Adaptive learning and labor market dynamics

Di Pace, Federico ; Mitra, Kaushik; Zhang, Shoujian

DOI:
10.1111/jmcb.12764

DOI: 10.1111/jmcb.12764

F. DI PACE
K. MITRA
S. ZHANG

Adaptive Learning and Labor Market Dynamics

The standard search and matching model with rational expectations is well known to be unable to generate amplification in unemployment and vacancies. We document a new feature that cannot be replicated: properties of wage forecasts published by institutions in the near term. A parsimonious model with adaptive learning can provide a solution to both of these problems. Firms choose vacancies by forecasting wages using simple autoregressive models; they have greater incentive to post vacancies at the time of a positive productivity shock because of overoptimism about the discounted value of expected profits.

JEL codes: E24, E32, J64
Keywords: adaptive learning, bounded-rationality, search and matching frictions

The Diamond-Mortensen-Pissarides (DMP) search and matching model has become the standard theory of equilibrium unemployment.

The views expressed here are those of the authors, and not necessarily those of the Bank of England or its policy committees. Shoujian Zhang was involved in the early stages of this work. We would like to thank participants of the conference on "Expectations in Dynamic Macroeconomic Models" (Bank of Finland), the 1st Birkbeck Centre of Applied Macroeconomics conference (BBK), the 20th Annual Computing in Economics and Finance Conference (Oslo), the 7th Conference of the Centre of Economic Growth and Business Cycles (Manchester), the 2014 Asian Meeting of Econometric Society (Taipei), 46th Annual Conference of the MMF (Durham), 32nd Annual Congress of the EEA (Lisbon), 14th Dynare Conference (ECB), 43rd Simposio of the Spanish Economic Association (Madrid), and seminar participants at the Universities of Exeter, Cardiff, the Technical University of Vienna, the Bank of England, and the Bundesbank. We would like to thank William Branch, George Evans, Ioannis Karavias, Pei Kuang, Lien Laureys, Bruce Preston, Joseph Pearlman, and, especially, the editor and two anonymous referees for useful comments and suggestions. Mallory Yeromonahos provided excellent research assistance.

[Correction added on 6 May, 2021, after first online publication: The copyright line was changed.]

F. Di Pace *is with Bank of England, United Kingdom* (*E-mail*: federico.dipace@bankofengland.co.uk). K. Mitra *is with University of Birmingham, United Kingdom* (*E-mail*: k.mitra@bham.ac.uk). S. Zhang *is with Erste Group Bank AG, Austria (E-mail*: shoujian.g.zhang@gmail.com).

Received January 26, 2018; and accepted in revised form March 9, 2020.
Journal of Money, Credit and Banking, Vol. 53, Nos. 2–3 (March–April 2021)

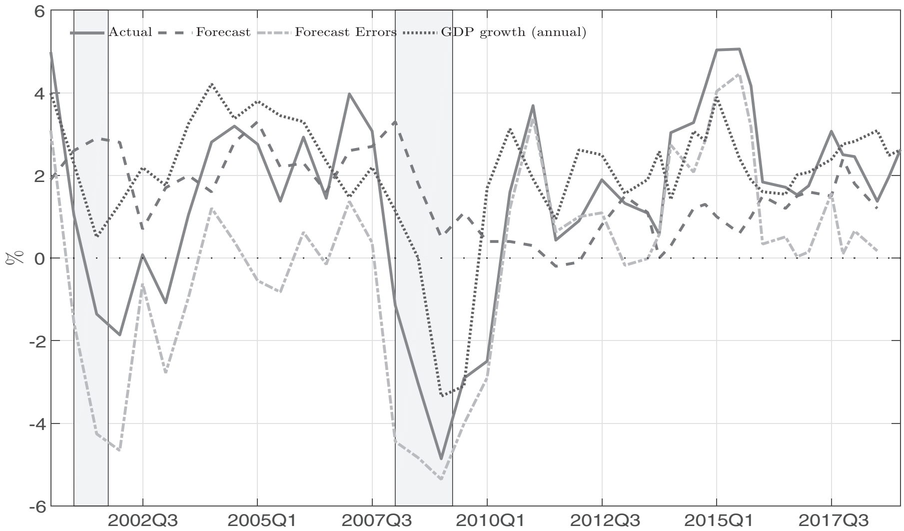

Fig 1. Wage Forecast Errors in the United States.

NOTES: Forecast errors of (real) annual growth of wage compensation are expressed in percentage points, that is, real wages minus forecasted wages in the previous year over the period 2000Q3–2018Q3. The actual realization of annual wage growth (compensation per employee) is denoted by the solid line, the forecasted annual wage growth by the dashed line and the forecast errors by the dash-dotted line. The dotted line shows annual GDP growth. Shaded areas indicate NBER recession dates. Note that forecast errors are irregularly spaced, this is evident from the number of black dots in the zero line.

Given its popularity, one might expect strong evidence of the model being consistent with key business cycle facts. However, Shimer (2005) shows that the standard search and matching model, driven by total factor productivity (TFP) innovations, has a hard time replicating the cyclical behavior of its central elements, namely, the amplification of labor market variables, such as unemployment, vacancies, and the measure of labor market tightness present in the U.S. data and other developed countries. Under the common assumption that wages are negotiated through Nash bargaining every period, wages tend to absorb most of the productivity innovations, generating little amplification in profits per hire.[1] This is referred to as the *unemployment volatility puzzle* in the literature.

We highlight in this paper, for the first time to our knowledge, another important feature of the data that the standard search and matching model under rational expectations (RE) is unable to replicate. One of the central variables in the search and matching model is the wage rate, which firms must forecast to make vacancy posting decisions. Figure 1 shows the forecasts of annual wage growth for the United States from European Commission (EC) (available biannually for the period 1999Q3 to 2018Q3). It shows that wage forecasts exhibit a great deal of persistence; forecast

1. Mortensen and Nagypal (2007) argue that the performance of the standard model featuring RE beliefs depends on the variability of profits per hire rather than on the assumption of wage cyclicality.

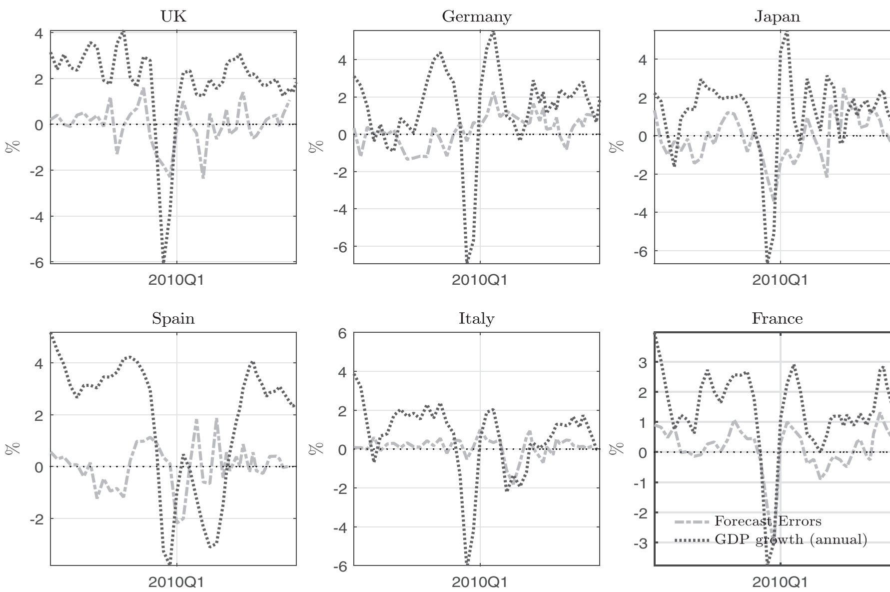

Fig 2. Annual Wage Forecast Errors and GDP Growth in Some Developed Economies.

NOTES: Forecast errors of (real) annual wage compensation per employee are expressed in percentage points, that is, real wages minus forecasted wages in the previous year over the period 2000Q3–2018Q3. The dash-dotted lines show forecast errors and the dotted line annual GDP growth.

TABLE 1

CORRELATION COEFFICIENT BETWEEN WAGE FORECAST ERRORS AND GDP GROWTH FOR A SUBSET OF DEVELOPED COUNTRIES

United States	United Kingdom	Germany	Japan	Spain	Italy	France
0.63***	0.56***	0.16	0.43***	0.00	0.39**	0.87***

NOTE: ***in denotes significance at 1% and **significance at 5%.

errors display systematic over/underprediction over the business cycle and are strongly procyclical. A similar pattern is observed for a number of developed economies (see Figure 2), whereby forecast errors are systematic and larger relative to what a standard model featuring RE would generate. For instance, Table 1 shows that for five out of the seven economies (including the United States) forecast errors are strongly procyclical and significant.[2]

Since standard models assume RE, they are less aligned with forecast data because agents do not make systematic errors in such models. The search and matching model with RE is unable to match the properties of these forecast errors. This paper provides

2. Wage forecast properties are discussed in Section 4. In an earlier version of the paper, Di Pace, Mitra, and Zhang (2016) show that similar systematic patterns arise in the forecast errors of unemployment available from the Survey of Professional Forecasters (SPF) over the longer period 1968–2015.

a solution to this puzzle by examining the role of expectation formation for hiring decisions. We replace the RE assumption with a set of simple autoregressive subjective beliefs for firms in line with the adaptive learning (AL) literature. This simple modification enables the model to match the general features present in the data: it presents a solution to the *unemployment volatility puzzle* and matches the statistical properties of wage forecast errors observed in the United States. This paper is the first one to highlight the role of expectation formation in the study of hiring decisions and its ability to replicate important features of labor market data.

We develop a simple search and matching model where wages are negotiated period by period with the assumption of RE being replaced by subjective beliefs as in the AL literature. Firms form forecasts of wage rates up to the infinite future to make hiring decisions using simple autoregressive models. Agents in our model face a dynamic problem due to long-lasting employment relations modeled through search frictions. We assume that agents have incomplete knowledge about the structure of the economy in that they do not know the technological constraints faced by other agents, and that wages are observed by agents with one period delay. Thus, in the baseline model, hiring decisions depend crucially upon the perception of future profits.

We find that not all forms of AL provide a solution to the unemployment volatility puzzle. In particular, when agents form their forecasts based on perceptions that take the form of the rational expectations equilibrium (REE) solution or when agents make only one-step ahead forecasts (rather than infinite horizon [IH] forecasts), then the amplification results are not too dissimilar from RE.[3] However, strikingly, AL models result in a much better fit to the data when agents use small forecasting models in their learning.[4] The AL models considered are in a sense a small departure from the RE solution. We consider two variants of small forecasting models motivated by the nature of the RE solution. The first model is a minimal departure from the RE solution in that agents assume that wages are AR(2) processes while the second model assumes that wages follow an AR(1) process.

