The distribution of individual productivity among employed workers is history dependent: $Pr(X < x \big| a^t)$, where a^t represents the history of aggregate productivity shocks $\{a_0, a_1, ..., a_t\}$ realized up to time t. This distribution is determined by the measure of employed workers over individual productivity $e_t(x)$, which follows a law of motion determined by flows between unemployment and employment and within employment. For values of x in the OJS interval $(x^r_t, x^s_t]$:

$$e_{t+1}(x) = p(\theta_t)[1 - e_t(x_H)][F(x) - F(x^r_{t+1})] + p(\theta_t)[F(x) - F(x^r_{t+1})]e_t(x^s_t) + (1-s)\Big\{\lambda[F(x) - F(x^r_{t+1})]\Big[e_t(x_H) - p(\theta_t)\overline{F(x^r_{t+1})}e_t(x^s_t)\Big] + (1-\lambda)\Big[e_t(x) - e_t(x^r_{t+1})\Big]\Big[1 - p(\theta_t)\overline{F(x^r_{t+1})}\Big]\Big\}. \quad (16)$$

[13] As $S(a_t, x)$ is monotonically increasing in x, the individual productivity threshold $x^r(a_t)$ is unique, and $S(a_t, x) > 0 \ \forall \ x > x^r(a_t)$. Appendix E.1 provides a detailed discussion.

For values in the non-searching region $x > x_t^s$:

$$
\begin{aligned}
e_{t+1}(x) = p(\theta_t)[1 - e_t(x_H)][F(x) - F(x_{t+1}^r)] + p(\theta_t)[F(x) - F(x_{t+1}^r)]e_t(x_t^s) \\
+ (1-s)\Big\{\lambda[F(x) - F(x_{t+1}^r)]\Big[e_t(x_H) - p(\theta_t)\overline{F(x_{t+1}^r)}e_t(x_t^s)\Big] \\
+ (1-\lambda)\Big[e_t(x) - e_t(x_t^s) + (1 - p(\theta_t)\overline{F(x_{t+1}^r)})[e_t(x_t^s) - e_t(x_{t+1}^r)]\Big]\Big\}. \quad (17)
\end{aligned}
$$

Gross flows and transition rates. Gross flows from employment to unemployment represent the total mass of workers separated from a job between two periods:

$$
\begin{aligned}
EU_{t+1} = s\Big[e_t(x_H) - p(\theta_t)\overline{F(x_{t+1}^r)}e_t(x_t^s)\Big] \\
+ (1-s)\Big\{\lambda F(x_{t+1}^r)\Big[e_t(x_H) - p(\theta_t)\overline{F(x_{t+1}^r)}e_t(x_t^s)\Big] \\
+ (1-\lambda)e_t(x_{t+1}^r)\Big[1 - p(\theta_t)\overline{F(x_{t+1}^r)}\Big]\Big\}. \quad (18)
\end{aligned}
$$

The job separation rate is then defined as the probability that an employed worker in period t is not employed in period $t+1$: $SR_t = EU_{t+1}/e_t(x_H)$. Similarly, the gross unemployment to employment (UE) flow is the total mass of workers who start a new job from unemployment:

$$UE_{t+1} = u_t p(\theta_t)(1 - F(x_{t+1}^r)),$$

and the job finding rate (JFR) is defined as the probability that an unemployed worker in period t is not unemployed in period $t+1$:

$$JFR_t = UE_{t+1}/u_t = p(\theta_t)(1 - F(x_{t+1}^r)). \quad (19)$$

The job-to-job rate (JJR) is measured as the ratio of gross employment to new employment (EE) flows over total employment:

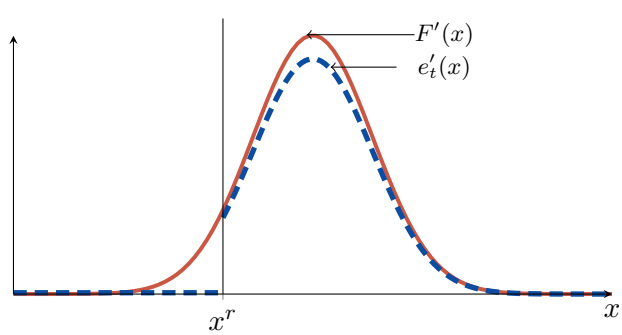
$$JJR_{t+1} = \frac{EE_{t+1}}{e_t(x_H)} = \frac{e_t(x_t^s)p(\theta_t)\overline{F(x_{t+1}^r)}}{e_t(x_H)}. \quad (20)$$

4 Mechanisms for state-dependent fluctuations

In this section, we discuss the two main mechanisms that generate state dependence in the response of the unemployment rate to productivity shocks. The first mechanism is based on the changes in the reservation productivity, and the second mechanism is based on the non-linearities in the firm's hiring decision.

Reservation productivity. Changes in the individual productivity threshold $x^r(a)$ generate distinct response of both job creation and job destruction over states of aggregate productivity. The mechanism relies on the interaction of the idiosyncratic productivity threshold $x^r(a)$ with the distribution of workers over x. For the job finding rate, the relevant distribution is the that of newly-established matches from unemployment $F(x)$. For the separation rate, the relevant distribution is the measure of incumbent workers $e_t(x)$. Figure 5 shows an illustrative probability density function for the productivity of new matches (i.e., $F'(x)$) and the measure of incumbent ones (i.e. $e_t'(x)$) (solid and dashed lines, respectively). We make the standard assumption (see Moscarini, 2005) that $F(x)$ is continuous, unimodal, and twice differentiable, with x^r always lying below the mode. For new matches, only those with $x > x^r$ turn into employment. For incumbent workers, $e_t(x)$ is truncated below at x^r.

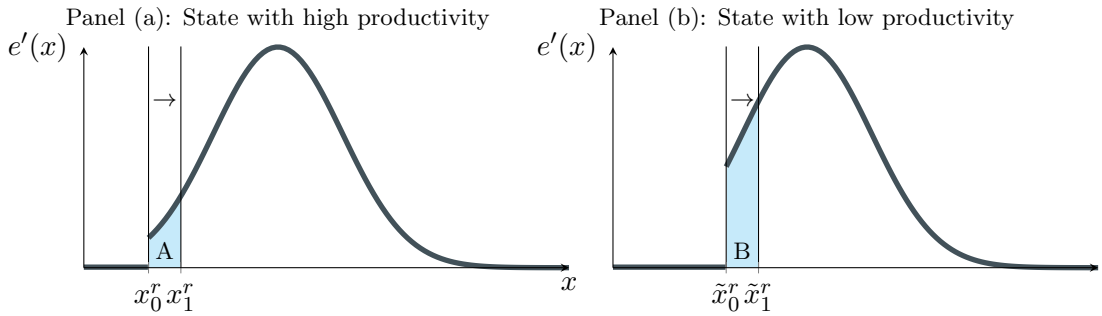
Figure 5: Distribution for $F'(x)$ and $e_t(x)$



Note. The figure shows an illustrative example of the p.d.f. of $F(x)$ (labelled $F'(x)$, solid red line) and the measure of employment over $e_t'(x)$ (dashed blue line).

Movements in the individual productivity threshold generate distinct responses in the job separation rate in relation to the state of aggregate productivity. Panel (a) in Figure 6 shows an

Figure 6: States of aggregate productivity and the job separation rate



Note. An increase of the threshold of individual productivity from $\tilde{x}_0^r$ to $\tilde{x}_1^r$ generates a larger response in the job separation rate in states with low aggregate productivity than an equivalent increase of the threshold of individual productivity from x_0^r to x_1^r in states with high aggregate productivity. The shaded area shows the mass of jobs sensitive to job separation in response to the change in the threshold.

illustrative example of the initial productivity threshold x_0^r on the distribution for continuing jobs $e(x)$ that is associated with a high level of aggregate productivity.[14] In response to a fall in aggregate productivity, the individual productivity threshold increases from x_0^r to x_1^r, leading to a rise in the job separation rate equal to shaded area A. Panel (b) shows the effect of an equivalent fall in aggregate productivity from an initially low level of aggregate productivity. In this instance, the individual productivity threshold is high and located in the domain of the distribution with high density. The same fall in aggregate productivity increases the individual productivity threshold from $\tilde{x}_0^r$ to $\tilde{x}_1^r$, leading to a rise in the job separation rate equal to shaded area B, which is larger than area A. Thus, the effect of a shock on the mass of jobs exposed to movements in the individual productivity threshold differs across levels of aggregate productivity, and the response of the job separation rate to the aggregate productivity shock is larger when aggregate productivity is low. By the same principle, the job finding rate may exhibit larger responses in the states with low aggregate productivity. When the initial level of aggregate productivity is low, x^r is located in a region of $F(x)$ with a higher density. Hence, shifts in x^r entail larger fluctuations in the fraction of new matches that do not turn into employment.

A similar mechanism is outlined in Zanetti (2011) and Sedláček (2014) for the study of

[14] Given that the surplus is increasing in both x and a, equation (15) implies that a high level of aggregate productivity is associated with a low individual productivity threshold. Assuming that the threshold always lies below the distribution mode, a lower threshold is located in a region of the distribution associated with lower density.

fluctuations in labor market variables over the business cycle. Ferraro (2018) shows the interplay between this mechanism and worker heterogeneity to study the skewness in the separation and job finding rates.

Non-linearity in recruiting decisions. In a standard DMP model, Hagedorn and Manovskii (2008) show that the firms' optimal decision on labor market tightness may exhibit strong non-linearity with respect to aggregate productivity if the surplus from forming a match is sufficiently small and responsive to productivity. Petrosky-Nadeau and Zhang (2017) show that deep recessions may even involve no vacancy postings if the match surplus becomes lower than the expected recruitment cost. They show that the response of the job finding rate to a productivity shock is amplified in states of low aggregate productivity, leading to greater volatility in the unemployment rate. This alternative mechanism generates state dependence in job creation by directly affecting the labor market tightness without involving changes in the threshold of idiosyncratic productivity x^r. Therefore, this channel is present even in the version of the DMP model that abstracts from the heterogeneity in productivity across workers and endogenous job separation.

5 Model simulation and quantitative results

This section presents the calibration of the model and the quantitative results. It first simulates the benchmark model to assess the performance in replicating the empirical state dependence of the unemployment rate, the job finding rate, and the job separation rate. It then uses generalized IRFs to study the dynamic responses of labor market variables in different states of aggregate productivity and compares them to the responses in the TVAR model.

5.1 Calibration and simulated moments

We solve the model non-linearly, iterating over the policy function on a discretized state space for aggregate productivity, following the approach in Tauchen (1986). Appendix E.2 describes the solution procedure.

We calibrate the model at a monthly frequency using standard values in the literature, except for the parameters s, σ_x, λ, b, c_s, and γ, whose values are set to match specific empirical statistics. The parameter values taken from other studies are summarized in Table 2.[15] The discount factor β is set equal to $0.953^{(1/12)}$, as in Shimer (2005). The cost of posting a vacancy κ is set equal to 0.17, following Fujita and Ramey (2012), who derive it from estimates of application screening costs and average vacancy duration reported by Barron and Bishop (1985) and Barron et al. (1997). The elasticity of the matching function with respect to vacancies η is set equal to 0.5 in the middle range of empirical estimates in Petrongolo and Pissarides (2001). To satisfy the Hosios (1990) condition, which ensures that the equilibrium of the decentralized economy is Pareto efficient, we assume that the elasticity of labor market tightness with respect to vacancies is equal to the firm's bargaining power, $(1-\phi)$, such that $\eta = (1-\phi) = 0.5$. The upper bound of the distribution of x is set to 1.5, a value sufficiently high for $F(x^H)$ to be effectively equal to 1. The autoregressive parameter ρ is set equal to 0.973, as in Hagedorn and Manovskii (2008).

Table 2: Calibrated parameter values for the baseline model.

Parameter	Description	Value
Externally Calibrated		
β	Discount factor	$0.953^{(1/12)}$
κ	Vacancy cost	0.17
η	Elasticity of matching with respect to vacancies	0.5
ϕ	Worker's bargaining power	0.5
ρ	Persistence parameter of aggregate productivity	0.973
σ	Standard deviation of aggregate productivity shocks	0.0063
x^H	Upper bound of $F(x)$	1.5
Internally Calibrated		
b	Flow value of unemployment	0.75
γ	Matching function efficiency	0.41
s	Exogenous separation probability	0.03
σ_x	Standard deviation of x draw	0.115
λ	Probability of new x draw	0.05
κ_s	Cost of OJS	0.12

We now describe the parameter values calibrated to match specific empirical statistics. We choose σ so that the quarterly standard deviation of the exogenous process is equal to the variance of labor productivity that is explained by productivity shocks in the TVAR. On

[15] We retain these values for the calibration of the alternative models in the later sections.

average, over the two regimes, the fraction of the stationary variance in the forecast error of labor productivity that is explained by productivity shocks is 80 percent. Thus, instead of matching the full variance of productivity that also reflects other shocks, we match the fraction that is explained by the TVAR, corresponding to a standard deviation of 0.023 at quarterly frequency. The resulting value of σ is 0.0063. We choose values for the flow value of unemployment b and the standard deviation of individual productivity shocks σ_x to match the volatility of the job finding and separation rate that the TVAR model attributes to productivity shocks. The fraction of the variance of the two rates that is explained by shocks in the TVAR is 30 and 40 percent, respectively, of their unconditional variances. To match the standard deviations of the job finding separation rates corresponding to these fractions, we set $b = 0.75$ and σ_x =0.115. Since the distribution of individual log productivity shocks is a truncated log-normal, we set the mean equal to $-\sigma_x^2/2$, to normalize the expected value of x to 1. The probability of receiving a new individual productivity shock λ is set at 0.05 to target the quarterly autocorrelation of the separation rate, detrended using the Hamilton (2018) filter, of 0.80. The matching function parameter γ is set equal to 0.41 to replicate the empirical average job finding rate of 0.45, as in Hagedorn and Manovskii (2008). The exogenous separation probability s is set equal to 0.03 to match the empirical average job separation rate of 0.033. The cost of searching on the job k^s is set to equal 0.13 to match the mean monthly job-to-job transition probability of 3.2 percent, calculated from U.S. data by Moscarini and Thomsson (2007). Table 2 summarizes these parameters.