We show that both of these models generate much more amplification in labor market variables compared to their RE counterpart. For instance, for the AR(2) model, unemployment and vacancies are 5.33 and 6.87 times more volatile than the corresponding measure of output, which is more in line with U.S. data; these numbers are in fact five times higher than those in the corresponding RE model. Our simple model matches well the statistical properties of forecast errors on wage data taken from the EC. In particular, the absolute volatility and procyclical behavior of wage forecast errors generated by these models are consistent with the data. This finding is in sharp contrast with the RE model and supports the type of perceived beliefs assumed in the learning model.

3. The former is often called learning based on correctly specified laws of motion and was popular in the early AL literature of the 1990s/2000s since the primary focus there was on theoretical convergence to the REE, often examined in simple *ad hoc* models (see, e.g. Evans and Honkapohja 2001). Subsequently, other forms of (underparameterized) beliefs have been explored see, for example, chapters 13 and 14 of Evans and Honkapohja (2001) and other references cited in footnotes 4 and 12.

4. In this sense, the results are consistent with Slobodyan and Wouters (2012), who too find that AL models give a good fit to U.S. inflation data for 1966–2008 and are closely related to the survey evidence on inflation expectations when agents form expectations using small forecasting models.

In the search and matching literature, the job creation condition represents the optimal decision rule for vacancy posting. Firms post vacancies until the expected marginal cost of posting a vacancy equals the benefits of hiring an additional worker, which can be expressed in equilibrium as the (infinite) sum of expected future profits generated at the margin.[5] Agents with incomplete knowledge about the structure of the economy tend to become optimistic after a positive TFP innovation and this in turn leads to more vacancy creation; the optimism is greater since firms forecast infinite periods ahead, which results in more vacancy creation and greater amplification. This means that the impact effect of productivity shocks on the present discounted value of profits is large because agents make systematic forecast errors about the path of future wages. Since firms forecast infinite steps ahead to choose how many vacancies to post, their *discounting* of future marginal products and wages turns out to be central.

Incidentally, there are a large number of studies that have attempted to provide solutions to the *unemployment volatility puzzle* under the assumption of RE. Two prominent solutions make relatively simple modifications to the standard search and matching model to generate greater amplification. The first approach proposed by Hall (2005) and Shimer (2005) introduces real wage rigidities. This means that, as wages cannot fully reflect productivity shifts, there is further incentive for vacancy creation.[6] The second popular approach by Hagedorn and Manovskii (2008) carries out a simple calibration exercise that sets the value of nonmarket activity close to the value of search.[7]

The main assumption in the paper that economic agents engage in "learning" behavior has been incorporated into macro-economic theory and used in a wide range of applications, see Evans and Honkapohja (2001 ,2003, 2006), Bullard and Mitra (2002), and Preston (2005, 2006, 2008). The standard AL approach treats economic agents like econometricians who estimate forecast rules, updating the parameter estimates over time as new data become available. Agents update their forecasts of future variables and resolve their dynamic optimization problem in order to make their decisions. This learning approach can be viewed as a version of the anticipated utility approach formulated by Kreps (1998) and used by Eusepi and Preston (2011) and Kuang and Mitra (2016) within the context of the RBC model; see the first two paragraphs of section 2 in Evans, Honkapohja, and Mitra (2012) for further discussion.

5. The solution to the model with RE beliefs is the same regardless of whether the decision rule for vacancies is specified recursively or as an infinite sum of future profits per hire.

6. Menzio (2005), Gertler, Sala, and Trigari (2008), Christoffel and Kuester (2008), Gertler and Trigari (2009), Blanchard and Gali (2010), and Hertweck (2013), among others, extend this idea to a general equilibrium setting.

7. These characterizations of the labor market have not been free from criticism. See, for example, Pissarides (2009) and Haefke, Sonntag, and van Rens (2013). There are of course other solutions to the unemployment volatility puzzle; see Menzio and Shi (2011), Quadrini and Trigari (2008), Gomes (2011), Reiter (2007), Guerrieri (2008), Robin (2011), Petrosky-Nadeau (2013), Petrosky-Nadeau and Wasmer (2013), and Alves (2018).

The point that there is an important divergence between the implied expectations in macro-economic models with RE and the expectations drawn from survey data has been made in the context of other models; see, for example, Adam, Marcet, and Beutel (2017), Kuang and Mitra (2016), Slobodyan and Wouters (2012), Milani (2011), and Ormeño and Molnar (2015). These papers show that systematic errors/gaps in variables such as GDP and interest rates are evident in survey forecasts (like the SPF). Other recent studies such as Malmendier and Nagel (2016) (for inflation expectations from the Reuters/Michigan Survey of Consumers), Gennaioli, Ma, and Shleifer (2016) (for expectations of earnings growth from Duke University's quarterly survey of Chief Financial Officers (CFOs)) and Greenwood and Shleifer (2014) (for expectations of returns from six different surveys of investors including a Gallup survey, investor newsletters, and the survey of CFOs of large corporations) also find evidence that agents use simple extrapolative rules to compute expectations.

The remainder of the paper is organized as follows. Section 1 describes the model. Section 2 states the solution of the model under RE and AL. Section 3 presents the main results of the baseline models. Section 4 evaluates the statistical properties of forecast errors. Section 5 examines whether agents can detect *ex post* misspecification in their beliefs. Section 6 analyzes the robustness of the results to three model extensions. Section 7 highlights the importance of the main assumptions. Section 8 concludes.

1. MODEL

We propose a model featuring labor market search and matching frictions as in Mortensen and Pissarides (1994) and a form of AL following Preston (2005) and Mitra, Evans, and Honkapohja (2013). Our model economy is inhabited by a continuum of firms. There is also a continuum of workers that search for jobs if unemployed and work for firms if employed. Firms post job vacancies and employ workers with a lag so as to produce output using labor as the only input of production. Agents form their expectations by updating their beliefs as new information becomes available. They make IH forecasts about the future path of wages by running simple autoregressive models in order to make vacancy posting decisions.

1.1 Labor Market

The labor market is frictional in that, from the perspective of the firm, it is costly to post vacancies and, from the standpoint of workers, searching for jobs is a time-consuming process. Every period firms create new vacancies, sought by unemployed workers who are looking for new job opportunities. Following Shimer (2010), we assume that workers that are matched at time t become productive at the beginning of next period, $t + 1$. Worker–firm matches break up at the exogenous rate, $\rho \in (0, 1)$. The aggregate number of matches, m_t, depends positively on both the unemployment rate, u_t, and aggregate vacancies, v_t. We assume that the matching process is guided by the following function:

$$m(v_t, u_t) = \bar{m} v_t^{\sigma} u_t^{1-\sigma}, \tag{1}$$

where $\bar{m}$ denotes the level of matching efficiency, σ the elasticity of the matching function with respect to aggregate vacancies, and the unemployment rate is defined as

$$u_t = 1 - n_t. \tag{2}$$

We define the measure of labor market tightness as

$$\theta_t = v_t / u_t. \tag{3}$$

Due to the assumption of constant returns to scale, the job finding rate is given by $m(v_t, u_t)/u_t = m(v_t/u_t, 1) = p(\theta_t)$ and the job filling rate by

$$m(v_t, u_t)/v_t = m(1, u_t/v_t) = q(\theta_t). \tag{4}$$

The job finding rate, $p(\theta_t)$, is increasing in θ_t and the job filling rate, $q(\theta_t)$, decreasing in θ_t.

1.2 Firms

Our model economy features a continuum of large firms of measure $f \in [0, 1]$. We assume that firms are large to aid comparison with the model extensions in Section 6.[8] Since posting vacancies is costly, period profits (π_{ft}) at time t may be written as

$$\pi_{ft} = z_t n_{ft} - w_{ft} n_{ft} - \kappa v_{ft}, \tag{5}$$

where v_{ft} and n_{ft} are the number of job openings and the level of employment at the firm and κ is the cost of posting a vacancy. Productivity shocks (z_t) follow the exogenous process given by

$$\ln z_{t+1} = \varrho \ln z_t + \epsilon_{t+1} \quad \text{with} \quad \epsilon_t \sim N(0, \varsigma), \tag{6}$$

where $\varrho \in (0, 1)$ denotes the persistence of the technology process and ϵ_t is an i.i.d. innovation with mean zero and standard deviation ς. The problem of each firm is to

8. The assumption of large firms together with constant returns to scale in employment yields an equivalent job creation condition relative to a setting in which there are one-worker firms (small firms). As shown by Krause and Lubik (2007b), the aggregate effects of intrafirm bargaining within a large firm environment are negligible in a standard search and matching framework with concave production functions. Thus, we abstract from intra-firm bargaining in the analysis.

choose v_{ft} so as to maximize the present discounted value of expected profits, which may be written as

$$\max_{v_{ft+j}} \quad \pi_{ft} + \beta \mathbb{E}^*_{ft} \left\{ \sum_{j=1}^{\infty} \pi_{ft+j} \right\} \quad \text{for} \quad j \geq 0, \tag{7}$$

subject to the law of motion of employment

$$n_{ft+1} = (1-\rho) n_{ft} + v_{ft} q_t. \tag{8}$$

The parameter β denotes the discount factor and $\mathbb{E}^*_{ft}$ the subjective expectation operator. The first-order condition with respect to v_{ft} is

$$\kappa = q(\theta_t) \beta \mathbb{E}^*_{ft} \left\{ \mathcal{V}_{ft+1} \right\}, \tag{9}$$

where $\mathcal{V}_{ft} = \mathcal{J}'(n_{ft})$ denotes the value of having an additional worker employed at the firm. Equation (9) states that the marginal cost and benefit of posting a vacancy must be equal. The envelope condition with respect to n_{ft} is

$$\mathcal{V}_{ft} = z_t - w_{ft} + (1-\rho) \beta \mathbb{E}^*_{ft} \left\{ \mathcal{V}_{ft+1} \right\}. \tag{10}$$

This condition simply states that the value of having an additional worker employed at the firm must be equal to flow profits—the marginal productivity of employment net of wage costs—plus the continuation value of employment at the firm.

Combining equations (9) and (10), we obtain the following job creation condition:

$$\frac{\kappa}{q(\theta_t)} = \beta \mathbb{E}^*_{ft} \left\{ z_{t+1} - w_{ft+1} + (1-\rho) \frac{\kappa}{q(\theta_{t+1})} \right\}. \tag{11}$$

This condition is central to our analysis since it determines the optimal number of vacancies that firm f would like to post. The expression simply states that the expected cost of a filled vacancy must be equal to its marginal benefit, which consists of expected profits and savings generated from the additional match. This way of formulating the firm's problem means that the choice of current vacancies is based on the forecast of future labor market conditions. Alternatively, the job creation condition can be rewritten as,

$$\frac{\kappa}{q(\theta_t)} = \beta \sum_{j=1}^{\infty} (1-\rho)^{j-1} \beta^{j-1} \mathbb{E}^*_{ft} \left[z_{t+j} - w_{ft+j} \right]. \tag{12}$$

Note that this expression is very intuitive; it says that the expected cost of a filled vacancy must be equal to the sum of the future stream of profits that the job is expected to generate. In order to post the optimal number of vacancies, firm f must make forecasts up to the infinite future (assuming the perceived transversality condition holds).

1.3 Wage Negotiation

Wages are negotiated according to a Nash bargaining protocol. The wage w_{ft} maximizes the joint surplus of a match between workers and firms,

$$\arg\max_{w_{ft}} \left[\mathcal{W}_{ft} - \mathcal{U}_t\right]^{\xi} \left(\mathcal{V}_{ft}\right)^{1-\xi},$$

where $\mathcal{W}_{ft}$ and $\mathcal{U}_t$ denote the employment (at firm f) and unemployment values for a worker at date t and $\xi \in (0, 1)$ denotes the workers' bargaining power.[9] The first-order condition of this problem then yields the standard sharing rule that characterizes the optimal split of the aggregate surplus,

$$(1-\xi)\left(\mathcal{W}_{ft} - \mathcal{U}_t\right) = \xi \mathcal{V}_{ft}. \tag{13}$$

To derive an expression for the bargained wage (w_{ft}), we assume that expression (13) holds in expectations. Thus,

$$w_{ft} = \xi(z_t + \kappa\theta_t) + (1-\xi)b. \tag{14}$$

The bargained wage is a weighted average of the marginal product of employment, the cost of replacing the worker and the opportunity cost of working (b).

1.4 Aggregation and Linearization

We make the assumption that in the temporary equilibrium workers and firms share the same set of beliefs about the future. This assumption is reasonable because workers and firms coordinate on expectations during the (*ex post*) wage negotiation process. The assumption of *ex post* homogeneity in expectations across workers and firms is standard in the learning literature ($\mathbb{E}_t^* = \mathbb{E}_{ft}^* = \mathbb{E}_{wt}^*$) and it implies a symmetric equilibrium (i.e., $n_{ft} = n_t$ and $v_{ft} = v_t$).

We further define aggregate output (y_t) as output net of vacancy costs,

$$y_t = z_t n_t - \kappa v_t. \tag{15}$$

In line with Mitra, Evans, and Honkapohja (2013), we assume that firms use a vacancy posting rule based on linearization of equation (12) around the steady-state values of $\bar{w}$ and $\bar{n}$. The firms' behavioral rule is

$$-\frac{\kappa}{\bar{q}^2}(\sigma-1)m\bar{\theta}^{\sigma-2}\tilde{\theta}_t = \mathcal{S}_{z,t} - \mathcal{S}_{w,t} = \mathcal{S}_t, \text{ where} \tag{16}$$

$$\mathcal{S}_{z,t} = \mathbb{E}_t^* \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j \tilde{z}_{t+j} \quad \text{and} \quad \mathcal{S}_{w,t} = \mathbb{E}_t^* \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j \tilde{w}_{t+j}.$$

9. Online Appendix A.1 contains the formulations of $\mathcal{W}_{ft}$ and $\mathcal{U}_t$ and Appendix A.2 the derivation of w_{ft}.

Note that a tilde over a variable x denotes the deviation of the variable from its steady-state value (i.e., $\tilde{x} = x - \bar{x}$, where $\bar{x}$ denotes the steady-state value of x).

We turn to the other key equations in the model. We linearize equations (6), (8), and (14) around the steady state and integrate over f to get

$$\tilde{z}_{t+1} = \varrho \tilde{z}_t + \epsilon_{t+1}, \tag{17}$$

$$\tilde{n}_{t+1} = (1-\rho)\tilde{n}_t + \bar{v}\tilde{q}_t + \bar{q}\tilde{v}_t, \tag{18}$$

$$\tilde{w}_t = \xi\left(\tilde{z}_t + \kappa\tilde{\theta}_t\right). \tag{19}$$

A temporary equilibrium is a set of values for the variables $\tilde{n}_t$, $\tilde{\theta}_t$, $\tilde{w}_t$ that, given the exogenous stochastic process $\{\tilde{z}_j\}_{j=t}^{\infty}$ and the initial condition $\tilde{n}_0$, satisfies the system of equations consisting of equations (16) and (17)–(19). This equilibrium is determined as follows. In every period t, given their forecasts, firms enter the market with their vacancy posting rule. To complete the description of temporary equilibrium one specifies how forecasts are formed. As argued before, it is plausible to assume that forecasts of firms are predetermined when they are brought to the market. The temporary equilibrium for the current period provides a new data point for agents. Given these new data, agents update their forecasts at the start of the following period using versions of RLS algorithm.

Finally, we linearize the auxiliary equations (2), (3), (4), and (15)

$$0 = \tilde{n}_t + \tilde{u}_t, \tag{20}$$

$$\tilde{v}_t = \bar{\theta}\tilde{u}_t + \bar{u}\tilde{\theta}_t, \tag{21}$$

$$\tilde{q}_t = (\sigma - 1)m\bar{\theta}^{\sigma-2}\tilde{\theta}_t, \tag{22}$$

$$\tilde{y}_t = \bar{n}\tilde{z}_t + \bar{z}\tilde{n}_t - \kappa\tilde{v}_t. \tag{23}$$

2. BELIEFS AND INFORMATION ASSUMPTIONS

The RE solution of the model is of the following form:

$$\begin{aligned} n_{t+1} &= \bar{b}_n + \bar{a}_{nn}n_t + \bar{a}_{nz}\tilde{z}_t, \\ w_t &= \bar{b}_w + \bar{a}_{wn}n_t + \bar{a}_{wz}\tilde{z}_t, \end{aligned} \tag{24}$$

given that employment and productivity are the only state variables in this model. The RE solution is expressed in levels rather than in deviations from their steady-state values. Under RE agents know that the productivity innovation follows an autoregressive process, its persistence and dispersion of the TFP innovation. Rational agents have complete information about the structure of the economy and know the true parameter values of policy functions (denoted with a bar over the parameter). Thus, rational agents make no systematic mistakes. As is well known, the RE model is not able to match the amplification present in labor market data since productivity shocks are unable to increase profits per hire by much. More detailed explanations are provided in Section 3.4.

We note that, following Campbell (1994), the RE solution (24) may be written in an equivalent way. This RE solution involves n_t as an AR(2) process and w_t as ARMA(2,1) processes. In particular, this solution can be shown to be of the following form:

$$\begin{aligned} n_{t+1} &= \vartheta_1 \bar{b}_n + \vartheta_2 n_t - \vartheta_3 n_{t-1} + \bar{a}_{nz}\epsilon_t, \\ w_t &= \vartheta_4 \bar{b}_w + \vartheta_2 w_{t-1} - \vartheta_3 w_{t-2} + \bar{a}_{wz}\epsilon_t + (\bar{a}_{wn}\bar{a}_{nz} - \bar{a}_{wz}\bar{a}_{nn})\epsilon_{t-1}, \end{aligned} \tag{25}$$

where $\vartheta_1 = (1 - \bar{\varrho})$, $\vartheta_2 = (\bar{\varrho} + \bar{a}_{nn})$, $\vartheta_3 = \bar{\varrho}\bar{a}_{nn}$, and $\vartheta_4 = (1 - \varrho)[(1 - \bar{a}_{nn})\bar{b}_w + \bar{a}_{wn}\bar{b}_n]$.

Using beliefs of the form (25) (or (24)) requires a lot of knowledge from agents: they need to know that solutions for the endogenous variables are exactly of the form above. In practice, ARMA-type processes are significantly more difficult to estimate. Pure AR models have been advocated by time series analysts as parsimonious models for forecasting (over ARMA models) on the grounds of being simpler to estimate and easier to specify because no identifiability problems arise in a procedure of "testing down" (when going from a general to a specific approach) to see if the model for forecasting could be simplified; see, for example, Granger and Newbold (1986) and Harvey (2008), pp. 78–80, for a discussion. Our agents adopt parsimonious models like AR(1) or AR(2) for their forecasting exercises in a similar vein.

In addition, agents endowed with these AR beliefs do not stop at this stage. Loosely speaking, they use a model selection strategy popularized by Box and Jenkins. In Section 5, it is shown that the baseline model featuring autoregressive beliefs in wages is consistent with the behavior of actual wages in the US data. We find that, using alternative approaches, agents are unable to detect misspecification in their forecasting methods *ex post* for very long periods of time.

As mentioned in the Introduction, recent studies have also found evidence for the use of extrapolative rules, for example, Malmendier and Nagel (2016), Gennaioli, Ma, and Shleifer (2016), Greenwood and Shleifer (2014), and Slobodyan and Wouters (2012). Moreover, recent experimental evidence by Hommes et al. (2005) and Heemeijer et al. (2009) suggests agents estimate simple univariate autoregressive models to make forecasts about future variables. Using AR-type perceived law of motion (PLM) is, therefore, very appealing.[10]

Thus, we deviate from RE in assuming that agents forecast the values of the variables of interest based on simple AR belief specifications and examine whether these belief specifications have the ability to get the search and matching model closer

10. Using AR beliefs obviates the need for agents to use productivity in their regression equations (34). We remark that Slobodyan and Wouters (2012) use AR(2) PLMs as their preferred specification in estimating medium-scale DSGE model based on small forecasting models. Their model features a wider set of nominal and real frictions and the dynamics of their model is driven by multiple innovations.

to the data. Agents have incomplete knowledge about the structure of the economy and they observe only their own objectives and constraints but do not observe other agents' production functions and beliefs. Thus, they do not know that their decisions are identical to those of other agents.

2.1 Learning with Autoregressive Beliefs

We propose alternative belief systems of autoregressive form.[11] The simplest forecasting model that is closest to the RE model is one where all endogenous variables are forecasted using univariate autoregressive processes of order 2, that is, AR(2) processes. This serves as the benchmark for agents' beliefs. Agents have PLMs of the form

$$w_t = a_0 + a_1 w_{t-1} + a_2 w_{t-2} + \mu_t, \tag{26}$$

where μ_t is a white noise process. Like RE, the belief system is expressed in levels. This belief specification represents only a modest departure from RE: the key difference being that the moving average (MA) term is dropped from w_t.[12] This seems like a reasonable assumption since wages are determined simultaneously by equilibrium considerations and depend on aggregate variables such as θ_t. We also examine AR(1)-type beliefs system, which is nested in equation (26).