Table E.1 in the Appendix shows that the simulated target moments are close to the empirical counterparts in the data. Table 3 reports the standard deviation, the correlation with productivity, and the autocorrelation at quarterly frequency for the main variables in the model that are comparable to the TVAR. The moments are based on a set of 1,000 simulations of the same length as the empirical data. The model replicates accurately the sign of the correlations of the variables with productivity from the IRFs of the TVAR model. The unemployment rate and the job separation rate are negatively correlated with productivity whereas the job finding rate is positively correlated.[16]

[16] In the appendix, Table E.2 reports the simulated statistics for all the variables in the model. Importantly,

Table 3: Simulated labor market statistics in the baseline model.

Data	P	U	JFR	SR
$\bar{X}$	-	5.85	43.2	3.32
σ_X	2.43	1.38	6.63	3.09
σ_X|TVAR	2.09	0.86	3.34	0.17
Model	P	U	JFR	SR
$\bar{X}$	1.1	7.3	44.4	3.34
σ_X	2.17	0.82	3.28	0.17
$\text{Corr}(p_t, X_t)$	1.00	-0.95	0.99	-0.95
$\text{Corr}(X_t, X_{t-1})$	0.88	0.90	0.89	0.82

Note. The table reports cyclical statistics for average labor productivity (P), the unemployment rate (U), the job-finding rate (JFR), the job-separation rate (SR). $\bar{X}$ is the long-run mean of the variable X, σ_X the standard deviation, $\text{Corr}(p_t, X_t)$ the correlation with productivity, and $\text{Corr}(X_t, X_{t-1})$ the autocorrelation. The model moments are computed as the mean of 1,000 simulations in which the simulated data are treated in the same way as the empirical series. See Appendix E.2 for details. The empirical standard deviations explained by the TVAR (σ|TVAR) are derived by multiplying the fraction of the forecast error variance explained by productivity shock in the TVAR (weighted average across the two regimes) with the raw variance of the data (σ^2) and taking the square root. Standard deviations in both the model and the data are multiplied by 100.

5.2 Assessing state dependence

In this section, we examine the extent to which labor market dynamics differ across states of high and low aggregate productivity. We investigate the issue in two ways. First, we compute the standard deviations of labor market variables when productivity is below and above the 58th percentile, which is the median of the posterior distribution of the threshold identified in the TVAR. We compare these standard deviations in the model to those in the TVAR that are explained by productivity shock in the low- and high-productivity regimes. Second, we compute generalized IRFs in the model for the 10th and 90th percentile of aggregate productivity. The first method provides a general assessment on the changes in the response of the two transition rates to productivity over the business cycle, while the second approach focuses on two specific percentiles of aggregate productivity that are directly comparable to the GIRFs in the TVAR.

Table 4 compares the standard deviation of productivity, the unemployment rate, the job finding rate, and the separation rate explained by productivity shocks in the TVAR model (left panel) and the theoretical model (right panel) for states of productivity below (Columns

the negative correlation between unemployment and vacancies yields an empirically consistent Beveridge Curve. Additionally, the job-to-job rate is also procyclical, which is consistent with empirical evidence from Mukoyama (2014) and Molloy et al. (2016). Furthermore, while the exercise of our analysis is to compare the TVAR model with the theoretical model, in Pizzinelli et al. (2018) we show that a different calibration replicates closely unconditional business cycle statistics.

(1) and (4)) and above (Columns (2) and (5)) the 58th percentile.[17] Columns (3) and (6) report the ratio of the standard deviations between the two regimes.[18] For productivity, the ratios of the standard deviations between the two regimes in the TVAR and theoretical model are similar and equal to 1.29 and 1.13, respectively. For the unemployment rate, the ratio of standard deviations above and below the threshold is 2.04 in the TVAR. The counterpart in the model is 1.92, implying that the model captures 88 percent of the state dependence from the TVAR.[19] For the job finding rate, the ratio is equal to 2.01 in the TVAR and 1.37 in the model, respectively, indicating that the theoretical model accounts for approximately 36 percent of the higher volatility in low-productivity states from the TVAR. For the separation rate, the ratios are 2.28 and 1.98 in the TVAR and theoretical model, respectively, implying that the theoretical model captures approximately 76 percent of the higher volatility elicited by the TVAR.

Table 4: State dependence for the baseline model relative to the TVAR.

	σ|TVAR			Model		
	$\sigma_{p<58pct.}$ (1)	$\sigma_{p>58pct.}$ (2)	Ratio (3)	$\sigma_{p<58pct.}$ (4)	$\sigma_{p>58pct.}$ (5)	Ratio (6)
Prod	1.61	1.25	1.29	1.35	1.23	1.13
U	1.22	0.6	2.04	0.66	0.35	1.92
JFR	5.01	2.49	2.01	2.24	1.68	1.37
SR	0.40	0.17	2.28	0.14	0.07	1.98

Note. $\sigma_{p<(>)58pct.}$ represents the standard deviation of the variable for the productivity state below (above) the 58th percentile of the productivity distribution, which corresponds with the estimated threshold in the TVAR. Columns (1) and (2) report the standard deviation of the variables explained by productivity shocks in the TVAR. Column (3) reports the ratio of (1) and (2). Columns (4) and (5) report the corresponding standard deviations from simulations of the baseline model. Entries are averages of 1,000 simulations over 768 monthly periods aggregated at quarterly frequency and have the same length as the period 1950:I-2014:IV. Column (6) reports the average of the ratio of the standard deviations across the simulations. All standard deviations are multiplied by 100 for presentation purposes.

To illustrate the extent to which the dynamic responses of labor market variables differ across states with high and low aggregate productivity, Figure 7 plots IRFs for the unemployment rate (left panels), the separation rate (middle panels) and the job finding rate (right panels) to a one-standard-deviation productivity shock (solid line) together with the 5th-90th percentiles (shaded area).[20] The top panels report the response of the variables in states in which productivity

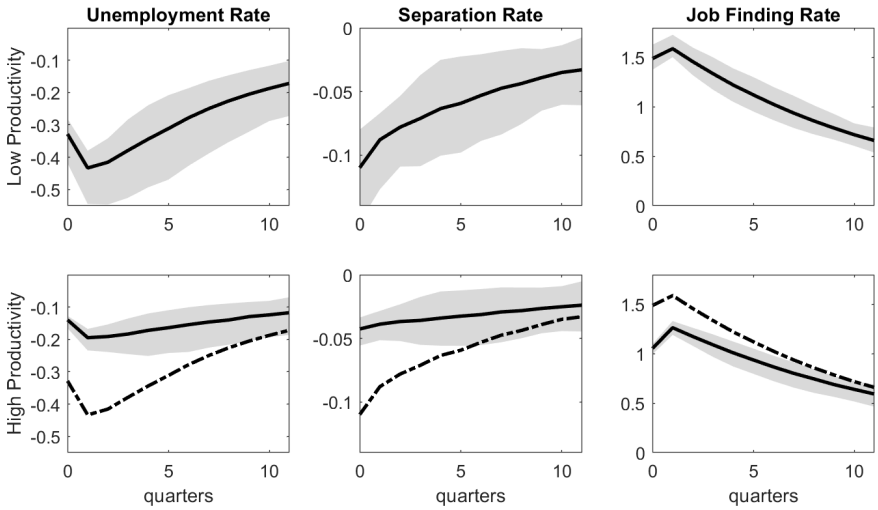
[17] As mentioned, the 58th percentile corresponds to the median of the posterior distribution of the threshold identified in the TVAR.

[18] More precisely, Column (6) reports the average ratio across 1,000 simulation.

[19] This statistic is derived by taking the ratios of the "excess volatility" (i.e., (1.92-1)/(2.04-1)).

[20] Appendix E.3 describes the computational method to derive generalized IRFs.

Figure 7: Generalized IRFs: baseline model with OJS and endogenous separations.



Note. The solid line represents the mean IRF value in each period. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points. In the second row, the dashed line reports the mean IRF from the first row for comparison.

is below the 10th percentile of its distribution while the bottom panels consider an initial productivity level above the 90th percentile. In the lower panels, the dashed line reports for comparison the mean IRF from the states of productivity below the 10th percentile reported in the first row.

The left panels in Figure 7 show that the response of the unemployment rate is more than twice as large in the state with low aggregate productivity compared to the state with high aggregate productivity. Since, the response of the unemployment rate results from movements in the job separation rate and the job finding rate, we study the dynamics of the transition rates to assess their separate contribution to the state dependence of the unemployment rate.

The middle panels in Figure 7 show that the response of the job separation rate is more than twice as large in the state with low productivity compared to the state with high productivity. This magnitude is similar to the GIRFs from the TVAR. As discussed in Section 4, the mechanism that generates these distinct dynamics is straightforward. The different responses with respect to the state with aggregate productivity originate from the effect of shifts in the reservation threshold for individual productivity on the job separation rate. In the state with high aggregate productivity, the threshold is low and located in a region of individual

productivity distribution with low density. Aggregate shocks that move the threshold displace a limited number of workers and therefore have a limited effect on the job separation rate. By contrast, in the state with low aggregate productivity, the threshold of individual productivity is high and located in a region of the distribution of individual productivity with high density. Thus, an identical aggregate productivity shock that moves the threshold similarly displaces a larger fraction of workers, thereby generating a large shift in the job separation rate.

The right panels in Figure 7 show that the response of the job finding rate is 40 percent larger in the state with low aggregate productivity compared to the state with high aggregate productivity. This magnitude is smaller than in the GIRFs from the TVAR. As discussed in Section 4, the asymmetry in the job finding rate may result from two forces related to the change in the reservation productivity and the non-linearity of the market tightness with respect to productivity. To examine the contribution of each channel to the response in the job finding rate, Appendix E.4 constructs two counterfactual IRFs, in which only one channel at a time is active, showing that changes in the reservation productivity x^r are the predominant mechanism for the state dependence in the job finding rate, and explaining most of the difference in the response of the job finding rate across states.[21]

Overall, the analysis suggests that the baseline model generates large state dependence in the separation rate while the job finding rate plays a limited role. Appendix E.5 develops an alternative approach to quantitatively assess the model's ability to capture the empirical state dependence by comparing the model IRFs to the GIRFs from the the TVAR. It measures the fraction of the difference in the low- and high-productivity GIRFs in Figure 4 that is replicated by the difference in the model's IRFs at various time horizons. This alternative approach yields similar findings to the baseline analysis, and it shows that model explains approximately 75 percent in the state dependence of the unemployment rate at 8 quarters after a productivity shock. The corresponding metrics for the separation rate and the job finding rate are equal to 84 and 33 percent, respectively.

[21] See Appendix E.4 for a detailed explanation of the exercise. Figure E.1 shows the main findings of the exercise.

6 Alternative calibrations and reduced models

In this section, we investigate whether state dependence in the baseline model is robust to: (i) alternative calibrations of the model, (ii) abstracting from OJS, and (iii) assuming no OJS and exogenous separations. Each subsection briefly outlines the calibration for the alternative models and presents the main findings, leaving supplementary material to Appendix F.[22] The exercise points to the importance of the calibration of key parameters, in particular the probability of drawing a new idiosyncratic productivity (λ), and the presence of OJS to generate state dependence.[23]

6.1 Alternative calibration of baseline model

The baseline calibration targets the conditional volatility of the job separation rate and job finding rate. As an alternative calibration strategy, we follow the approach in Krause and Lubik (2007) and target the volatility of job destruction *relative* to the volatility of employment (i.e., σ_{SR}/σ_e), which is equal to 0.22 over the sample period 1950-2014 in the baseline calibration.[24] The value of σ_x in the alternative calibration is equal to 0.12, which is only slightly higher than the baseline calibration. We maintain the value of b from the baseline calibration.

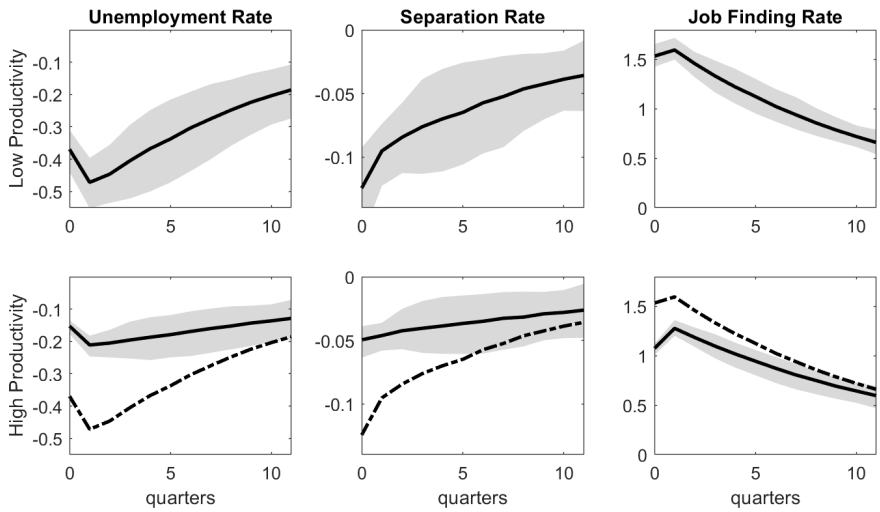
Figure 8 shows IRFs for the model with the alternative calibration. Both the job separation rate and job finding rate exhibit a larger response to the shock in low-productivity states relative to high-productivity states. The difference in responses across states is similar in magnitude to the baseline calibration. The same result holds for the simulated standard deviations of the unemployment rate and job-transition rates, reported in Table 5. The ratios of the standard deviations above and below the thresholds are very similar to those in the baseline calibration in Table 4.