For economy of space, we only explain parameter updating with AR(2) beliefs. Let $\Phi_t = [a_0\, a_1\, a_2]'$ and $\Psi_t = [1\, w_{t-1}\, w_{t-2}]'$. Agents use a constant gain recursive least squares (RLS) algorithm (widely used in the AL literature) to update their beliefs

$$\begin{aligned} \Phi_t &= \Phi_{t-1} + \gamma R_t^{-1} \Psi_{t-1} \left(w_{t-1} - \Phi_{t-1}' \Psi_{t-1} \right)', \\ R_t &= R_{t-1} + \gamma \left(\Psi_{t-1} \Psi_{t-1}' - R_{t-1} \right), \end{aligned} \tag{27}$$

where $\gamma \in (0, 1)$ denotes the constant gain learning parameter and R_t is the precision (3×3) matrix associated with each equation.

Agents update their beliefs over time by revising the value of parameters using the constant gain RLS algorithm. At the beginning of each period, agents inherit the parameters of their belief system from the previous period, make forecasts, and compute the present discounted sums that allows them to form vacancy posting decisions at every point in time. At the end of each period, agents are informed about wages

11. These alternative belief specifications are similar in spirit to the *simple wage rule models* proposed by Christiano, Eichenbaum, and Trabandt (2016) within the RE literature, where the wage rules are autoregressive. For further details on the learning algorithms, see Online Appendices A.3 and A.4.

12. Under the proposed belief system the economy converges to a restricted perceptions equilibrium (RPE) as in Sargent (1999), Cho, Williams, and Sargent (2002), and Branch and Evans (2006a) that is different from the RE equilibrium; see chapter 13.1 of Evans and Honkapohja (2001), Branch (2006), for a discussion and Huang, Liu, and Zha (2009) for an application to the growth model. As our interest is in matching unemployment volatility and wage forecasts, we do not study the nature of this RPE. We do, however, show the evolution of the parameters in agents' PLM and what values they converge to; see Figure 11 in the Online Appendix.

and they update their beliefs in the following period. In the learning literature, it is standard to assume that for the parameter estimation agents use data available up to period $t - 1$.

2.2 Informational Assumptions

The standard assumption in the learning literature is to assume that agents' forecasts at t from period $t + 1$ onward are based on past endogenous variables (here w_{t-1}). This approach conveniently avoids the simultaneous determination of forecasts and endogenous variables. An interpretation, owing to Evans and Honkapohja (2001) and Evans and Honkapohja (2006), is that these forecasts are computed before going to the market place. We therefore assume that firms make wage forecasts at the beginning of period t and the productivity shock is also revealed at the beginning of the period, that is, firms do not observe observe wages (w_t) at the time of making their wage forecasts. This informational assumption is particularly appealing here since the standard search literature is calibrated to monthly frequency and the data are generally available with a lag.

In addition, wage determination in the search and matching literature, which is different from the standard neoclassical model, entails a complex decision problem involving expectations. Under RE, workers and firms have the same expectations and information set. However, in our context, agents may not know this to be true *ex ante*. *Ex post*, the bargained wage $\tilde{w}_t$ will depend on aggregate variables such as the measure of labor market tightness $\tilde{\theta}_t$; see equation (14) or (19). The assumption of time-to-hire is such that workers are matched to employees in period t but can only start working in period $t + 1$. Therefore, the assumption of predetermined forecasts in this model is particularly plausible.

3. NUMERICAL RESULTS FOR MODEL AND DATA

3.1 Calibration

We set the structural values of the parameters in the model following a standard calibration exercise. First, we choose some parameter values using *a priori* information. Second, the choice of the remaining parameters ensures that the stationary equilibrium of the model matches a number of stylized facts as observed in the post-WWII U.S. economy. As is standard in the search and matching literature, a period in our model corresponds to a month in the data.

The parameters chosen using *a priori* information are the subjective discount factor (β), the exogenous separation rate (ρ), the worker's bargaining power (ξ), and the elasticity of the matching function with respect to vacancies (σ). The value of β is set to 0.996, which implies an annual real interest rate of about 4%. The value of ρ is calibrated to 0.033 in order to match the evidence that jobs last on average two and a half years as estimated in Davis, Haltiwanger, and Schuh (1996). We set the value of σ at 0.5 in line with the literature. This value lies within the plausible interval

TABLE 2

Calibrated Parameters: Monthly

Description	Parameter	Value
Discount factor	β	0.996
Replacement ratio	rr	0.6
Efficiency of the matching technology	$\bar{m}$	0.379
Elasticity of the matching function	σ	0.5
Bargaining power	ξ	0.5
Separation rate	ρ	0.033
Productivity level	$\bar{z}$	1
Persistence of productivity shocks	ϱ	0.98
St. dev. of productivity shocks	ς	0.005
Gain parameter	γ	0.002

of [0.5, 0.7] as surveyed by Petrongolo and Pissarides (2001). In order to facilitate comparability with the existing literature, ξ is chosen to be 0.5. Following Shimer (2010), we choose the persistence of the technology shock (ϱ) to be 0.98 and the standard deviation of the innovation to be 0.005.

The remaining two labor market parameters, namely, κ are $\bar{m}$ are set to match (i) a vacancy filling rate of 27.8% as estimated by Shimer (2005), which is consistent with a quarterly rate of 70% as in Trigari (2006) and den Haan, Ramey, and Watson (2000); (ii) an unemployment rate of 6%, which corresponds with the standard ILO definition of unemployment for the post-WWII U.S. average.

The replacement ratio is set to 60%, which is slightly below the value suggested by Mortensen and Nagypal (2007). According to a study by Hagedorn and Manovskii (2008), a total replacement ratio of around 95% can generate labor market fluctuations that are in line with the empirical evidence. Their study argues that, if the outside option of the worker is high (this happens when both ξ is low and the replacement ratio high), then (steady-state) firm's profits are small and can generate greater amplification in labor market variables. The resulting replacement ratio ensures that our results are not driven by the Hagedorn and Manovskii effect. Table 2 provides a summary of the parameters used in the baseline calibration of our hypothetical model economy.

We choose the gain parameter in the learning algorithm to be $\gamma = 0.002$ (equivalent to a value of 0.006 in the corresponding quarterly model), which implies that agents use past data to update their beliefs for around 42 years ($1/0.002 =$ 500 months or 167 quarters). There is lack of consensus in the learning literature concerning the constant gain parameter, which ranges from 0.002 to 0.035 at quarterly frequencies. See, for example, Eusepi and Preston (2011), Branch and Evans (2006b), Milani (2007), and Orphanides and Williams (2007). The value chosen for this parameter is relatively small because we exclude policy considerations (e.g., as in Mitra, Evans, and Honkapohja (2017) or Mitra, Evans, and Honkapohja (2013) where a higher gain parameter is used) but lies within the range of parameters suggested in the literature. The smaller the gain parameter the longer it takes to learn the long-run equilibrium.

TABLE 3

SUMMARY STATISTICS, QUARTERLY U.S. DATA

		$\hat{y}_t$	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$
$\sigma_{\widehat{x}_{1t}}/\sigma_{\widehat{y}_t}$		1	0.57	9.76	8.80	18.21
$\rho(\widehat{x}_{1t}, \widehat{x}_{1t-1})$		0.81	0.90	0.90	0.89	0.90
	$\hat{y}_t$	1	0.79	0.82	−0.78	0.82
	$\hat{n}_t$	–	1	0.92	−0.97	0.96
$\rho(\widehat{x}_{1t}, \widehat{x}_{2t})$	$\hat{v}_t$	–	–	1	−0.93	0.98
	$\hat{u}_t$	–	–	–	1	−0.98
	$\hat{\theta}_t$	–	–	–	–	1

NOTES: Relative standard deviations, autocorrelation, and correlation coefficients in this table correspond to quarterly data series detrended using a Hodrick–Prescott filter with smoothing parameter 1600. Each data series x_{1t} corresponds to a variable in the model. The term $\rho(x_{1t}, x_{2t})$ stands for the correlation coefficient between variables x_{1t} and x_{2t}.

3.2 U.S. Data

In this section, we compare the main statistical properties of the simulated labor market series generated from the model with the corresponding series in the U.S. data, in particular focusing on second moments.

The seasonally adjusted series of (un)employment is taken from the Bureau of Labor Statistics (BLS). As a proxy for vacancies, we merge the seasonally adjusted help-wanted advertising index released by the Conference Board with the vacancy series calculated by Barnichon (2010). Aggregate output is measured as seasonally adjusted real GDP, which is drawn from the National Income and Product Account (NIPA; tables 1.1.6 and 1.1.5). All data series are quarterly and cover the period ranging from 1951Q1 to 2016Q4 (due to the data availability on vacancies). Table 3 summarizes the main cyclical properties of the logged detrended series.

One of the most salient features in the data is the high volatility of unemployment, vacancies, and labor market tightness as reported in Table 3. In particular, both vacancies and unemployment are about 9.76 and 8.80 times more volatile than the aggregate output, respectively. Moreover, the measure of labor market tightness is around 18.21 times more volatile than output. Another well-known stylized fact is the negative relationship between vacancies and unemployment, also known as the Beveridge curve.