[22] In Pizzinelli et al. (2018), we calibrate the model to replicate unconditional moments in the data, obtaining similar findings regarding the relation of these alternative models to the baseline model. In the Online Appendix, we perform additional robustness analysis, focusing in particular on the Hagedorn and Manovskii (2008) calibration approach.

[23] Appendix F reports the calibration of the system for all versions of the model. Table F.1 reports internally calibrated parameters, and Table F.2 reports the simulated target moments.

[24] In Krause and Lubik (2007), the targeted value is 7. The difference comes from the treatment of the data. We focus on fluctuations in the raw rate in percentage points while these authors focus on percent deviations of the employment level from the steady state. In the model, we define the employment rate as $e_t = 1 - u_t$.

Figure 8: Generalized IRFs: model with OJS and endogenous separations, under the alternative calibration.



Note. The solid line represents the mean IRF value in each period. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels.

Table 5: State dependence for the alternative models.

	Alternative Calibration			No OJS			No OJS and exogenous separation		
	$\sigma_{p<58pct.}$	$\sigma_{p>58pct.}$	Ratio	$\sigma_{p<58pct.}$	$\sigma_{p>58pct.}$	Ratio	$\sigma_{p<58pct.}$	$\sigma_{p>58pct.}$	Ratio
Prod	1.34	1.22	1.14	1.34	1.21	1.14	1.42	1.25	1.17
U	0.69	0.37	1.91	0.62	0.33	1.95	0.36	0.24	1.52
JFR	2.24	1.70	1.36	2.20	1.79	1.26	2.23	1.84	1.24
SR	0.15	0.08	1.96	0.14	0.06	2.38	-	-	-

Note. $\sigma_{p<(>)58pct.}$ represents the standard deviation of the variable for the productivity state below (above) the 58th percentile of the productivity distribution, which corresponds with the estimated threshold in the TVAR. Entries are averages of 1,000 simulations over 768 monthly periods aggregated at quarterly frequency and have the same length as the period 1950:I-2014:IV. Columns (3), (6), and (9) report the average of the ratio of the standard deviations across the simulations for the corresponding model. All standard deviations are multiplied by 100.

6.2 Model without OJS

To assess the role of OJS for state dependence, we consider a version of the model that abstracts from OJS. Without OJS, workers cannot search for employment opportunities and therefore the surplus function becomes $S^{n,c}(a_t, x)$, as reported in equation (12). Target statistics for the calibration are the same as in the baseline model. The main difference in the calibration with the baseline model is the higher value for the arrival rate of productivity shocks λ, implying that workers receive a new productivity draw more frequently than in the benchmark model.

Intuitively, the higher value for λ implies that even in the absence of OJS a large fraction of workers who are in the lower tail of the productivity distribution have a chance to avoid job separation by drawing a new idiosyncratic productivity x in each period while remaining in the same job.

Figure 9 shows the IRFs. While this model replicates well the larger response in the job separation rate in the low-productivity state, the response of the job finding rate is similar in low- and high-productivity states. Table 5 shows the simulated standard deviations in the model without OJS (middle panel) and corroborates the findings from the IRFs. The ratio of the standard deviation for the job separation rate between the state of productivity below and above the 58th percentile is 2.38, similar to the value of 2.28 in the TVAR. However, for the job finding rate, the ratio is equal to 1.26, lower than the value of 2.01 in the TVAR.

In the version of the model without OJS, job creation responds to the two channels for state dependence (i.e., reservation productivity and non-linearities in the firm's recruitment decisions) discussed in Section 4. In Appendix F, we separate the contribution of each of these channels, showing that the effect of the reservation productivity, the primary driver for state dependence in the baseline model, is quantitatively small in absence of OJS. This absence is largely responsible for the decrease in the overall degree of state dependence in the model.[25]

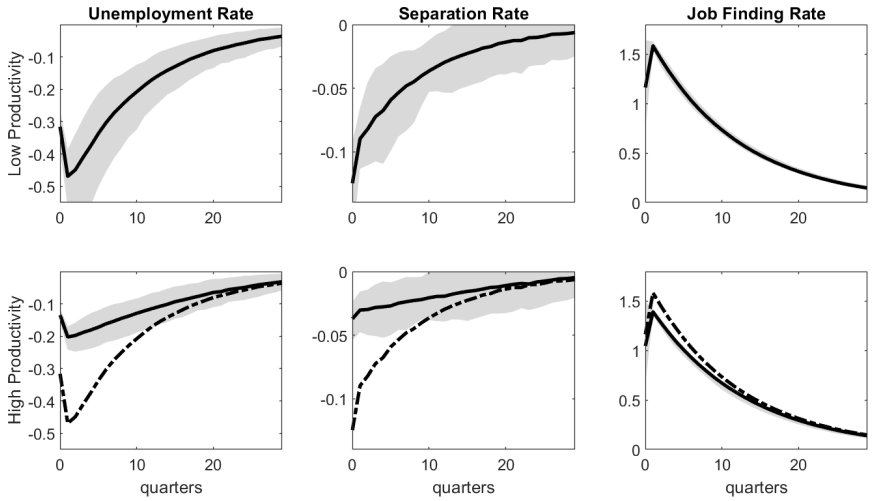
6.3 Model without OJS and with exogenous separations

As a final robustness check, we assess the extent to which the DMP model with no OJS and with exogenous separations generates state-dependent fluctuations. To abstract from endogenous job separation, we assume that workers have the same level of match-specific productivity (i.e., $x = 1$) and that the unique source of state dependence stems from the nonlinearity in the firm's choice with respect to market tightness.

Table 5 shows that the ratio of the standard deviation for the job finding rate and the

[25] Appendix F.2 discussed further the limited asymmetries in the job finding rate in this version of the model. Without OJS, the firms' reservation threshold x^r is lower and less responsive to changes in aggregate productivity a_t (Figure F.3). Intuitively, when OJS is possible, the firm decides not to create matches with low productivity, as these workers are likely to quickly move to another firm. If OJS is not feasible, the firm is more willing to take low-productivity workers as they are now more likely to obtain another productivity draw while remaining with the firm.

Figure 9: Generalized IRFs for the model without OJS.

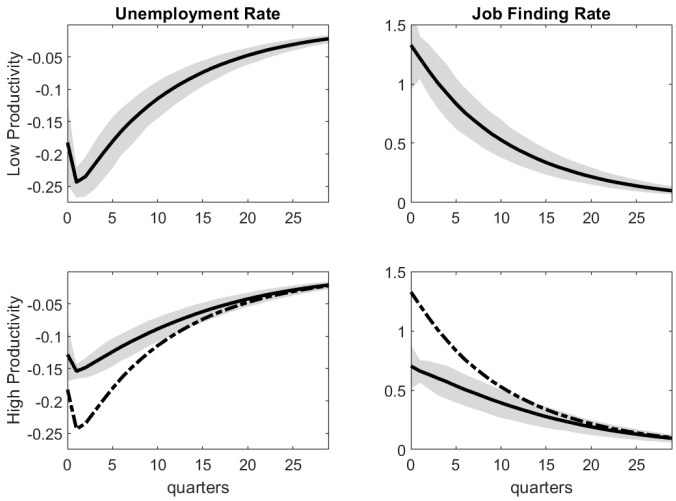


Note. The solid line represents the mean IRF value in each period. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points. In the second row, the dashed line reports the mean IRF from the first row for comparison.

unemployment rate between states of low and high aggregate productivity are equal to 1.24 and 1.52, respectively. Thus, this version of the model produces limited state dependence compared to the baseline model when considering the overall magnitude of fluctuations above and below the 58th percentile of productivity. However, focusing on state dependence at specific points of the business cycle, the 10th and 90th productivity percentiles, provides a different picture. Figure 10 shows the IRFs for the unemployment rate (left panels) and the job finding rate (right panels) at the two percentiles. The response of the job finding rate is twice as large in the 10th percentile than the 90th percentile, a difference that closely resembles the peak of the GIRFs in the TVAR.

Since the job separation rate is exogenous and workers are identical, the channel of state dependence related to reservation productivity is inactive. Thus, the sole source of state dependence in the model is the nonlinearity in firms' recruiting decisions. As shown in Table 5, under our calibration, this channel is insufficient to produce significant differences in volatilities of labor market variables over the business cycle. However, the IRFs at the specific 10th and 90th percentiles suggest that the model entails strong nonlinearities in downturns with low productivity. In other words, the state dependence generated by this version of the model is strong

Figure 10: Generalized IRFs: exogenous separation model.



Note. The solid line represents the mean IRF value in each period. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points. In the second row, the dashed line reports the mean IRF from the first row for comparison.

at specific low percentiles of the distribution of idiosyncratic productivity, but it weakens substantially at higher percentiles. This finding corroborates the results in Petrosky-Nadeau and Zhang (2017) who, using a calibration similar to Hagedorn and Manovskii (2008), show that a standard DMP model with exogenous job separation generates large nonlinearities during deep recessions.

7 Conclusion

This paper isolates important state dependence in labor market fluctuations over the business cycle. Using a TVAR model, we find that the unemployment rate and the job transition rates exhibit a stronger response to productivity shocks in periods of low aggregate productivity. A DMP model with endogenous job separation and on-the-job search captures this state dependence through the interaction between the distribution of match-specific productivity and the reservation threshold of firms for efficient matches. When calibrated to match the overall labor market volatility explained by the TVAR shocks, the model replicates 76 percent of the state dependence in the separation rate and 36 percent of that of the job finding rate.

The analysis may be extended in several dimensions. In particular, since the DMP framework is a key building block of richer models, it would be relevant to study the interaction of the DMP asymmetries with frictions from other sides of the economy. To this end, the mechanism generating state dependence may be recast in a comprehensive model that accounts for a broader range of real and nominal rigidities needed to replicate several business cycle properties in the data. The more general framework may unveil important interactions between state dependence of labor market dynamics and a broad set of macroeconomic variables. This research will prove challenging, however, because it requires a non-linear solution to a complex model. It also would be interesting to use the framework to study the design of optimal labor market reforms.[26] These investigations remain outstanding tasks for future research.

[26] In Pizzinelli et al. (2018), we study the effect of labor market reforms on aggregate fluctuations and welfare across phases of the business cycle.

8 Bibliography

Abbritti, M. and S. Fahr (2013): "Downward wage rigidity and business cycle asymmetries," *Journal of Monetary Economics*, 60, 871–886.

Altissimo, F. and G. L. Violante (2001): "The non-linear dynamics of output and unemployment in the U.S," *Journal of Applied Econometrics*, 16, 461–486.

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

Banbura, M., D. Giannone, and L. Reichlin (2010): "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, 25, 71–92.

Barattieri, A., S. Basu, and P. Gottschalk (2014): "Some Evidence on the Importance of Sticky Wages," *American Economic Journal: Macroeconomics*, 6, 70–101.

Barnichon, R. (2012): "Vacancy posting, job separation and unemployment fluctuations," *Journal of Economic Dynamics and Control*, 36, 315–330.

Barron, J. M., M. C. Berger, and D. A. Black (1997): "Employer Search, Training, and Vacancy Duration," *Economic Inquiry*, 35, 167–192.

Barron, J. M. and J. Bishop (1985): "Extensive Search, Intensive Search, and Hiring Costs: New Evidence on Employer Hiring Activity," *Economic Inquiry*, 23, 363–382.

Benigno, P., L. A. Ricci, and P. Surico (2015): "Unemployment and Productivity in the Long Run: The Role of Macroeconomic Volatility," *The Review of Economics and Statistics*, 97, 698–709.

Bianchi, F. (2013): "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," *Review of Economic Studies*, 80, 463–490.

Bianchi, F., C. L. Ilut, and M. Schneider (2018): "Uncertainty Shocks, Asset Supply and Pricing over the Business Cycle," *Review of Economic Studies*, 85, 810–854.

Bianchi, F. and L. Melosi (2017): "Escaping the Great Recession," *American Economic Review*, forthcoming.

Caggiano, G., E. Castelnuovo, and N. Groshenny (2014): "Uncertainty shocks and unemployment dynamics in U.S. recessions," *Journal of Monetary Economics*, 67, 78 – 92.

Canova, F., D. Lopez-Salido, and C. Michelacci (2012): "The Ins and Outs of unemployment: an analysis conditional on technology shocks," *Economic Journal*, 123, 515–539.

Carter, C. K. and R. Kohn (1994): "On Gibbs sampling for state space models," *Biometrika*, 81, 541–553.

Chen, C. (1998): "A Bayesian analysis of generalized threshold autoregressive models," *Statistics & Probability Letters*, 40, 15–22.

Chen, C. W. S. and J. C. Lee (1995): "Bayesian inference of threshold autoregressive models," *Journal of Time Series Analysis*, 16, 483–492.

Erceg, C. J., L. Guerrieri, and C. Gust (2005): "Can Long-Run Restrictions Identify Technology Shocks?" *Journal of the European Economic Association*, 3, 1237–1278.

Fernald, J. G. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *Federal Reserve Bank of San Francisco Working Paper Series*, 2012-19.

Ferraro, D. (2018): "The Asymmetric Cyclical Behavior of the U.S. Labor Market," *Review of Economic Dynamics*, 30, 145–162.

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

Gottfries, A. (2018): "Bargaining with renegotiation in models with on-the-job search," Working Paper.

Hagedorn, M. and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98, 1692–1706.

Hamilton, J. D. (2018): "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, 100, 831–843.

Hosios, A. J. (1990): "On The Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57, 279–298.

Justiniano, A. and G. E. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98, 604–41.

Kim, C.-J. and C. R. Nelson (1999): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, vol. 1 of *MIT Press Books*, The MIT Press.

Kohlbrecher, B. and C. Merkl (2016): "Business Cycle Asymmetries and the Labor Market," Working Paper.

Koop, G., M. Pesaran, and S. M. Potter (1996): "Impulse response analysis in nonlinear multivariate models," *Journal of Econometrics*, 74, 119–147.