3.3 Simulation Results

We simulate the search and matching model under the different belief specifications and compare results. We use standard methods to solve and simulate the RE model. We initiate the simulations of our learning models from values consistent with the deteministic steady state and then generate a series for 10,900 periods using the learning algorithm previously stated. The first 10,000 periods ensure convergence to the long-run equilibrium and are, as a result, discarded. We keep the remaining 900 observations, which correspond to 75 years of data, so as to guarantee that the simulated series are free from any transitional dynamic considerations. We then repeat

TABLE 4

Summary Statistics: RE, AR(1), and AR(2) Learning Models

Model	Statistics	$\widehat{y}$	$\widehat{n}$	$\widehat{u}$	$\widehat{v}$	$\widehat{\theta}$	$\widehat{q}$
	$\sigma_{\hat{x}_t}/\sigma_{\hat{y}_t}$	1.00	0.07	1.08	1.29	2.32	1.16
RE	$\rho(\hat{x}_t, \hat{v}_t)$	0.98	0.96	−0.92	1.00	0.98	−0.98
	$\rho(\hat{x}_t, \hat{x}_{t-1})$	0.94	0.95	0.95	0.88	0.94	0.94
	$\sigma_{\hat{x}_t}/\sigma_{\hat{y}_t}$	1.00	0.48	7.55	9.52	16.43	8.21
$IH - AR(1)$	$\rho(\hat{x}_t, \hat{v}_t)$	0.76	0.93	−0.85	1.00	0.97	−0.97
	$\rho(\hat{x}_t, \hat{x}_{t-1})$	0.93	0.92	0.93	0.71	0.88	0.88
	$\sigma_{\hat{x}_t}/\sigma_{\hat{y}_t}$	1.00	0.34	5.33	6.87	11.61	5.80
$IH - AR(2)$	$\rho(\hat{x}_t, \hat{v}_t)$	0.74	0.91	−0.81	1.00	0.96	−0.96
	$\rho(\hat{x}_t, \hat{x}_{t-1})$	0.93	0.92	0.93	0.70	0.87	0.87

Notes: Relative standard deviations, autocorrelation, and correlation coefficients in this table correspond to the quarterly simulated series expressed in percentage deviations from steady-state value. The term $\rho(x_{1t}, x_{2t})$ stands for the correlation coefficient between variables x_{1t} and x_{2t}.

this procedure 100 times and report the mean values of the variables of interest. We check the stability of each learning model at every point in time by examining the highest eigenvalue of the coefficient matrix, disregarding the entire draw when the stability condition is not met. The number of draws we discard is small.[13]

Since our model is calibrated for monthly frequencies and GDP data are reported only in quarterly frequencies, we then convert the monthly simulated series into quarterly frequencies following Gertler and Trigari (2009). We transform the simulated series from absolute deviations into percentage deviations. Table 4 reports the statistical properties of the simulated series of interest under learning for all three belief specifications. In all tables and figures, a hat over a variable denotes the percentage deviation of the variable from its steady-state value (e.g., $\hat{y} = \tilde{y}/\bar{y}$). Table 4 shows that the RE model generates very little amplification in labor market variables. The table indicates that the search and matching model under learning can replicate the second moments of the U.S. labor market remarkably well. The two learning models provide a good match for the relative volatility in vacancies, unemployment and labor market tightness. For instance, in the AR(2) specification, vacancies and unemployment are about 6.87 and 5.33 times more volatile than output respectively with the corresponding numbers in U.S. data being 9.30 and 8.84. The dispersion of vacancies and unemployment relative to output in the AR(1) model are closer to the data, 7.55 and 9.52, respectively. The learning models do significantly better than the RE model in matching amplification in the data. The correlations and autocorrelations of the autoregressive models are in line with the RE model and data. In

13. This choice of dealing with instability at each point in time is similar to the procedure used to estimate Bayesian vector autoregressive (BVAR) models. We do not implement a projection facility because we prefer to avoid contaminating the statistics with unstable draws. The initial values of parameter estimates in each regression equation is typically set at zero with the intercept terms set equal to the corresponding deterministic steady value of the variable. The precision matrices are set equal to the identity matrix to start the simulations. An advantage of setting initial beliefs in this manner is that it makes them consistent (symmetric) across the different belief specifications, namely, AR(1) and AR(2) beliefs. We have also experimented with other initial values and our results do not change significantly.

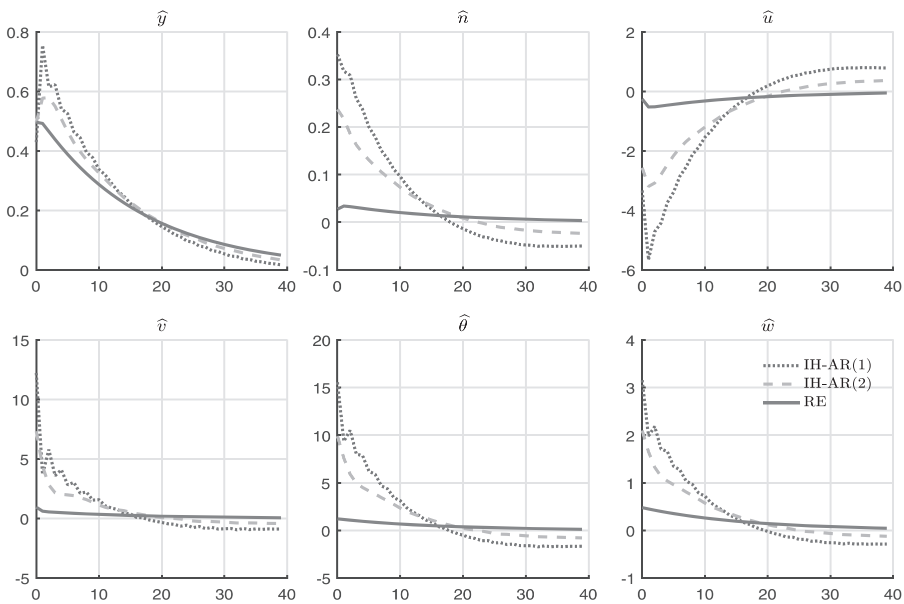

Fig 3. Impulse Responses to Labor Market Variables.

NOTES: Impulse responses to a one standard deviation (productivity) shock. The solid line is RE, dotted line AR(1), and dashed line AR(2) beliefs. Percentage deviations from steady-state values reported along the vertical axis. The horizontal axis displays the number of quarters after the shock.

particular, all models generate a negatively sloped Beveridge curve (-0.92 for RE, -0.85 for AR(1) and -0.81 for AR(2)).

3.4 Impulse Responses

In this subsection, we study how labor market variables respond to a TFP innovation and compare the dynamics under RE and learning. Following Eusepi and Preston (2011), the impulse responses of the learning model are computed by simulating the model twice over $10,000 + 120$ periods. We add to the first simulation a positive one standard deviation (productivity) shock in period $10,001$ and compute the impulse responses as the difference between the two resulting set of impulse responses from period 10,001 onward. This experiment is then repeated 100 times and the mean impulse responses of the variables of interest are reported. The simulated series are *converted into quarterly frequencies* and then *expressed in percentage deviations from steady-state values*.

Figure 3 shows the impulse responses of aggregate output, (un)employment, vacancies, labor market tightness, and the wage rate to a positive TFP innovation under learning and RE. The impulse responses under RE display negligible amplification relative to the learning models. Firms endowed with RE beliefs correctly understand

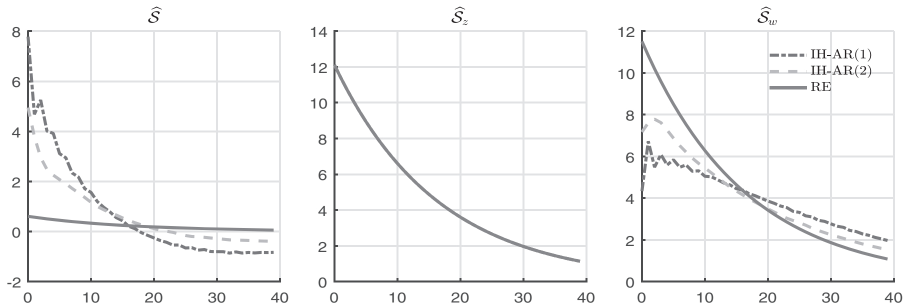

Fig 4. Impulse Responses of Infinite Sums in Job Creation Condition.

NOTES: Impulse responses to a one standard deviation (productivity) shock. The solid line is RE, dashed dotted line AR(1), and dashed line AR(2) beliefs. Percentage deviations from steady-state values reported along the vertical axis. The horizontal axis displays the number of quarters after the shock.

the equilibrium restrictions that determine future wages. For this reason, expected wages tend to absorb most of the productivity increase and, as a result, labor market variables respond only marginally to a TFP innovation.

Very distinct dynamic responses are observed under learning to a positive TFP innovation for all two belief specifications. Following a positive technology shock, the incentive for vacancy creation increases sharply on impact, leading to more employment and a sharp fall in unemployment. The response of employment is largest for the AR(1) beliefs followed by AR(2) beliefs; the amplification generated in the learning models is, however, much larger than that under RE. Quantitatively, the magnitude of the responses is preserved for the next 4 years, that is, it is greater under AR(1) than AR(2). After 4 years, the magnitude of these responses is reversed, that is, it is smaller under AR(1) than under AR(2). We observe over- and undershooting of the variables under AL compared to RE as they converge to the steady state.

Figure 3 shows that learning increases the internal propagation mechanism of the model. In particular, the response of output significantly exceeds that of the TFP innovation on impact. In contrast, under RE the magnitude of the response of output is approximately the same as that of the TFP innovation, which is another feature of the *unemployment volatility puzzle*. The reason behind the greater response in output under learning has to do with a larger expansion in employment.

To disentangle further these effects under learning and gain intuition, we rearrange the job creation condition, equation (16), in linearized form,

$$-\frac{\kappa}{\bar{q}^2}\tilde{q}_t = \mathcal{S}_t \equiv \mathcal{S}_{z,t} - \mathcal{S}_{w,t}, \tag{28}$$

and we then plot the impulse responses of the infinite sums. As shown earlier, the infinite sum $\mathcal{S}_t$, which denotes the present discounted value of profits per hire, responds much more strongly under AL relative to RE. Figure 4 decomposes this infinite sum ($\mathcal{S}_t$) into the two components ($\mathcal{S}_{z,t}$ and $\mathcal{S}_{w,t}$). In the figure, a tilde over $\mathcal{S}_t$ denotes the infinite sum expressed in percentage deviation from the steady-state value (analogously for $\mathcal{S}_{z,t}$ and $\mathcal{S}_{w,t}$). In the subsequent discussion, we refer to respective variables expressed in percentage deviations. Note that $\mathcal{S}_{z,t}$ is the same under RE and learning and, hence, the lines overlap. The size of the responses of both $\mathcal{S}_{z,t}$ and $\mathcal{S}_{w,t}$ under RE are of equal magnitudes, which explains the negligible amplification generated under RE. In sharp contrast, the magnitude of the response of $\mathcal{S}_{z,t}$ is much greater than that of $\mathcal{S}_{w,t}$ under learning for the two belief specifications. The difference between $\mathcal{S}_{z,t}$ and $\mathcal{S}_{w,t}$ explains the strong amplification mechanism of the learning models; this difference depends primarily on the forecasting models used by agents.

Mechanism of amplification. We now examine the transmission mechanism of TFP innovations under different belief specifications. We first describe the mechanism under RE beliefs in order to understand why the model fails to generate large fluctuations in labor market quantities and then explain how unemployment volatility is generated under AL.

A TFP shock has well-understood implications in the standard search and matching model under RE beliefs. A positive innovation (ϵ_t) shifts the production frontier and the labor demand schedule, by increasing labor productivity, which raises marginal profits per hire. Since wages in the standard model are flexible and negotiated through the process of Nash bargaining, a shift in technology increases both future marginal products of employment and wage costs, leading to little vacancy creation.

We rewrite equation (28) approximately as the presented discounted value of profits, which consists of the difference between infinite sums of revenues and costs,

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\bar{q}}{\kappa}\mathbb{E}_t\sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j\big[\tilde{z}_{t+j}-\tilde{w}_{t+j}\big] \equiv \frac{\bar{q}}{\kappa}\big(\mathcal{S}_{z,t}-\mathcal{S}_{w,t}\big). \tag{29}$$

We replace $\mathbb{E}_t^*$ by $\mathbb{E}_t$, the expectation operator under RE. Since unemployment is a predetermined variable, this expression pins down how many vacancies are posted in equilibrium. Note that vacancy creation depends on the present discounted value of profits per hire, which responds only little under RE. The RE solution expressed in terms of deviations from steady-state values may be written as

$$\tilde{n}_{t+1} = \bar{a}_{nn}\tilde{n}_t + \bar{a}_{nz}\tilde{z}_t \quad \text{and} \quad \tilde{w}_t = \bar{a}_{wn}\tilde{n}_t + \bar{a}_{wz}\tilde{z}_t,$$

where $\bar{a}_{x_1x_2}$ denotes the elasticity of variable x_1 with respect to variable x_2. Plugging the RE solution into (29) yields

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\beta\bar{q}}{\kappa}(\bar{\varphi}_1\tilde{z}_t - \bar{\varphi}_2\tilde{n}_t - \bar{\varphi}_3\tilde{z}_t), \tag{30}$$

where

$$\bar{\varphi}_1 = \frac{\varrho}{1 - \beta(1-\rho)\varrho}, \quad \bar{\varphi}_2 = \frac{\bar{a}_{wn}\bar{a}_{nn}}{1 - \beta(1-\rho)\bar{a}_{nn}} \quad \text{and,} \tag{31}$$

$$\bar{\varphi}_3 = \frac{\bar{a}_{wn}\bar{a}_{nz}(1 - \beta(1-\rho)\varrho) + \varrho\bar{a}_{wz}(1 - \beta(1-\rho)\bar{a}_{nn})}{[1 - \beta(1-\rho)\bar{a}_{nn}]\big[1 - \beta(1-\rho)\varrho\big]}. \tag{32}$$

This expression provides the key insight into the lack of amplification under RE. At the outset, vacancy posting does not respond to employment because $\tilde{n}_t$ is predetermined; it is next period employment that affects vacancy posting. There is a direct effect from productivity innovations, $\tilde{z}_t$, to the vacancy posting decision through shifts in future returns from employment (term $\bar{\varphi}_1\tilde{z}_t$ in equation (30), which corresponds to the first infinite sum in (29)) and an indirect effect from productivity innovations through shifts in future wage costs (the terms $\bar{\varphi}_2\tilde{n}_t$ and $\bar{\varphi}_3\tilde{z}_t$ in equation (30), which correspond to the second infinite sum in (29)). For the baseline calibration, it can be shown that $\bar{\varphi}_1$ and $\bar{\varphi}_3$ in equation (30) are very close to each other, which implies that productivity innovations have little impact on the present discounted value of profits per hire.[14] In other words, when firms correctly forecast future profits per hire, the variability of profits per hire is dampened because wages absorb large part of the innovation. As a result, the RE model is unable to generate sufficient amplification in vacancies.

We now turn to the explanation of amplification under learning. We first note that under RE firms have knowledge of the equilibrium restrictions that determine future wages and employment; in particular, that a positive productivity shock leads to higher wages and employment in the future. Under learning, on the other hand, firms are not aware of these equilibrium restrictions, that is, that future wages and employment will be higher due to higher productivity. In the learning model agents are unable to understand the implications of future wage negotiations and instead estimate simple autoregressive models to forecast future wages using past information.

We restrict ourselves to discussing the results with AR(1) beliefs. A similar logic applies to AR(2) beliefs. Under AR(1) beliefs, firms estimate the parameter a_1 by running the following regression: $\tilde{w}_t = a_1\tilde{w}_{t-1} + \hat{\eta}_{wt}$, to make their forecasts of future wages. Combining this expression with (29) yields

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\beta\bar{q}}{\kappa}(\varphi_1\tilde{z}_t - \varphi_2\tilde{w}_{t-1}), \tag{33}$$

where $\varphi_1 = \frac{\varrho}{1-\beta(1-\rho)\varrho}$ and $\varphi_2 = \frac{a_1^2}{1-\beta(1-\rho)a_1}$. A TFP innovation increases the discounted value of future marginal products of employment, the term $\varphi_1\tilde{z}_t$ in

14. For the baseline calibration, apart from $\bar{a}_{wn} = 0$, the values of the elasticities $0 < \bar{a}_{x_1x_2} < 1$ for the other possible combinations of x_1 and x_2. In the general equilibrium version of the model presented in Section 6.3, a similar intuition holds ($0 < \bar{a}_{x_1x_2} < 1$ for all possible combinations of x_1 and x_2 there). The baseline calibration ensures that our results are not driven by the strategy in Hagedorn and Manovskii (2008).

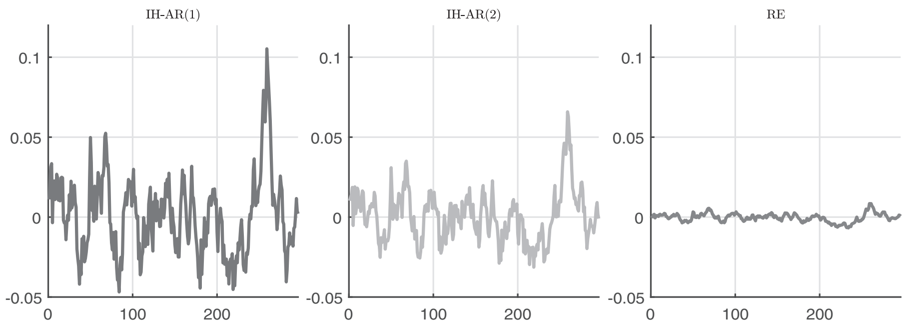

Fig 5. Simulated Path of Infinite Sums.

NOTES: Infinite sums are expressed in percentage deviations from steady-state values. The first 10,000 months are discarded and the remaining 900 months averaged over 100 replications under RE and AL. Monthly simulated series are converted into quarterly series reflecting 300 quarters.

equation (33), but has no direct impact effect on the present discounted value of future wage costs, the term $\varphi_2 \tilde{w}_{t-1}$. The high persistence of the TFP innovation, together with the low value of the separation rate, is key for generating amplification under learning because it raises the rate at which future marginal products of employment are discounted. The fact that expectations are predetermined implies that firms overestimate the present discounted value of profits per hire, which then translates into further incentives for vacancy creation. Thus, the labor market becomes tighter and the job filling rate falls after a TFP innovation, which then raises the expected costs of vacancy posting. After a productivity shock, the present discounted value of future revenues is much larger than that of future wage payments. More vacancies in equilibrium translate into higher employment and output.[15]

Figure 5 plots the path of the simulated infinite sum, $\mathcal{S} = \mathcal{S}_z - \mathcal{S}_w$, in percentage deviations from the steady state (i.e. the present discounted value of profits per hire) for the different models over 300 quarters. Under RE the two sums are very close to one another, explaining the lack of amplification in the model. The figure shows that the difference in the two sums is greatest under AR(1) beliefs followed by AR(2) beliefs. This means that learning models can better match the relative volatility of labor market variables in the U.S. data. The persistence of the simulated series is highest under AR(1) and AR(2) beliefs, helping explain the greater incentive for vacancies posting.[16]

15. Our study is concerned with understanding unemployment amplification over the post World War II period, as is done in the literature. If one focuses on the narrower period, in the years following the financial crisis in 2007, professional forecasters were consistently overly optimistic about output growth. An explanation has been that there was incomplete understanding of ongoing structural changes in the economy. In our model, this overoptimism would result in strong job creation, but what was observed, in contrast, is a jobless recovery. Of course, the Great Recession is an exceptional period and our model is not equipped to study this period missing several important features, for example, financial market imperfections.

16. See Section F of the Online Appendix for the coefficients of the actual law of motions in the different learning models.

4. FORECAST ERROR PROPERTIES

Vacancy posting decisions in search and matching models typically depend on the marginal profitability of long-term employment relationships. We argue that wage forecasts play a key role for job creation. Under RE firms can perfectly assign a value to a filled vacancy because they have full information and knowledge about the structure of the economy. As shown in Section 3, relaxing these assumptions can lead to more job creation. Given that wage forecasts are central for explaining amplification in labor market data, a first suitable test of the AL models would be to compare how well the statistical properties of wage forecasts generated by the AL models (and RE) fair against those in the data. We argue that the AL models are more successful than RE at matching wage forecast data.

Data published by the EC on annual nominal wage and price inflation for the United States is used to conduct the comparison.[17] These forecasts are released twice a year in Autumn and Spring (in Q1 and Q3). It is worth noting that wage forecast data are only available from 1999Q3 to 2017Q3 and the mean forecast of nominal wage compensation per head is only reported in annual growth space. To compute the equivalent of 1-year ahead *real* wage growth forecast, we simply subtract 1-year ahead price inflation forecast from the nominal wage inflation forecast. We then compute forecast errors as the difference between realized and forecasted annual real wage growth.[18] The 1-year ahead real wage forecasts for the United States are reported in Figure 1 and for some other developed economies in Figure 2. A caveat of these data is that it is reported biannually and the time span is relatively short. Another caveat is that the data are not entirely consistent with the definition of wages in the model because, while the model only considers the extensive margin, hours worked are part of real wage compensation per employee.[19] A third caveat is that average wages are subject to compositional effects and several studies, such as Pissarides (2009) and Haefke, Sonntag, and van Rens (2013), report that wages for newly hired workers are more cyclical than average wages.[20]

To be consistent with the availability of data, we calculate *1-year ahead* wage forecasts from the different models. We compute forecast errors as the difference between

17. The annual forecasts of nominal wage compensation per employee and CPI inflation are taken from Supplement *A* of Economic Trends, which are published by the EC, mostly biannually since Autumn 1999. The EC started publishing its forecasts three times per year only from 2013. We discard the additional five observations to compute the forecast error statistics with regularly spaced data as in the models.

18. The real wage growth forecasts are calculated using inflation forecasts based on the GDP deflator; the results are very similar if CPI inflation is used.

19. An earlier version of the paper, Di Pace, Mitra, and Zhang (2016), used unemployment forecast errors (derived from SPF forecasts) over the period 1968–2015. It was shown that AL models were able to replicate some of the unemployment forecast error properties. The focus here is on wage forecasts since this constitutes more direct empirical evidence for the key mechanism of unemployment amplification in the AL models studied here. We thank an anonymous referee for pointing us in this direction.