Krause, M. U. and T. A. Lubik (2007): "The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions," *Journal of Monetary Economics*, 54, 706–727.

Liu, P. and H. Mumtaz (2011): "Evolving Macroeconomic Dynamics in a Small Open Economy: An Estimated Markov Switching DSGE Model for the UK," *Journal of Money, Credit and Banking*, 43, 1443–1474.

Liu, P., K. Theodoridis, H. Mumtaz, and F. Zanetti (2018): "Changing Macroeconomic Dynamics at the Zero Lower Bound," *Journal of Business & Economic Statistics*, 0, 1–14.

Lopes, H. F. and E. Salazar (2006): "Bayesian Model Uncertainty In Smooth Transition Autoregressions," *Journal of Time Series Analysis*, 27, 99–117.

Merz, M. (1999): "Heterogeneous job-matches and the cyclical behavior of labor turnover," *Journal of Monetary Economics*, 43, 91–124.

Molloy, R., C. L. Smith, R. Trezzi, and A. Wozniak (2016): "Understanding Declining Fluidity in the U.S. Labor Market," *Brookings Papers on Economic Activity*, 47, 183–259.

Mortensen, D. T. and C. A. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, 397–415.

Moscarini, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73, 481–516.

Moscarini, G. and K. Thomsson (2007): "Occupational and Job Mobility in the US*," *Scandinavian Journal of Economics*, 109, 807–836.

Mueller, A. I. (2017): "Separations, Sorting, and Cyclical Unemployment," *American Economic Review*, 107, 2081–2107.

Mukoyama, T. (2014): "The cyclicality of job-to-job transitions and its implications for aggregate productivity," *Journal of Economic Dynamics and Control*, 39, 1–17.

Mumtaz, H. and F. Zanetti (2012): "Neutral Technology Shocks And The Dynamics Of Labor Input: Results From An Agnostic Identification," *International Economic Review*, 53, 235–254.

——— (2015): "Labor Market Dynamics: A Time-Varying Analysis," *Oxford Bulletin of Economics and Statistics*, 77, 319–338.

Neftci, S. N. (1984): "Are Economic Time Series Asymmetric over the Business Cycle?" *Journal of Political Economy*, 92, 307–28.

Panagiotidis, T. and G. Pelloni (2007): "Nonlinearity in the Canadation and U.S Labor Markets: Univariate and Multivariate Evidence from a Battery of Tests," *Macroeconomic Dynamics*, 11, 613–637.

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

Petrosky-Nadeau, N. and L. Zhang (2017): "Solving the Diamond-Mortensen-Pissarides model accurately," *Quantitative Economics*, 8, 611–650.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory, 2nd Edition*, vol. 1 of *MIT Press Books*, The MIT Press.

Pizzinelli, C., K. Theodoridis, and F. Zanetti (2018): "State Dependence in Labor Market Fluctuations: Evidence, Theory, and Policy Implications," Economics series working papers, University of Oxford, Department of Economics.

Postel-Vinay, F. and J.-M. Robin (2004): "To match or not to match?: Optimal wage policy with endogenous worker search intensity," *Review of Economic Dynamics*, 7, 297–330.

Ravenna, F. (2007): "Vector autoregressions and reduced form representations of DSGE models," *Journal of Monetary Economics*, 54, 2048–2064.

Sedláček, P. (2014): "Match efficiency and firms' hiring standards," *Journal of Monetary Economics*, 62, 123–133.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25–49.

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

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

Sims, C. A., D. F. Waggoner, and T. Zha (2008): "Methods for inference in large multiple-equation Markov-switching models," *Journal of Econometrics*, 146, 255–274.

Sims, C. A. and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 96, 54–81.

Tauchen, G. (1986): "Finite State Markov Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.

Thomas, C. and F. Zanetti (2009): "Labor market reform and price stability: An application to the Euro Area," *Journal of Monetary Economics*, 56, 885–899.

Uhlig, H. (2004): "Do Technology Shocks Lead to a Fall in Total Hours Worked?" *Journal of the European Economic Association*, 2, 361–371.

Zanetti, F. (2011): "Labor market institutions and aggregate fluctuations in a search and matching model," *European Economic Review*, 55, 644–658.

Appendix

A Data sources

The TVAR analysis uses the following time series: average labor productivity, the unemployment rate, total hours worked, the job finding rate, and the separation rate. Labor productivity is the series adjusted for capital utilization developed by developed by Fernald (2014). The unemployment rate and the total hours series work were downloaded from Federal Reserve Bank of St. Louis Database (FRED). The monthly job separation and job finding probabilities are computed following the continuous-time adjustment proposed by Shimer (2012), requiring the time series for unemployment with a duration of 5 weeks or less, which we also download from FRED. While we leave the details to the original paper, the essence of continuous time adjustment is to estimate the transition probabilities between unemployment and employment as discrete time probabilities derived from continuous-time hazard rates that are assumed to be constant within each month. This method controls for the bias of simultaneously estimating two related discrete probabilities. We used the original series provided on Robert Shimer's web page from 1950 to 2007, and extend them using the BLS data until 2014. The monthly series are then averaged over the respective quarters.

B TVAR: details,additional figures, and sensitivity analysis

B.1 Priors

The formulation of priors follows Banbura et al. (2010) and the same prior moments have been used for the parameters in both regimes. To be precise, it is assumed that the prior distribution of the VAR parameter vector has a Normal-Wishart conjugate form

$$b|\Sigma \sim N(b_0, \Sigma \otimes \Omega_0),\ \Sigma \sim IW(v_0, S_0). \tag{21}$$

where b is obtained by stacking the columns of the matrix of the autoregressive coefficients B. The prior moments of b are given by

$$E[(B_k)\, i,j] = \begin{cases} \delta_i & i=j, k=1 \\ 0 & \text{otherwise} \end{cases},\ Var[(B_k)\, i,j] = \lambda \sigma_i^2/\sigma_j^2,$$

and as it is explained by Banbura et al. (2010) they can be constructed using the following dummy observations

$$Y_D = \begin{pmatrix} \frac{diag(\delta_1\sigma_1 ... \delta_N\sigma_N)}{\lambda} \\ 0_{N\times(K-1)N} \\ ..... \\ diag\,(\sigma_1 ... \sigma_N) \\ ..... \\ 0_{1\times N} \end{pmatrix} \text{ and } X_D = \begin{pmatrix} \frac{J_K \otimes diag(\sigma_1 ... \sigma_N)}{\lambda} \\ 0_{N\times NK} \\ ..... \\ 0_{1\times NK} \end{pmatrix} \tag{22}$$

where $J_K = diag\,(1, 2, ..., K)$ and $diag$ denotes the diagonal matrix. The prior moments of (21) are just functions of Y_D and X_D, $B_0 = Y_D X_D' \,(X_D X_D')^{-1}$, $\Omega_0 = (X_D X_D')^{-1}$, $S_0 = (Y_D - B_0 X_D)\,(Y_D - B_0 X_D)'$ and $v_0 = T_D - NK$. Finally, the hyper-parameter λ controls

the tightness of the prior.

The values of the persistence – δ_i – and the error standard deviation – σ_i – parameters of the AR(1) model are obtained from its OLS estimation (as in Mumtaz and Zanetti, 2012, 2015)). Sensitivity analysis reveals that the results are robust to different selections of VAR lags. As we work with a five variables VAR and eight lags, we follow Canova et al. (2012) and select a value for λ that implies fast lag decay towards zero ($\lambda = 0.25$). It is shown in the next section that the results are robust when "looser" priors are used.

Finally, z^* is assumed to be normally distributed with zero mean and standard deviation σ_{z^*} calibrated to deliver an MCMC acceptance rate between 20% and 40%.

B.2 Max Variance VAR Shock Identification Method

Following Uhlig (2004), the benchmark identification scheme employed in this study amounts to finding the shock that explains most of the variation of the (adjusted for utilisation) labor productivity (Fernald, 2014) series over business cycle frequencies. The scheme consists of finding an orthonormal rotation matrix Q of the orthogonalized shocks that maximizes the sum of the forecast error variance of labor productivity from horizons 0 to h. The scheme is applied separately for the observation in each regime. Therefore, for simplicity we drop the regime specific notation $i = 1, 2$ in the following description. For a regime-specific VAR model, the conventional moving-average (MA) notation is

$$Z_t = \sum_{j=0}^{\infty} \tilde{B}_j v_{t-j},$$

where v_t's are reduced-form shocks, $\tilde{B}_0 = I_N$, and $\tilde{B}_j = \sum_{s=0}^{j} \tilde{B_{s-j}} B_j$ (with B_j being the reduced-form coefficients, such that $B_j = 0$ for $j > K$. Using the MA representation, the forecast error variance at horizon h is

$$\Sigma_h = \sum_{j=0}^{h-1} \tilde{B}_j A A' \tilde{B}_j'$$

where $\Sigma_v = AA'$ is the variance-covariance matrix of reduced form shocks, and A is a lower-triangular matrix obtained through a Choleski decomposition. Assume (without loss of generality) that labor productivity is the first series in the VAR and that the shock driving labor productivity between 0 and 40 quarters is ordered first. In this case the identification is achieved by finding the first column of matrix Q that solves the following maximisation problem:

$$\arg\max_{Q_1} e_1' \left[\sum_{h=0}^{H} \sum_{j=0}^{h-1} \tilde{B}_j A \; Q_1 \; Q_1' A' \tilde{B}_j' \right] e_1$$

such that $Q_1 Q_1' = 1$ and e_1 is a N-by-1 vector with 1 in the first entry and 0 in all the others. As shown by Uhlig (2004), identification of the productivity shock only requires finding the first column of Q (i.e. Q_1). Moreover, the maximization can be re-written as an eigenvalue eigenvector problem and a solution can be easily obtained. The identification of the productivity shock under this scheme requires no further restrictions.

B.3 Generalized Impulse Response Functions

The generalized impulse responses are calculated using parallel computing technology (MATLAB Distributed Computing Server/Parallel Computing Toolbox on 32 cores). Similar to Koop et al. (1996), the exact simulation steps are as follows:

1. We draw 1×48 structural shocks ω_t from the standard normal distribution (where $t = 1, .., 40$)
2. We simulate the model using the shocks from step 1, we denote the simulated data by y_t
3. We simulate the model using again from step 1 but now we increase the value of the structural shock of interest in period 1 by an amount x, namely

$$\tilde{\omega}_{j,1} = \omega_{j,1} + x \tag{23}$$

   where $x = 1$ standard deviation in our case. We denote the data obtained from this simulation by $\tilde{y}_t$
4. Steps 1 to 3 are repeated histories Th (all starting points in the data)
5. The GIRF is calculated as follows

$$GIRF = \frac{1}{Th} \sum_{i=1}^{Th} \left(\tilde{y}_t^i - y_t^i \right) \tag{24}$$

Steps 1-5 are implemented for all posterior draws.

B.4 Additional figures from the TVAR

Figure B.1: Prior and posterior distributions of the labor productivity threshold value.

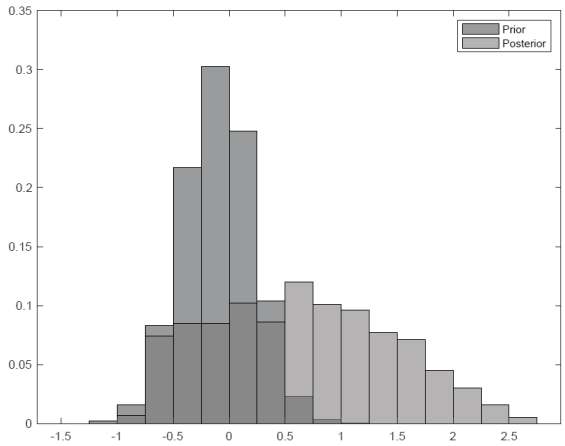
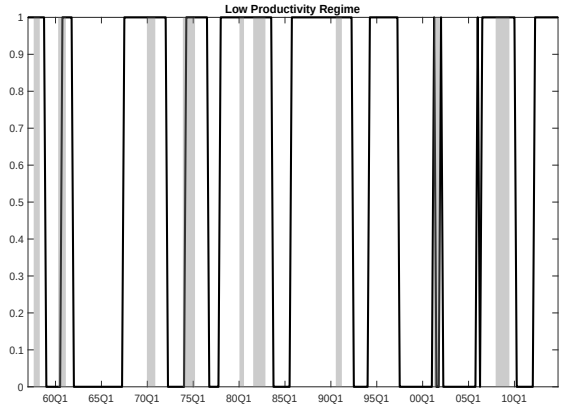
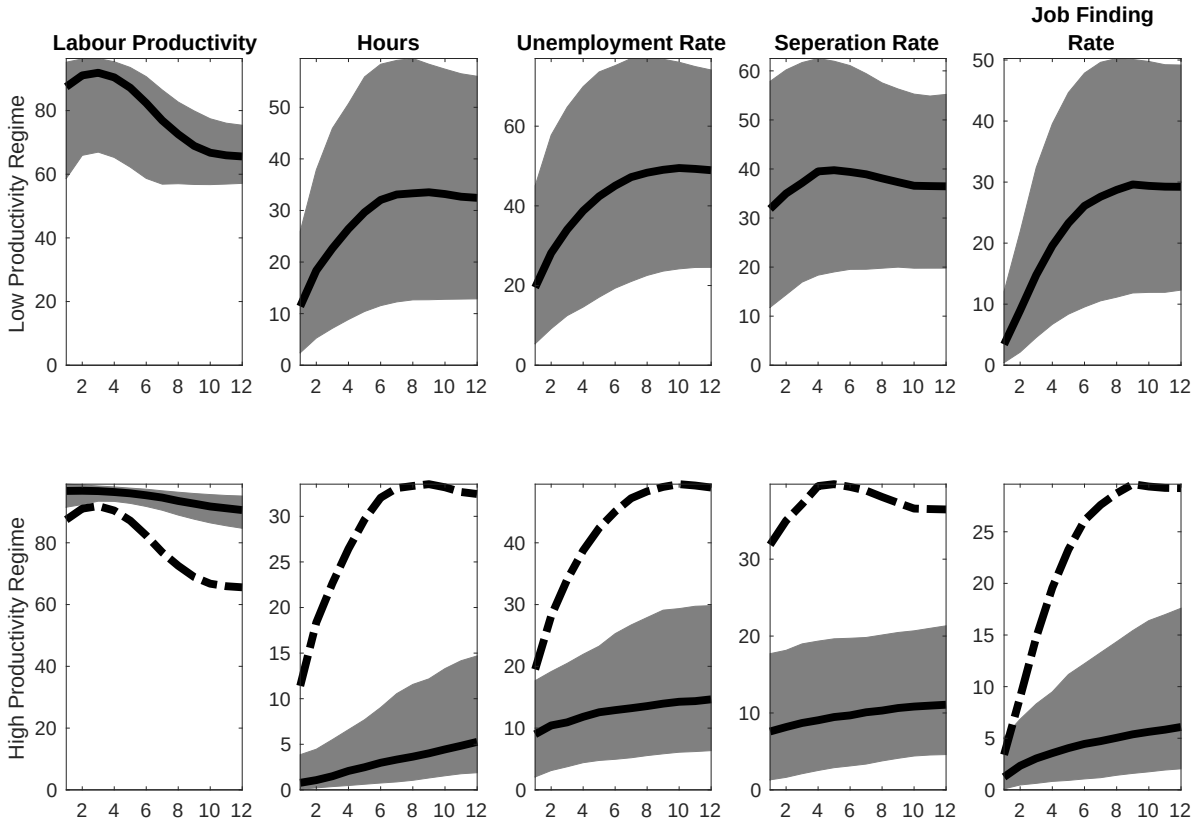


Figure B.2: Regimes based on the median of the posterior of the threshold.



Note. The solid line, switching from 0 to 1, represents the current regime as identified by the median of the threshold's posterior distribution. The value of 1 corresponds with the low-productivity regime. The shaded grey areas comprise NBER recession dates.

Figure B.3: Forecast Error Variance Decomposition from the TVAR



Note. The solid line reports the median fraction of the forecast error variance each variable that is explained by the identified productivity shocks. The shaded area comprises the 16th-84th percentile range. In the second row, the dashed line reports the fraction from the low-productivity regime for comparison.

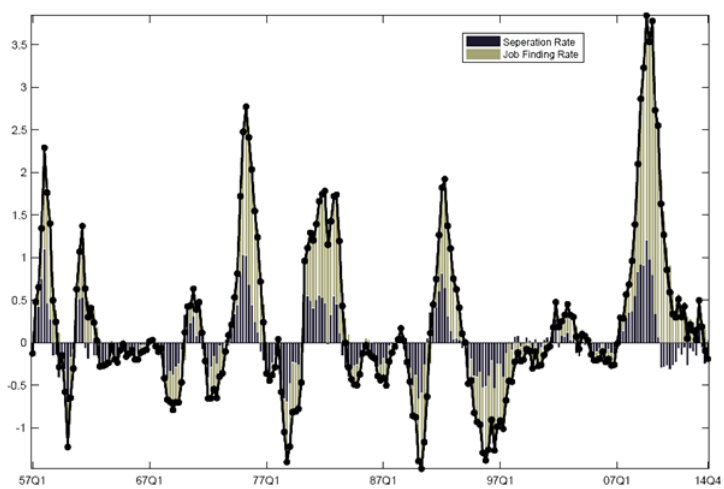
B.5 The steady-state approximation of the unemployment rate in the TVAR: historical decomposition

Historical decomposition of the approximation in the TVAR

We examine the role of productivity shocks for movements in the job finding and separation rates using the steady-state unemployment rate approximation proposed by Shimer (2012): $u_t^{ss} = SR_t/(SR_t + JFR_t)$, where SR_t and JFR_t are the separation rate and the job finding rate at time t, respectively. We focus on the component of these variables that is explained by the productivity shock, and thus use input from the TVAR. We denote with a hat the component of each variable that is explained by productivity shocks and with a bar the long-run mean of the variable. For instance, $\hat{SR}_t$ is the productivity-explained counterfcactual separation rate for its mean. As a result $\hat{u}_t^{ss} - \bar{u}^{ss} = \hat{SR}_t/(\hat{SR}_t + \hat{JFR}_t) - \overline{SR}/(\overline{SR} + \overline{JFR})$ is the explained deviation of steady-state approximation. We isolate the individual contribution of $\hat{SR}_t$ and $\hat{JFR}_t$ to $\hat{u}_t^{ss}$ by holding one of them constant at the long-run mean while allowing the other component to vary.[27]

Figure B.4 plots the movements in the unemployment rate that are explained by changes in productivity in the job finding rate (light grey bars) and the job separation rate (dark grey bar). Both transition rates contribute to changes in the unemployment rate. The job finding rate explains approximately two thirds of the total fluctuations in unemployment. The separation rate explains approximately one third of fluctuations, suggesting that job destruction also plays an important part in the business cycle dynamics of unemployment.

Figure B.4: Historical decomposition of productivity-induced fluctuations of steady-state approximation of the unemployment rate.



Note. The black line reports the deviations from trend of the steady-state approximation of the unemployment rate that are explained by productivity shocks in the estimated TVAR. The light grey bars report the part of those deviations that are explained by deviations of the job finding rate, while the dark grey bars report deviations explained by the separation rate.

[27] As the TVAR entails transition across aggregate regimes, where each variable has a regime-specific mean, and because productivity enters the model in de-trended form, the construction of the series outline above is actually done separately for each regime.

B.6 Assessing the contributions of the job finding rate and the separation rate to the state dependence of unemployment.

To assess the relative importance of each transition rate for the state dependence of the unemployment rate, we combine the steady-state approximation of the unemployment rate with the IRFs from the TVAR. We compute IRFs for the steady-state approximation in low- and high-productivity points using the median of the IRFs of the separation rate and job finding rate in each regime (reported in Figure 2). We then derive counterfactual IRFs where one of the two transition rates responds as in the high-productivity regime and the other responds as in the low-productivity regime. These counterfactual cases shed light on the fraction of the state dependence in the unemployment response accounted for by each transition rate at a given time horizon. Below we outline more precisely how these are constructed:

- We compute the mean steady-state unemployment rate using the mean value of the separation and job finding rates: i.e. $\overline{u}^{ss} = \overline{SR}/(\overline{SR} + \overline{JFR})$.
- Let $\Delta JFR^i(h)$ be the mean posterior IRF of the JFR at horizon h in regime i. Similarly define $\Delta SR^i(h)$.
- Then we define the regime-i IRF at horizon h as

$$\Delta u^{ss,i}(h) = \frac{(\overline{SR} + \Delta SR^i(h))}{(\overline{SR} + \Delta SR^i(h) + \overline{JFR} + \Delta JFR^i(h))} - \overline{u}^{ss}$$

.

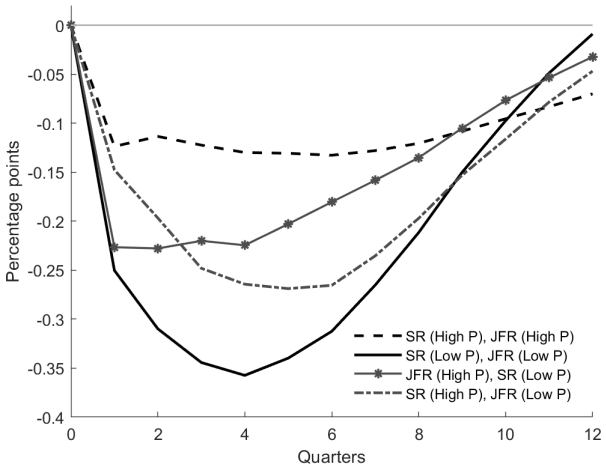
- Then we can also construct "intermediate" IRFs where either $\Delta JFR^i(h)$ or $\Delta SRR^i(h)$ is replaced with $\Delta JFR^j(h)$ or $\Delta SR^j(h)$ where $j \neq i$. For instance,

$$\Delta u^{ss,ij}(h) = \frac{(\overline{SR} + \Delta SR^i(h))}{(\overline{SR} + \Delta SR^i(h) + \overline{JFR} + \Delta JFR^j(h))} - \overline{u}^{ss}$$

.

Figure B.5 shows the resulting IRFs. The solid black line shows the IRF obtained from the response of both transition rates in the high-productivity regime. The dashed line shows the equivalent IRF for both transition rates in the low-productivity regime. Thus, the difference between the solid and dotted lines represents the average degree of state dependence, which is similar to the different responses of unemployment across the two regimes reported in Figure 2 (third row). The starred (dash-dotted) line shows the IRF in which the separation rate (job finding rate) responds as in the low-productivity regime and the job finding rate (separation rate) as in the high-productivity regime. The difference between the dashed line and the starred (dash-dotted) line thus indicates the contribution of the separation (job finding) rate to the total state dependence of unemployment at a given time horizon. For instance, in the first quarter in which the shock hits (Quarter 1 in the figure), the starred line covers approximately 80 percent of the difference between the solid and the dashed IRFs. Hence, the separation rate explains most of the state dependence in the unemployment rate in the first quarter. However, four quarters after the shock, the contribution of the two rates is very similar, as the starred and dash-dotted lines are both approximately mid-distance between the dashed and solid ones. By the eighth quarter, the job finding rate accounts for approximately the entire difference between the response of unemployment in the two regimes.

Figure B.5: Counterfactual IRFs for the approximation of the steady-state unemployment rate



Note. The solid-black line reports the constructed IRF of steady-state unemployment approximation where both transition rates follow the IRFs from the low-productivity regime. The dashed line reports the IRF based in which both transition rates use IRFs of the high-productivity regime. The dash-dotted (starred) line reports the intermediate case in which the IRF for the separation rate (job finding rate) uses the high-productivity regime while the IRF for the job finding rate (separation rate) uses the low-productivity regime. See Appendix B.6 for the details on the construction.

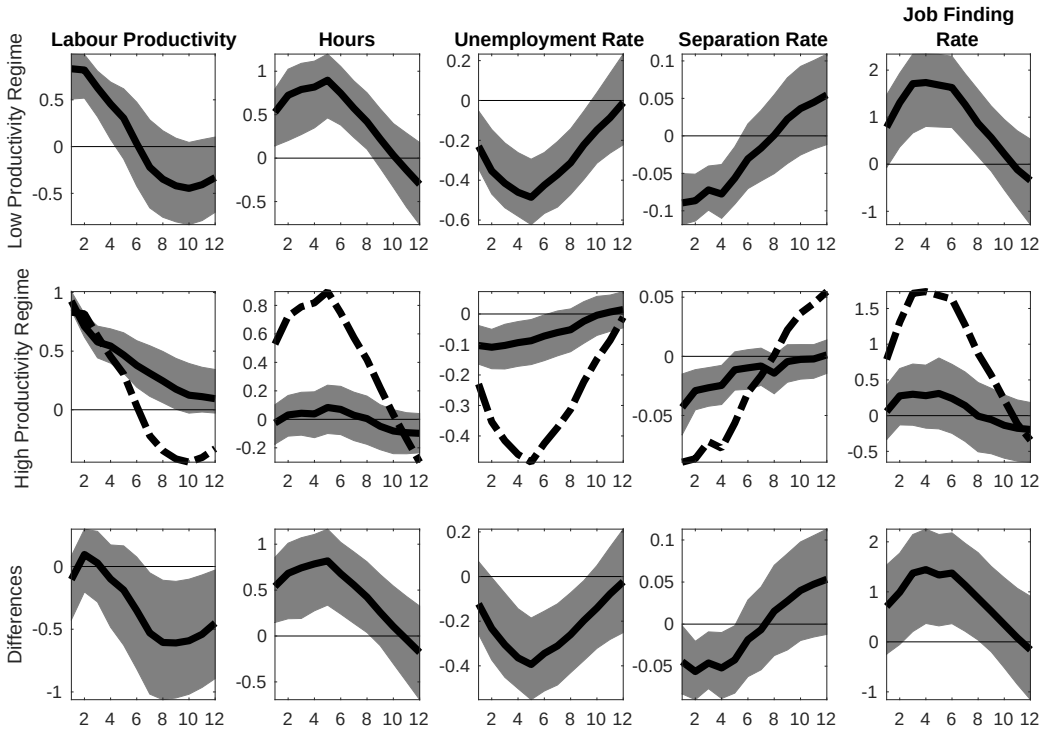
B.7 Sensitivity analysis of the TVAR

Figures B.6-B.10 report the IRFs from a battery of robustness checks. In Figure B.6 we estimate the TVAR with looser priors on the values of the matrix of reduced-form coefficients. The parameter determining the tightness of the prior distribution is changed from $\lambda = 0.25$ to $\lambda = 0.5$.[28] Figure B.7 estimates the TVAR using 4 lags instead of eight (we also experimented with 6 lags with similar results). Figure B.8 uses a raw series for labor productivity, not adjusted for capital utilization as proposed by Fernald (2014). Finally, Figures B.10 and B.9 report the results of the baseline TVAR estimated on particular time sub-periods: the pre-Great Recession sample (1950-2007), and the post-Volcker period (1980-2014).

The results generally hold well in all the robustness exercises. The only difference worthy of note is the response of the job finding rate in the post-Volcker sample. Here, the IRF for the high-productivity regime is not as muted as in the baseline results. However, it also does not exhibit a hump-shaped trajectory. Instead, it responds immediately to the shock but with low persistence. The response in the low-productivity regime on the other hand is smaller in magnitude compared to the baseline case. The main caveat in interpreting the robustness analysis on reduced time series is that the estimates are likely noisier, as the TVAR model requires large samples in order to estimate a tighter posterior distribution for the parameters and the threshold.