20. For instance, due to the cleansing effect of recessions (Caballero and Hammour (1994)), many low paid workers are dismissed during recessions, and are then rehired next period. As a result, the real aggregate wage is much less cyclical than individual wages (Solon, Barsky, and Parker (1994) and Bowlus, Liu, and Robinson (2002)). Given that all workers in our model are paid the same wage, the response of the individual wage is as elastic as that of aggregate wages.

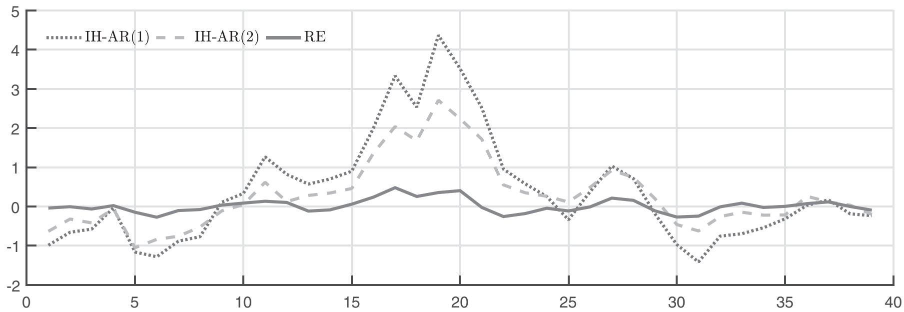

Fig 6. Simulated Annual Wage Forecast Errors in the RE and Learning Models.

NOTES: We compute the forecast errors as $\tilde{x}_{1t} = \hat{w}_{t+12} - \mathbb{E}_t^* \hat{w}_{t+12}$, and we convert monthly into quarterly forecast errors by averaging. Finally, we sample the last 76 quarters of simulated data and discard 2 quarters (Q2 and Q4) to aid comparability with the data. The solid line is RE, dotted line is AR(1), and dashed line is AR(2) belief.

TABLE 5

STATISTICAL PROPERTIES: FORECAST ERRORS OF ANNUAL WAGE GROWTH

	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	$\sigma_{\hat{x}_{1t}}$
Data	0.63	0.65	1.68	2.48
RE	0.20	0.63	0.84	0.18
$IH-AR(1)$	0.48	0.86	4.61	1.37
$IH-AR(2)$	0.42	0.85	3.20	0.86

NOTES: The term $\rho(x_{1t}, x_{2t})$ stands for the correlation coefficient between variables x_1 and x_2. Data from European Commission. Annual forecast errors are computed as $\hat{x}_{1t} = \hat{w}_{t+12} - \mathbb{E}_t^* \hat{w}_{t+12}$ and monthly forecast errors are transformed into quarterly forecast errors by taking a 3-month average. To compare the performance of the models with the data, we sample 76 quarters from the simulated series and discard two quarters out of four (Q2 and Q4).

actual and forecasted real wage growth, $\hat{x}_{1t} = (\hat{w}_{t+12} - \hat{w}_t) - \mathbb{E}_t^*(\hat{w}_{t+12} - \hat{w}_t) = \hat{w}_{t+12} - \mathbb{E}_t^* \hat{w}_{t+12}$. We average monthly forecasts into quarterly forecasts and sample the last 76 quarters of simulated data to have the same number of observations as in the actual data (i.e., we discard two quarters out of four, the equivalent of Q2 and Q4 in the data). For a visual feel, Figure 6 plots the simulated forecast errors of annual wage growth for the RE and AL models. Noticeably, the magnitude (fluctuation) in forecast errors generated by the AL models is much larger than under RE. Note that all forecast errors (including RE) exhibit some degree of serial correlation because these are annual forecasts (rather than shorter horizon forecasts).

We next compare the statistical properties of forecast errors generated by the different models with those in the data. The learning models get close to matching the data in several aspects. Table 5 reports the properties of forecast errors in the models and the data: the absolute standard deviation of forecast errors, the volatility of forecast errors (relative to annual output growth), the first-order autocorrelation, and the correlation with annual output growth. The absolute volatility of forecast errors and its correlation with output growth in the learning models are more in line with the data than RE. In particular, the forecast errors generated by the AR(1) model is much more volatile in absolute terms (1.37) than the RE (0.18) counterpart though they are still less volatile than the data (2.48). The procyclicality of forecast errors in the learning models (0.48 and 0.42) is more in line with the data (0.63) than the RE model (0.20). Over the business cycle, the forecast errors are procyclical, meaning that agents tend to underpredict wages during expansions and overpredict them in recessions.

Forecast errors fluctuate more than output in the data. Unlike RE, there is excess volatility of forecast errors vis-à-vis output in the learning models relative to the data. The RE model predicts about the right amount of serial correlation in the forecast errors whereas the AL models deliver higher auto-correlation. Arguably, it is surprising that even the simple and stylized AL models are able to match forecast data along these dimensions.

5. DISCIPLINING AGENTS' BELIEFS

In this section, we analyze whether agents' beliefs are consistent with actual data and with those in the model *ex post*.[21] We focus on monthly wage data for two reasons: (i) actual wage data are available at monthly frequencies and (ii) the model is calibrated to monthly frequencies.[22]

Loosely speaking, use is made of a model selection strategy popularized by Box and Jenkins. Pure AR models have been advocated by time series analysts (e.g., Granger and Newbold (1986)) as parsimonious models (over ARMA models) on the grounds of being simpler to estimate and (more importantly) easier to specify because no identifiability problems arise in a procedure of "testing down" (using a general-to-specific approach) to see if the model could be simplified (see Harvey 2008, pp. 78–80 for a discussion).

We first summarize the different methods used on monthly real wages for the United States (which are suggestive that wages be modeled as an AR(1) process). First, partial autocorrelation plots are used to identify the order of the autoregressive model that best fits the U.S. data. Partial autocorrelation plots (Box et al. (2015)) are a commonly used tool for identifying the order of an autoregressive model. In addition, use is made of a different approach to model selection, the Bayesian information criterion (BIC), which is based on a goodness-of-fit criterion. Finally, another method (namely the Wald test) is used to determine the lag order of a generic autoregressive process. All of these approaches favor an AR(1) model for monthly real wages in the data.

21. Adam, Kuang, and Marcet (2012) follow a similar approach to discipline agents' beliefs in their model.

22. Monthly real wage data are computed using data from FRED. We calculate monthly nominal wages as the ratio between seasonally adjusted wage compensation of employees (W209RC1) and all employees: total nonfarm payrolls (PAYEMS). Nominal wages are deflated by Consumer Price Index for All Urban Consumers (CPIAUCSL). The monthly (real) wage data are detrended using a Hodrick–Prescott filter with smoothing parameter 14400.

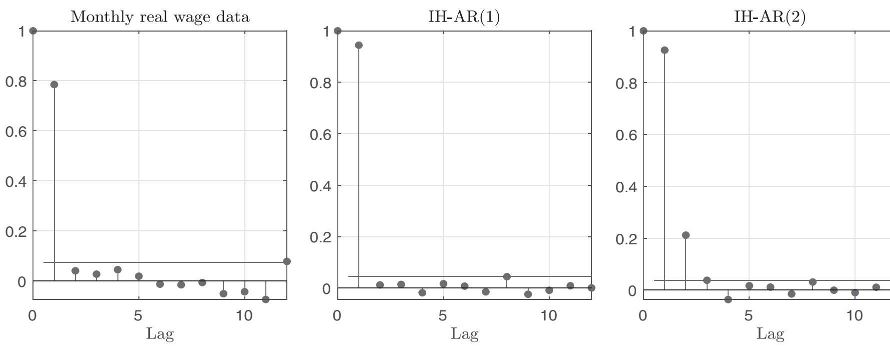

Fig 7. Partial Autocorrelation Coefficients (Data vs. Learning Models).

NOTES: Monthly real wage data series is detrended using the HP filter with smoothing parameter 14,400. First panel reports the partial autocorrelation coefficients of the detrended data series. Partial autocorrelation coefficients of the simulated wage series originating from the learning models with autoregressive beliefs are reported in the second and third panels. We use different sample sizes for the learning model because learning is faster when there are fewer parameters to be estimated (2,000 periods under AR(1) and 3,000 periods under AR(2) beliefs). Blue lines show an indication of sampling uncertainty.

We next show that the baseline model featuring autoregressive beliefs generate simulated wage series whose behavior is consistent with the U.S. data. Motivated by the model selection strategy outlined above, agents try to find an initial specification of the model based on the sample partial autocorrelation function. The partial autocorrelation of an AR(p) process is zero at lag $p+1$ and greater. Figure 7 shows the sample partial autocorrelation coefficients for the U.S. monthly wage data (first panel) and compares it with the learning models (remaining panels). The sample partial autocorrelation plots indicate an autoregressive process of order 1 for the U.S. data. The middle and right panels show that, while learning toward their long-run equilibria, agents beliefs' are in line with the simulated wage data *ex post*. Thus, using partial autocorrelations, agents would fail to detect misspecification for very long periods of time.

We further investigate whether agents would be able to statistically detect any misspecification *ex post*. Agents run autoregressions of a higher order than their beliefs using simulated data to select the order of the autoregressive process (a general-to-specific approach). Here, agents use the BIC to optimally select the order of the autoregression and they conduct Wald tests to test down the autoregressive beliefs for an extended period of time. Note that under constant gain learning agents discount past data more heavily. Consistent with the assumption of $\gamma = 0.002$, we assume that agents use 500 periods worth of wage data and run rolling regressions up to and including period 4, 000.