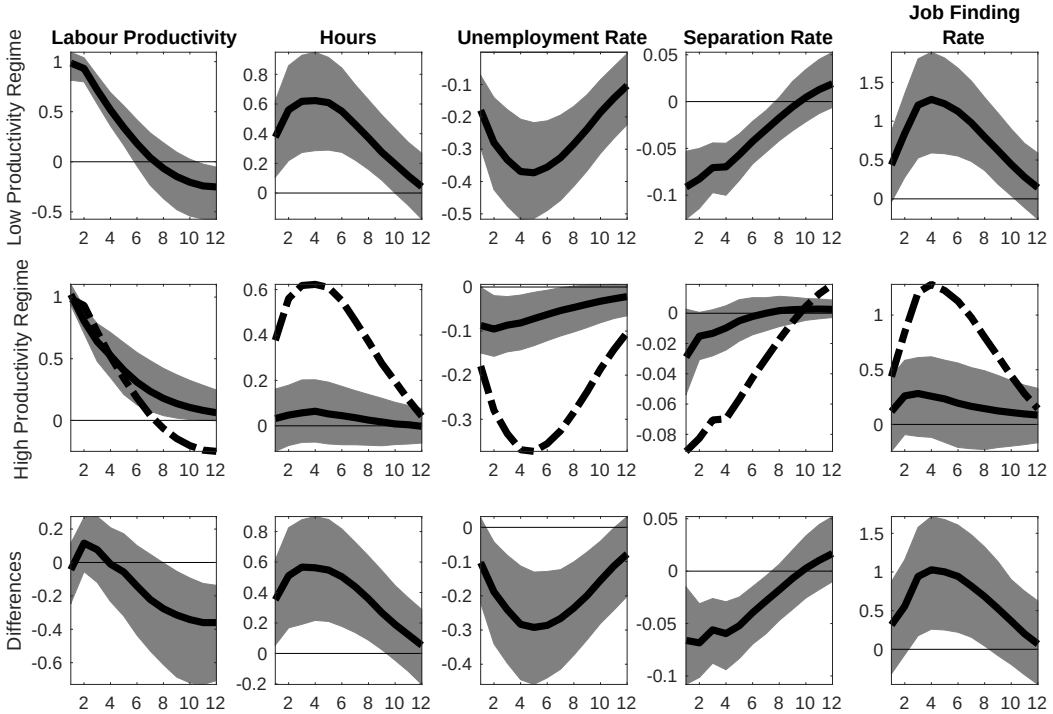
[28] As it is discussed by Balleer (2012) and Canova et al. (2012), the large dimension of the reduced-form parameter vector calls for tight priors (given the time span of the available data) that "shrink" the number of the estimated parameter towards zero relatively "fast".

Figure B.6: TVAR Model with looser priors: Impulse Responses



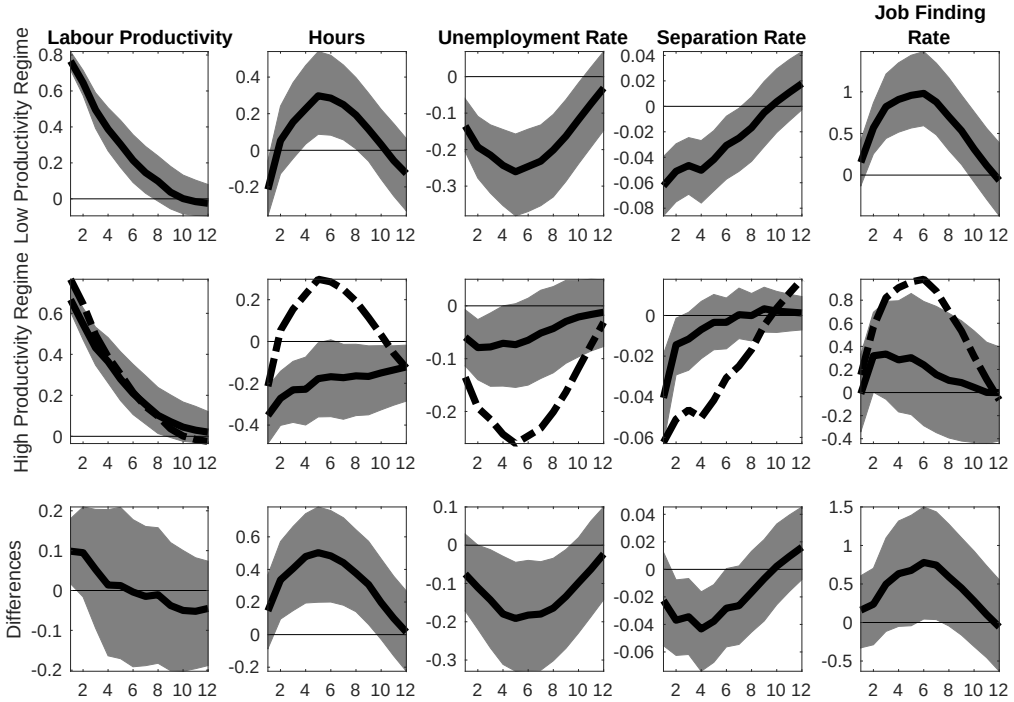
Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The red line in the second row is the pointwise median from the Low Productivity Regime.

Figure B.7: TVAR Model with four lags: Impulse Responses



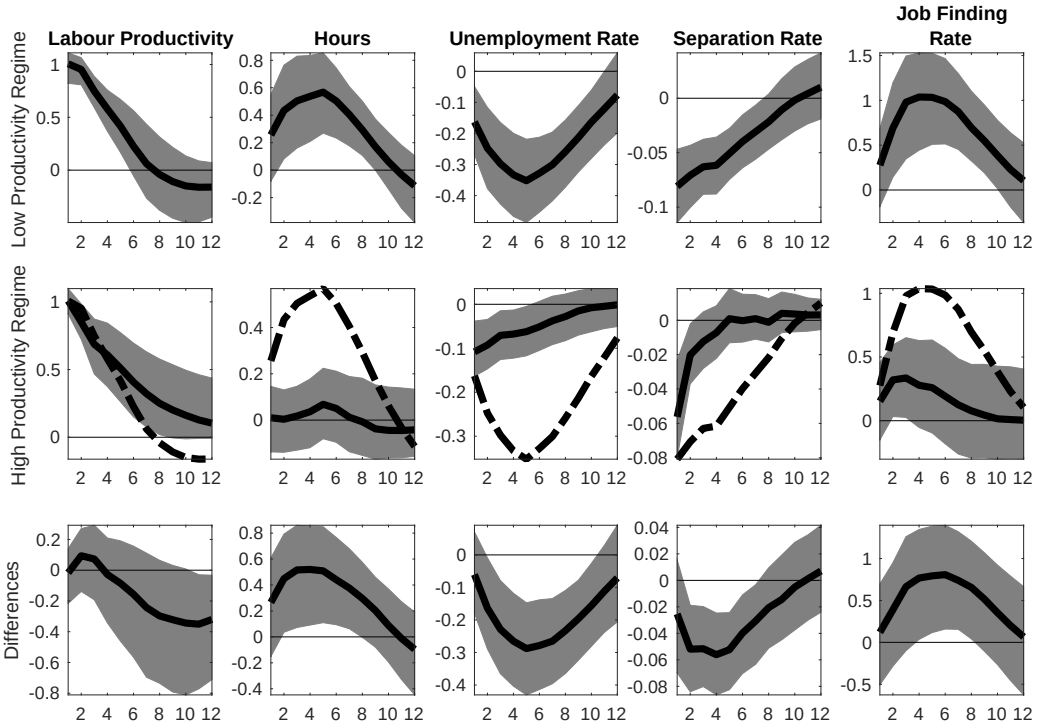
Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The red line in the second row is the pointwise median from the Low Productivity Regime.

Figure B.8: TVAR Model estimated using the labor productivity series not adjusted for capital utilization: Impulse Responses



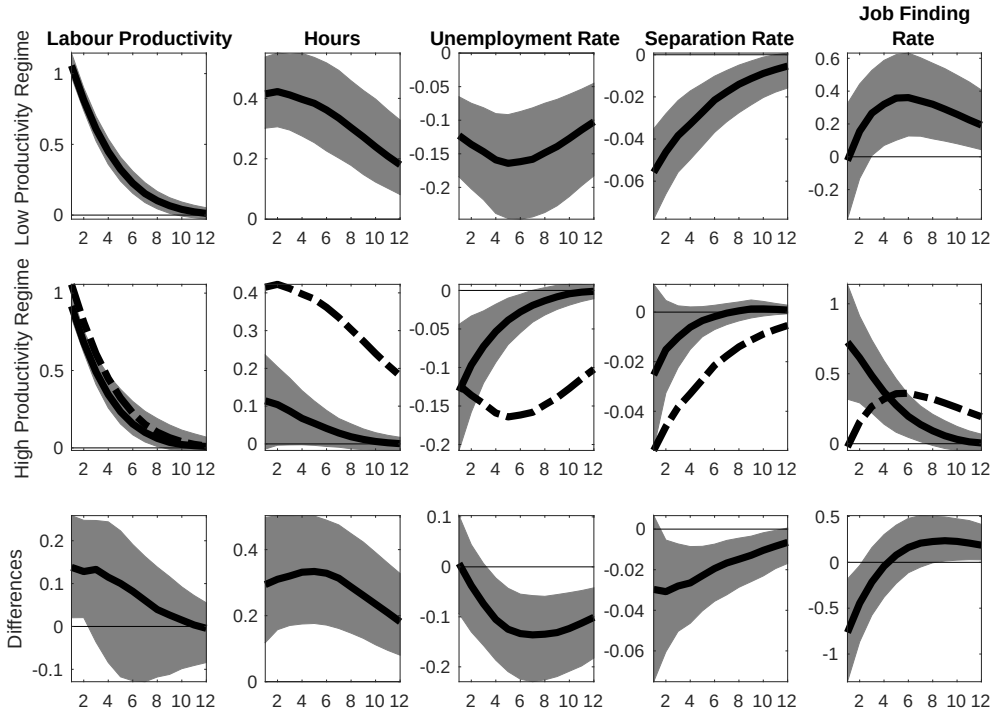
Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The red line in the second row is the pointwise median from the Low Productivity Regime.

Figure B.9: TVAR Model estimated one reduced time sample 1950:I-2007:4: Impulse Responses



Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The red line in the second row is the pointwise median from the Low Productivity Regime.

Figure B.10: TVAR Model estimated one reduced time sample 1980:I-2014:4: Impulse Responses



Note. The solid black line represents the pointwise median IRF, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. Horizontal axes report quarters, vertical axes report percentage deviations from the trend. The third row displays the posterior distribution of the difference between impulse responses for low (first row) and high (second row) states of productivity. The red line in the second row is the pointwise median from the Low Productivity Regime.

C Markov-Switching VAR (MS-VAR) model

C.1 The Specification

The functional form of the MS-VAR is very similar to the benchmark T-VAR model:

$$Z_t = c_{\xi_t} + \sum_{k=1}^{K} B_{k,\xi_t} Z_{t-k} + \Omega_{\xi_t}^{1/2} \varepsilon_t \tag{25}$$

where the regime switches in the model are governed by the latent variable ξ_t that follows a first order Markov chain with two regimes and transition matrix $P = \begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}$, where $p_{ij} = p\left(\xi_t = j | \xi_{t-1} = i\right)$. Z_t consists of the same variables as in the T-VAR and we also use the same number of lags.

C.2 Priors

C.2.1 Coefficient and Covariance Matrices

The formulation of the priors follows that of the TVAR.

C.2.2 Regime Probabilities

As standard in the literature (Liu and Mumtaz (2011), Bianchi (2013), Bianchi et al. (2018), Bianchi and Melosi (2017), Liu et al. (2018) among others), the prior distribution assumed for the columns of the matrix P is the Dirichlet distribution ($D(\alpha)$), where α has been calibrated to ensure that there are enough observations for regime so the parameters can be estimated.

C.3 Estimation

C.3.1 Steps

The estimation of the model consists of two stages. In the first stage, the posterior mode of the above model is obtained via maximisation techniques. As this step is computational demanding, we follow the methodology in Sims et al. (2008) and we 'break' the maximisation of the objective function into three blocks (autoregressive, covariance and probabilities). Once the posterior mode is obtained, we 'launch' the following Gibbs Sampler:

1. Conditional on P, Σ_{ξ_t} and ξ_t, sample B_{ξ_t} from its conditional Normal posterior distribution
2. Conditional on P, B_{ξ_t} and ξ_t, sample Σ_{ξ_t} from its conditional Inverse-Wishart posterior distribution
3. Conditional on $\Sigma_{\xi_t}, B_{\xi_t}$ and ξ_t, sample P from its conditional Dirichlet posterior distribution
4. Conditional on $\Sigma_{\xi_t}, B_{\xi_t}$ and P sample ξ_t from its conditional posterior distribution

As it is explained in Kim and Nelson (1999), the conditional distribution of ξ_t is derived using techniques similar to Carter and Kohn (1994) algorithm. This step requires the use of the Kalman filter and smoother.

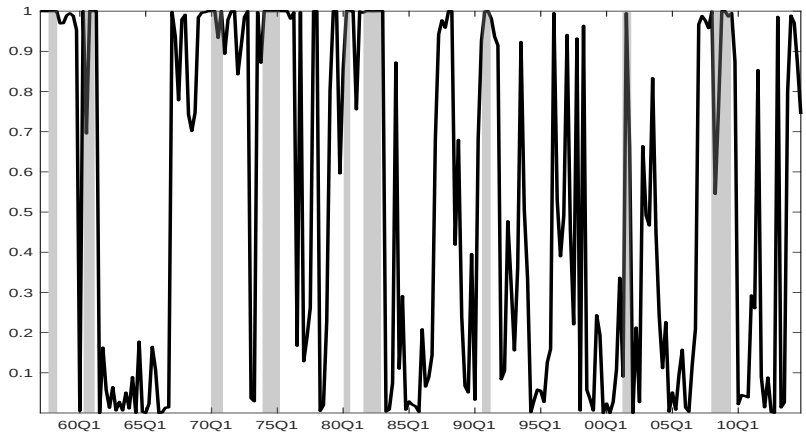
C.3.2 Label Switching

As the likelihood of the model does not depend on the specific label that we attach to a regime, each regime can change during the estimation of the model, and thus the posterior distribution of the model may be multi modal. To avoid this possibility, we restrict the mean of the productivity implied by the MS-VAR model in the Regime 1 (μ_1) to be lower than the mean of the productivity in Regime 2 ($\mu_1 < \mu_2$). In other words, Regime 1 is labelled as the low-productivity regime.

C.4 Results

Figure C.1 illustrates the evolution of the low-productivity regime. Please note in the MS-VAR case, this is inference of a latent variable using filtering techniques. Despite the very different way of estimating this process, it is very similar to the one derived using the TVAR model.

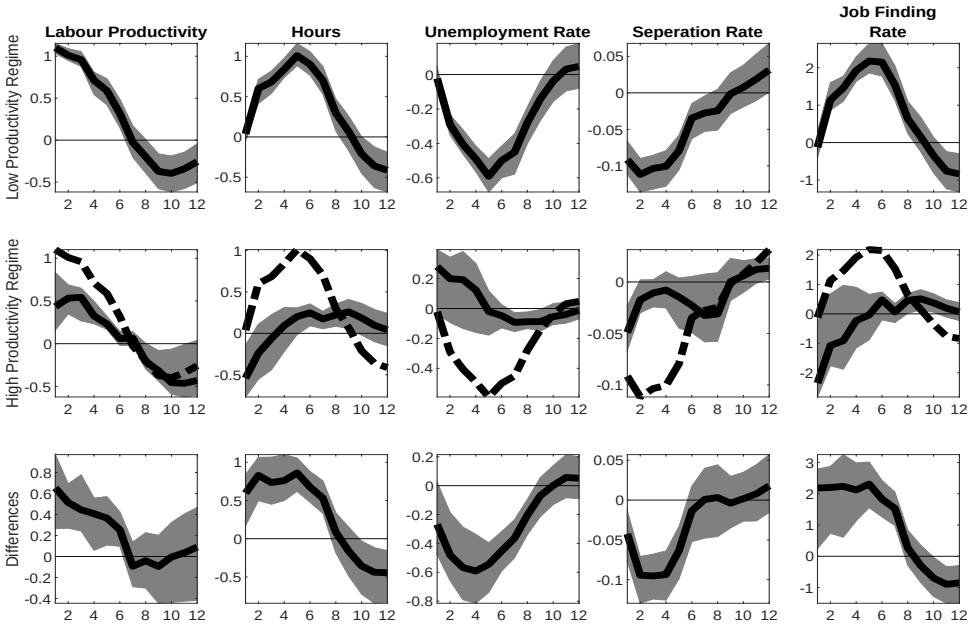
Figure C.1: Regime Probabilities



Note. The black line illustrates the evolution of the low productivity regime. The shaded bars represent the recession dates.

Figure C.2 plots the responses of the variable in our system after a productivity shock identified as in the TVAR case. Despite the use of a methodologically different model, the inference are similar our benchmark TVAR model. The responses of the variables to a productivity shock are stronger in the low-productivity regime.

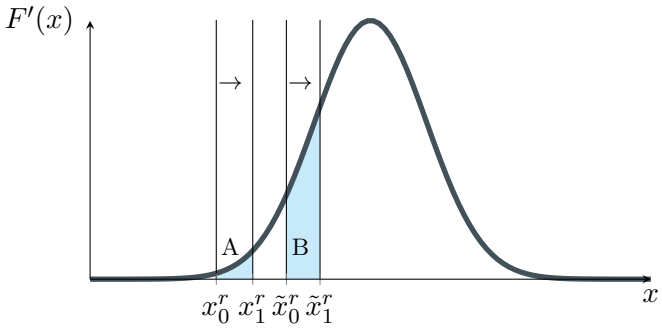
Figure C.2: Impulse Responses



Note. The solid line represents the pointwise median impulse response function, and the shaded area is the corresponding 16th and 84th percentiles of the posterior distribution. The horizontal axes are in quarters, the vertical axes are in percentage deviations from the trend. The third row displays the posterior distribution of the difference between Low (first row) and High (second row) Productivity Regime Impulse Responses. The red line in the second row is the pointwise median from the Low Productivity Regime.

D Additional diagrams

Figure D.1: Illustrative diagram of the mechanism driving asymmetries with respect to state of the economy in the job finding rate.



E Model: additional information

E.1 Uniqueness of productivity threshold $x^r(a_t)$

Assuimng the threshold for OJS $x^s(a)$ lies above the threshold $x^r(a)$, and setting $S^s(a,z) = 0$, we can rearrange (13) into

$$b + k^s = a_t x_t^r + \beta \mathbb{E}_t \Big\{ (1-s)\Big(1 - p(\theta_t)\overline{F(x_{t+1}^r)}\Big)\Big((1-\lambda)S(a_{t+1}, x) + \lambda \int S(a_{t+1}, x')dF(x')\Big)\Big\}$$

The first term on the RHS is continuous, strictly increasing, and bounded below in x over the support $(0,\ x_H]$. For a given θ_t, the term $\Big(1 - p(\theta_t)(1 - F(x_{t+1}^r))\Big)$ is continuous and strictly increasing in x. By the properties of $S(a,x)$, the second term on the RHS is bounded below by 0 and is continuous and weakly increasing in x. Therefore, the RHS is bounded below by 0, strictly increasing and continuous in x. These conditions are sufficient for the uniqueness of the value $x^r(a_t)$. Furthermore, as the LHS of the equation is constant, and $S(a,x)$ is increasing in a, then $x^r(a)$ must be decreasing in a.

E.2 Solution of the model and simulated moments

We solve the model nonlinearly using an iterative procedure. To generate an accurate solution, we set the number of grid points of the state space to 45 for the discretized AR(1) process a_t, following the approach described in Tauchen (1986), and 800 for the individual productivity x. The iteration starts with a guess for the policy function of the market tightness $\theta^{(0)}(a)$. Using the guess, we compute the match surplus for all values of x and a and derive the individual productivity threshold $x^r(a)$ and the OJS threshold $x^s(a)$. Using these results, $\theta^{(1)}(a)$ is computed through the free entry condition and used again to compute the surplus. The process is repeated until the norm of $||\theta^{(0)}(x) - \theta^{(1)}(x)||$ is below a chosen critical value.[29] Using the result from Pissarides (2000), under a linear production function, hiring and layoff decisions do not depend on the aggregate level of employment or on the distribution of individual productivity. Hence, the only relevant state variable for the policy function is the aggregate productivity factor.

To obtain business cycle statistics, we run 1,000 simulations of the model. Each simulation comprises 1,380 monthly periods. After discarding the first 600 periods, we take quarterly averages of all the simulated series to create a time series of the same length as the data. For each simulation we compute the relevant business cycle moments after taking logs and filtering the series with the Hamilton (2018) filter. The simulated business cycle statistics are the average of each moment across the simulations.

Table E.1 compares the empirical targets for the calibration with corresponding simulated moments for each version of the model.

E.3 Computation of Generalized IRFs for the model

We resort to numerical simulations of the model to produce the response of the variables at different points in the state space. To implement the computation, it is critical to establish

[29] The value we use is 10^{-8}.

Table E.1: Empirical targets and simulated moments from the baseline model.

		Targets	Baseline Model
SR	Mean	3.3	3.3
SR	Standard Deviation	1.7	1.7
SR	Autocorrelation	0.79	0.81
JFR	Mean	45	44
JFR	Standard Deviation	3.3	3.3
J2J	Mean	3.2	3.2

Note. The first column reports the targets for the moments of the separation rate (SR), the job finding rate (JFR), and the job-to-job rate. The targeted standard deviations of the separation rate and the job finding rate are derived by taking the share of the stationary variance explained by the TVAR (weighted across regimes) multiplied by the unconditional variance. Means and standard deviations are multiplied by 100 for presentational purposes.

Table E.2: Simulated moments for a broader set of labor market statistics in the baseline model.

	p	U	JFR	SR	E	V	V/U	Y	JJR
σ_X*100	2.17	0.82	3.28	3.03	0.17	6.17	17.02	0.82	0.42
Corr(p_t, X_t)	1.00	-0.95	0.99	-0.95	0.95	0.97	0.99	1.00	0.97
Corr(X_t, X_{t-1})	0.88	0.90	0.89	0.82	0.89	0.81	0.88	0.90	0.85

Note. The table reports cyclical statistics for average labor productivity (p), the unemployment rate (U), the job finding rate (JFR), the job separation rate (SR). $\bar{X}$ is the long-run mean of the variable, σ_X the standard deviation, Corr(p_t, X_t) the correlation with productivity, and Corr(X_t, X_{t-1}) the autocorrelation. The model moments are computed as the mean of 1,000 simulations in which the simulated data are treated in the same way as the empirical series. See Appendix E.2 for details.

the starting points for the IRFs, which we describe below together with the procedure used to compute the IRFs.

After obtaining the firm's policy functions for job separation and job finding rates, we simulate the model for 10,000 monthly periods by generating a random sequence of the Markov process for productivity and then computing the relevant policy variables and state variables.[30] We then obtain the stationary distribution of labor productivity and compute the 10th and 90th percentiles. We split the simulated time series into two samples. The first sample includes all of the periods in which productivity is equal to 10th percentile: this is the low-productivity sample. The second sample includes all periods in which productivity is equal to the 90th percentile: this is the high-productivity sample. For each observation in a given sample, we collect the following variables: u_t, a_t, $e_t(x)$ and $e_{t-1}(x)$. We compute four IRFs as combinations of the following conditions: productivity is either "high" or "low", and the economy is hit by a positive one-quarterly standard deviation productivity shock.

Each IRF is obtained through a series of 1,000 simulations. As a starting point for each simulation, we draw a random observation from the relevant sample (either low or high productivity) with replacement: $\{u_0, a_0, e_0(x), e_{-1}(x)\}$. We then simulate a continuous Markov path of productivity from the preset starting value a_0 to a_T, where $T = 90$ months. We compute corresponding values of $\{u_t, e_t(x), x_t^r, sr_t, jfr_t, jjr_t\}_{t=0}^{T}$, where sr_t is the separations rate, jfr_t the job finding rate, and jjr_t is the job-to-job transition rate. We compute the "alternative" history in which the path of productivity is initially hit by a further positive shock in the initial

[30] We start with an initial 200 observations that are then discarded.

period: i.e. $\{\tilde{a_t}\}_{t=0}^{T}$ such that $\tilde{a_0} = a_0 + \sigma$ or $\tilde{a_0} = a_0 - \sigma$. The corresponding variables under the alternative aggregate productivity path are $\{\tilde{u_t}, \tilde{e_t(x)}, \tilde{x_t^r}, \tilde{sr_t}, \tilde{jfr_t}, \tilde{jjr_t}\}_{t=0}^{T}$. Because the value of $\tilde{a_t}$ does not fall on one of the nodes of the discretized grid, we compute the paths of all variables using a linear interpolation of the policy functions. After taking quarterly averages, the IRF for a given variable is computed as the difference between the values under the alternative and the baseline history: e.g. $\mathrm{d}u_q = \tilde{u_q} - u_q$ for the unemployment response, where q represents a one-quarter period. The 10th, 50th, and 90th percentile IRFs are calculated as the relevant percentiles of $\mathrm{d}u_t$ (or any other variable) across all simulations at each $t = 0, T$.[31]

To compute the decomposed job finding rate IRF in Figure E.1, we construct two alternative series: $\widehat{jfr}_t = p(\theta_t)(1 - F(\tilde{x_t^r}))$ and $\widetilde{jfr}_t = p(\tilde{\theta_t})(1 - F(x_t^r))$. The former assumes that only the reservation productivity responds to the shock. The latter assumes that only tightness responds to the shock.

E.4 Decomposition of the job finding rate IRF

To compute the decomposed job finding rate IRF, we construct two alternative series. For a given simulation, the tilde symbol indicates the alternative the response of a variable including the one-standard deviation shock. The the two alternative series are: $\widehat{jfr}_t = p(\theta_t)(1 - F(\tilde{x_t^r}))$ and $\widetilde{jfr}_t = p(\tilde{\theta_t})(1 - F(x_t^r))$. The former assumes that only the reservation productivity responds to the shock. The latter assumes that only tightness responds to the shock.

The solid line of Figure E.1 shows the counterfactual response of the job finding rate assuming that the reservation productivity x^r does not respond to the shock, so that the change in the job finding rate originates only from the change in labor market tightness θ_t. Since the magnitude in the response of the job finding rate is similar between low- and high-productivity states, the non-linearity in tightness plays a limited role in the generation of state dependence. Conversely, the dashed line in the figure shows the response of the job finding rate assuming θ_t is irresponsive to the shock, so that the mechanism for state dependence is the change in reservation productivity x^r. The response of the job finding rate in the state of high productivity (left panel) is only one third of that from the state of low productivity (right panel). The exercise thus indicates that changes in the reservation productivity are the predominant mechanism for the state dependence in the job finding rate.

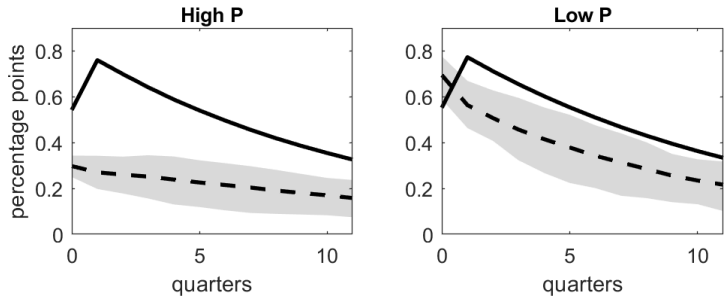
E.5 Assessing the state dependence through IRFs

To better quantify the extent to which the model explains the observed state dependence in the TVAR, we construct a metric that compares differences between the IRFs in the model and the TVAR. For each variable, we compare the difference between the IRFs at low- and high-productivity points in the model with the equivalent difference from the GIRFs in the TVAR. We define with y_h^p the h-period (median of the) IRFs of variable y that starts from the productivity percentile p in the model. Let $\Delta_{p'p}^h y$ represent the difference in the IRFs between the IRFs of variable y that start from the different percentiles in the distribution of productivity p' and p (i.e., $\Delta_{p'p} y(h) = y_h^p - y_h^{p'}$). Similarly, let $\Delta_{p'p} y^{\text{TVAR}}(h)$ represent the equivalent metric for the TVAR model (reported in the third row of Figure 4).

In our exercise, the percentiles of interest are the $p = 10^{\text{th}}$ and the $p' = 90^{\text{th}}$. Figure E.2 plots $\Delta_{p'p} y(h)$ (solid line) and $\Delta_{p'p} y^{\text{TVAR}}(h)$ (dashed line) for the unemployment rate (left panel), the job separation rate (middle panel), and job finding rate (right panel). The figure shows that

[31] It is important to note that the plotted IRFs therefore do not represent a specific and unique response, but simply the percentiles of the distribution of responses in period t.

Figure E.1: Generalized IRFs: decomposing the JFR response.



Note. The solid line reports the mean IRF of the job finding rate if only tightness θ_t responds to the shock. The dashed line reports the mean IRF of the job finding rate if only the reservation productivity for new matches x^r responds to the shock. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points.

the model replicates the difference in the IRFs for the unemployment rate in the early quarters, although it is unable to replicate the hump-shaped response. The peak difference across the responses at the 10th and 90th percentiles in the TVAR model realizes in the fourth quarter, while in the theoretical model it realizes in the second quarter. For the job separation rate the difference in the model closely tracks that of the TVAR at all horizons. For the job finding rate, the wedge in the model IRFs is more modest than in the data at medium and longer horizons.

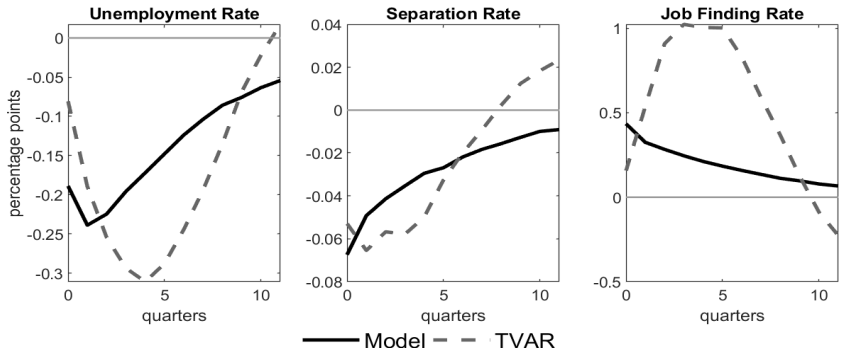
Next, we appraise the model's ability to capture the empirical state dependence at different time horizons quantitatively through a simple metric $\tilde{y}_{p'p}(H)$. For each variable and time horizon H, we compute the ratio of the average value of $\Delta_{p'p}y(h)$ and $\Delta_{p'p}y^{\text{TVAR}}(h)$ from $h = 0$ to H. In other words, we integrate over the solid and dashed lines of Figure E.2 and take the ratio of the two.

$$\tilde{y}_{p'p}(H) = \left(\frac{1}{\text{H}}\sum_{h=0}^{H}\Delta_{p'p}y(h)\right) \Big/ \left(\frac{1}{\text{H}}\sum_{h=0}^{H}\Delta_{p'p}y(h)^{\text{TVAR}}\right), \tag{26}$$

where H is the time horizon considered, p is the 10th percentile and p' the 90th. A value for $\tilde{y}_{p'p}(H)$ equal to 1 indicates that the theoretical model generates an average state dependence that is exactly the same as in the TVAR model at the given time horizon H.

The first row in Table E.3 reports values for $\tilde{y}_{p'p}(H)$ for the unemployment rate, the separation rate and the job finding rate in the baseline model (row (1)) for a time horizon equal to $H = 4, 8$, and 12, respectively. The metric ranges from 0.75 to 1.04 for unemployment rate, between 0.83 and 1.17 for the separation rate, and between 0.37 and 0.39 for the job finding rate. Overall, these findings show that the benchmark model replicates closely the state dependence in the unemployment rate and the job separation rate, while it is less powerful in replicating the state dependence in the job finding rate. The other rows of the table present the same metric for the alternative models with no OJS (row (2)) and no OJS and exogenous separation (row (3)). The second row in Table E.3 reports the state dependence metric in equation (26) for the model without OJS. This version performs worse than the baseline model, especially for the replication of the job finding rate. The state dependence metric for the job finding rate ranges between 0.18 and 0.24. The third row shows that the model without OJS and with exogenous separation performs worse than the baseline with respect to the state dependence of

Figure E.2: Difference in the IRFs at the 90th and 10th percentiles of productivity: baseline model vs. TVAR



Note. The solid line plots the difference in the mean simulated IRFs at the 10th and 90th percentiles of productivity in the model (as reported in Figure 7). The dashed line reports the difference in the median GIRFs at the 10th and $90th^{\text{th}}$ percentiles of productivity from the TVAR (as reported in Figure 4).

the unemployment rate. However, this version of the model performs better than the baseline model in replicating the state dependence of job creation rate, whose metric ranges between 0.53 ($h = 8$) and 0.79 ($h = 4$).

Table E.3: Percent of difference between TVAR GIRFs at 10th and 90th percentiles explained by model over different time horizons.

	Unemployment Rate			Separation Rate			Job Finding Rate		
	H=4	H=8	H=12	H=4	H=8	H=12	H=4	H=8	H=12
(1) Baseline	1.13	0.81	0.88	0.86	0.88	1.22	0.49	0.32	0.37
(2) No OJS	1.12	0.83	0.91	1.04	1.07	1.53	0.24	0.18	0.21
(3) No OJS, exogenous separation	0.37	0.28	0.31	-	-	-	0.79	0.53	0.61

Note. Each entry reports the value of the metric in equation (26) for the corresponding variable (column) in each version of the model (row) over different time horizons h.

F Details of alternative models

This section contains the main equations and the additional tables and figures pertaining to the alternative models. Table F.1 reports the values of the calibrated parameters. Table F.2 compares the simulated moments to the targets.

F.1 Alternative calibration

Figure F.2 shows the decomposition of the response of the job finding rate for model with OJS and endogenous separations under the alternative calibration.

F.2 Model without OJS and with endogenous separations

The surplus function is

$$S(a_t, x) = \max\{S^{n,c}(a_t, x), 0\}, \tag{27}$$

Table F.1: Parameter values to match empirical evidence for baseline and alternative models

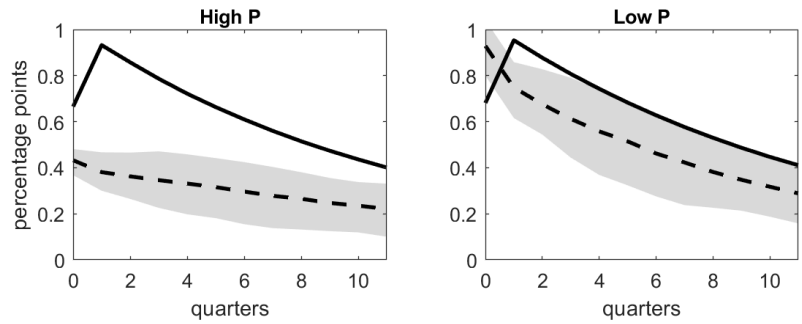
	Description	(1) Alternative Calibration	(2) No OJS	(3) Exogenous Separations
b	Flow value of unemployment	0.75	0.84	0.84
γ	Matching function efficiency	0.41	0.49	0.051
s	Exogenous separation probability	0.29	0.029	0.033
σ_x	Standard deviation of x draw	0.12	0.105	-
λ	Probability of new x draw	0.055	0.31	-
κ_s	Cost of OJS	0.12	-	-

Table F.2: Empirical targets and simulated moments from the baseline and alternative models.

		Targets	Alternative Calibration	No OJS	Exogenous Separations
SR	Mean	3.3	3.3	3.4	3.3
SR	Standard Deviation	1.7	-	1.7	-
SR	Standard Deviation relative to employment	0.22	0.22	-	-
SR	Autocorrelation	0.79	0.82	0.81	-
JFR	Mean	45	44	43	45
JFR	Standard Deviation	3.3	-	3.3	3.3
Job-to-Job	Mean	3.2	3.2	-	-

Note. The first column reports the targets for the moments of the separation rate (SR), the job finding rate (JFR), and the job-to-job rate. The targeted standard deviations of the separation rate and the job finding rate are derived by taking the share of the stationary variance explained by the TVAR (weighted across regimes) multiplied by the unconditional variance. Means and standard deviations are multiplied by 100 for presentational purposes.

Figure F.1: Generalized IRFs: decomposing the JFR response for the baseline model under the alternative calibration.



Note. The solid line reports the mean IRF of the job finding rate if only tightness θ_t responds to the shock. The dashed line reports the mean IRF of the job finding rate if only the reservation productivity for new matches x^r responds to the shock. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points.

where $S^{n,c}(a_t, x)$ is defined as (12) and the free-entry condition as in (7). The job fining rate is measured as $JFR_t = UE_{t+1}/u_t = p(\theta_t)(1 - F(x^r_{t+1}))$.

Figure F.2 decomposes the response of the job finding rate between the two alternative channels that generate state dependence.

Figure F.3 compares the value of $F(x^r)$ across aggregate productivity levels a_t for the baseline and the no-OJS model. $F(x^r)$ represents the share of unsuccessful matches from unemployment due to low individual productivity. In the baseline model (solid line) $F(x^r)$ is overall high, more strongly downward-sloped and convex with respect to aggregate productivity. In other words, x^r is very elastic with respect to shocks when aggregate productivity is low and more inelastic when it is high. Meanwhile, the same share in the no-OJS model is overall lower and flatter, implying that the strength of this channel is reduced and does not vary over the cycle. Hence, it does not contribute to state dependence in the job finding rate.

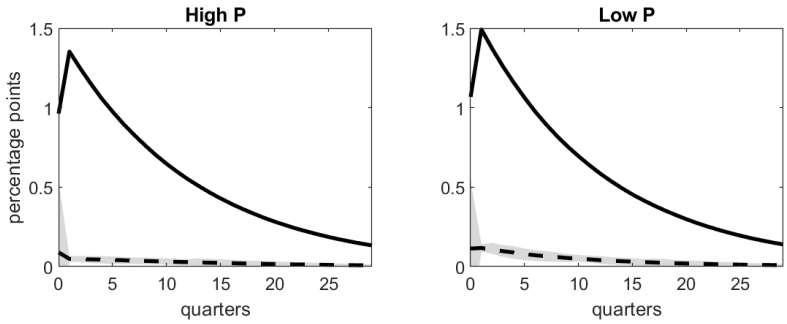
The intuition behind the different value of x^r in the two specifications of the model hinges on the role of OJS and the calibrated value of λ. When OJS is present, and the value of λ is 0.05, workers with a low idiosyncratic productivity x search for new jobs in other firms. The joint surplus of the match is low because the worker is likely to leave the firm for a better paid job. Given the cost of OJS, some matches are not worth forming since searching from unemployment is costless. Meanwhile, without OJS and with $\lambda = 0.31$, a low-x worker has high chances of obtaining a more productive job within the same firm, leading to a high joint continuation value of the current match. The firm is thus willing to accept matches with a low x that would not be convenient if the worker searched on the job.

F.3 Model without OJS and with exogenous separations

The DMP model with exogenous separations is defined by the surplus function, $S(a_t)$ which only has aggregate productivity as an argument:

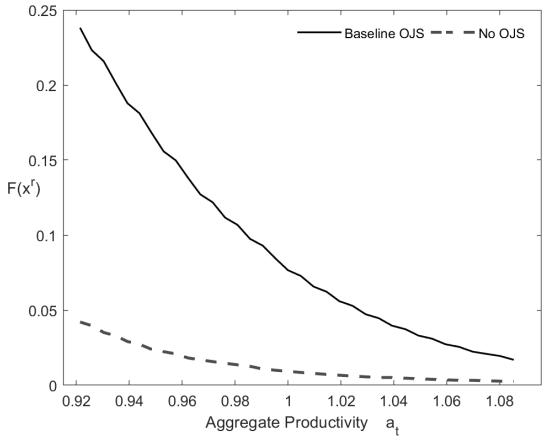
$$S(a_t) = a_t - b + \beta\mathbb{E}\Big\{(1 - s - p(\theta_t)\phi)S(a_{t+1})\Big\}.$$

Figure F.2: Generalized IRFs: decomposing the JFR response for the model without OJS.



Note. The solid line reports the mean IRF of the job finding rate if only tightness θ_t responds to the shock. The dashed line reports the mean IRF of the job finding rate if only the reservation productivity for new matches x^r responds to the shock. The shaded area represents the 5th and 95th percentiles of the IRF values. Responses of the variables in periods with high (low) aggregate productivity are in left (right) panels. Units on the y-axis are percentage points.

Figure F.3: Comparing reservation productivities: baseline and model without OJS.



Note. The solid black line reports the fraction of new matches that are not continued into employment in the baseline model for different values of aggregate productivity. The dashed grey line reports the corresponding fraction for the model without OJS.

The free entry condition becomes $k = q(\theta_t)\beta(1-\phi)\mathbb{E}\Big\{S(a_{t+1})\Big\}$ and the job finding rate is simply

$$JFR_t = UE_{t+1}/u_t = p(\theta_t).$$