Agents use the BIC to select the lag length of the autoregressive model. Figure 8 illustrates that agents with AR(1) beliefs are unable to select between AR(1) and AR(2) beliefs. The top left panel of the figure shows that only after 2,000 periods (months) can agents start differentiating between the two processes (since the BIC

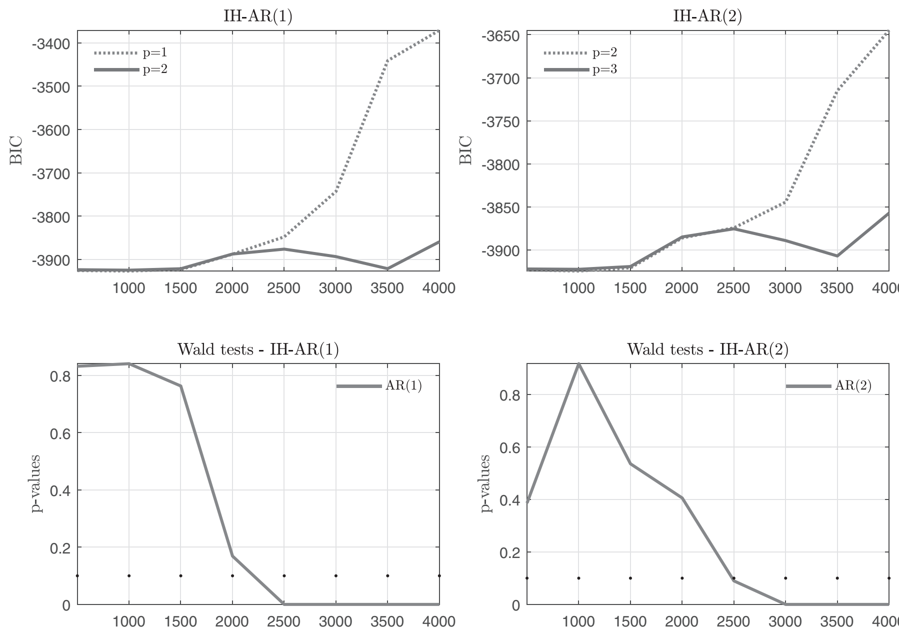

Fig 8. Detecting misspecification of beliefs.

NOTES: The top left panel indicates the Bayesian information criterion (BIC) over time from two autoregressive models (AR(1) and AR(2)) using wage data generated from the model with AR(1) beliefs. A lower value indicates the preferred model. The top right panel shows the BIC over time from two autoregressive models (AR(2) and AR(3)) using wage data generated by the model with AR(2) beliefs. The bottom left panel reports the *p*-values of the Wald test whose null hypothesis that the second order autoregressive coefficient of an AR(2) process is zero in the model with AR(1) beliefs. The bottom right panel reports the *p*-values of the Wald test where the null hypothesis is that the the third-order autoregressive coefficient of an AR(3) process is zero in the model with AR(2) beliefs. In the bottom panels, the dotted line indicates the 10% significance level.

indicates a lag length of order 2, i.e., $p = 2$). A similar pattern also arises for the model featuring AR(2) beliefs. With AR(2) beliefs, agents are unable to detect *ex post* misspecification for about 2,500 periods. It is only if agents continue to run regressions for a very extended period of time (200 years) would they be able to detect misspecification in their beliefs. Note, however, that the actual data used in our analysis are of a much shorter horizon (700 months), so agents would be unable to reject their beliefs based on the sample sizes that are in practice available.

Next, agents test the null hypothesis that the PLM is misspecified by computing the W-statistic.[23] Thus, agents with AR(1) beliefs would like to test the hypothesis that the higher order autoregressive coefficients of a generic AR(p) process are zero. The bottom left panel of Figure 8 reports the *p*-values of the Wald tests over time.[24]

23. See Adam, Kuang, and Marcet (2012) for a similar test.

24. We only report the result of the test in which agents with autoregressive beliefs of order p compare their model with an alternative model of order $p + 1$. If detected, the degree of misspecification is not higher than order 1.

Since the p-values are above the 10% threshold for around 2,000 periods, agents are yet again not able to detect misspecification *ex post* for very long periods. The same pattern arises for agents with AR(2) beliefs. The bottom right panel shows that agents cannot reject the hypothesis that the coefficient on the third lag of an AR(3) model is different from zero for around 2,500 periods. These results are consistent with the BIC analysis shown in the top panels of Figure 8.

6. MODEL EXTENSIONS

In this section, we examine the performance of the baseline model to three extensions, namely, by allowing for (i) search effort, (ii) endogenous separation, and (iii) general equilibrium considerations. We describe the new features of each model extension relative to the baseline model in Online Appendix B. We maintain the same informational assumptions outlined in Section 2 with agents continuing to forecast infinite steps ahead. For brevity, we only focus on AR(2) beliefs in Sections 6.1 and 6.2 (though the results are not confined to these beliefs). We find that the main results carry over to the different model extensions. All tables and figures we refer to in Sections 6 and 7 are in Online Appendices C, D, and E.

6.1 Search Effort

We assess the robustness of the results to introducing search effort to the standard search and matching model in line with recent work by Mukoyama, Patterson, and Şahin (2018). We make the assumption that the functional form of the matching function is Cobb–Douglas to keep the analysis consistent with the baseline model. We find that allowing for search effort does not alter the main message of the paper.

The impulse responses of labor market variables to a positive TFP innovation under IH-AR(2) learning and RE are broadly in line with Figure 3 (see Figure 1 in Online Appendix D). Following a positive productivity shock, the incentive for vacancy creation rises sharply on impact, resulting in more employment (and lower unemployment). Intuitively, as the labor market tightens and the job filling probability falls, both the marginal cost and benefit associated with posting an additional vacancy increase. As firms become more productive, however, workers exert more effort to find jobs (but in doing so they do not internalize that their actions increase average job search). The rise in job intensity in turn improves matching efficiency endogenously, which increases the likelihood of workers and firms to find each other. For this reason, the internal propagation mechanism of the model is stronger relative the baseline model.

Mukoyama, Patterson, and Şahin (2018) find that search effort under RE does not act as an amplifier of labor market fluctuations. In line with the baseline model, we find that amplification generated under AL is greater than that under RE (see Table 1, Online Appendix C). In particular, we notice that the extent to which the simulated path of the infinite sum $\mathcal{S}_t$ responds to productivity shocks is indicative of the strength of the AL mechanism. Relative to the baseline model, vacancies and the job filling probability are less responsive to a TFP innovation. However, this is counteracted by a highly cyclical search effort and greater unemployment amplification.

6.2 Endogenous Separation

Next, we examine the sensitivity of the results to allowing for endogenous separation. We follow Krause and Lubik (2007a) and Trigari (2009) in modeling the firms' problem but omit general equilibrium considerations.[25] The behavior of the model featuring endogenous separation under RE is very different from the baseline model in that a positive productivity innovation leads to a protracted *fall* in job creation (see Figure 2, Online Appendix D). Less jobs get destroyed, offsetting the contraction in job creation and reducing unemployment. Thus, the firing margin becomes dominant relative to the creation margin. Despite less vacancy creation, the labor market tightens and firms find it easier to fill vacancies. A tighter labor market strengthens the fall in the cutoff level of idiosyncratic productivity, which in turn reduces job destruction. It is well known that the model featuring endogenous separation generates an upward sloping Beveridge curve that is inconsistent with the data.[26]

Despite the clear dominance of the firing margin, the amplification generated after a TFP innovation is greater under learning. Job creation and destruction are much more volatile relative to RE (see Table 2, Online Appendix C). By comparing learning models (with and without endogenous separation), vacancies and the job filling probability become less cyclical. At the same time, a highly cyclical job separation rate helps generate more unemployment volatility (relative to the data). The behavior of tightness is roughly in line with the baseline model under AR(2) beliefs (but lower than the data). This is because vacancies and unemployment become less and more cyclical, respectively. The slope of the Beveridge curve under learning is flatter (0.47), though still positive, relative to the RE model (0.95).

6.3 General Equilibrium

We extend the standard search and matching frictions to allow for general equilibrium considerations. For more details, see Online Appendix B.3 and Di Pace, Mitra, and Zhang (2016). In this extension firms not only forecast the future path of wages but also those of employment, interest rates, and profits. For this reason, in addition to univariate autoregressive beliefs, we also study the sensitivity of the results to allowing for potential general equilibrium effects when forecasting the value of future variables. One way to introduce such interactions is for economic

25. Their set-up assumes that firms are large and hire multiple workers with different productivity levels. We make these additional assumptions to facilitate comparability with the baseline model whilst keeping the model relatively tractable. Further model details are provided in the Online Appendix B.2.

26. This anomaly results from the assumption that, whilst firing workers is instantaneous and costless, hiring workers takes time and is costly. This result arises irrespective of whether we assume rational or boundedly rational agents. For an example where firing costs are introduced in a search and matching model with endogenous separations and small firms, see Thomas (2006).

agents to compute their forecasts using vector autoregressive (VAR) models. We find that the results from the baseline model are not sensitive to general equilibrium considerations.

This version of the model under all three autoregressive beliefs generates a great deal of amplification relative to the data (see Table 3, Online Appendix C). It is worth noting, however, that under VAR(1) beliefs the Beveridge curve becomes flatter than under univariate autoregressive beliefs. The learning models generate impulse responses that are observationally equivalent to the baseline model (see Figure 3, Online Appendix D).

7. OTHER LEARNING SPECIFICATIONS

In this section, we examine the performance of the baseline model under different beliefs and alternative assumptions about the length of the forecast horizon and timing. First, we assume that firms' beliefs are of the RE form, and they learn the parameter values as new information becomes available. Second, we assume that firms, instead of forecasting infinite steps ahead, forecast only one-step ahead, so as to make vacancy posting decisions. Third, we assume that period wages are part of the information set of agents.[27]

7.1 Correctly Specified Beliefs

If agents have beliefs that are of the same form as equation (24) but have no knowledge about the value of the RE parameters ($\bar{b}_x$, $\bar{a}_{xn}$, and $\bar{a}_{xz}$), they would attempt to learn them as new information becomes available. *Correctly specified* beliefs are given by

$$x_t = b_x + a_{xn} n_t + a_{xz} \tilde{z}_t + \eta_{xt}, \quad \text{for} \quad x_t = \{n_{t+1}, w_t\}, \tag{34}$$

where η_{xt} are white noise processes. Agents estimate parameters b_x, a_{xn} and a_{xz} by running regressions of variables (such as wages) on previous period employment and current TFP in order to make vacancy posting decisions. We assume that agents continue to forecast infinite steps ahead with such beliefs.

We find that the AL model featuring correctly specified beliefs suffers from the same problems as the RE model; that is, this model is neither able to provide a solution to the unemployment volatility puzzle nor to generate large and systematic forecast errors. These findings should not be surprising because loosely speaking,

27. The model statistics are computed after convergence to the long-run equilibrium. In reality, agents do not know the nature of this equilibrium and would have to learn it over a long period of time. Irrespective of the assumptions and the form of learning, as agents learn their way to the equilibrium, it is important to note that AL delivers labor market statistics and forecast error properties that are more in line with the data, compared to RE. Thus, learning plays an important role. Of course, the disadvantage with computing statistics off the equilibrium is that the time period may matter, with its choice being arbitrary. Thus, we always report results once convergence to the equilibrium has taken place.

with constant gain learning, the model with correctly specified beliefs converges in distribution to the RE solution (Evans and Honkapohja 2001, chapter 7.4). The value of the parameters b_w, a_{wn}, and a_{wz} are very close indeed to those under RE ($\bar{b}_w$, $\bar{a}_{wn}$, and $\bar{a}_{wz}$). This may provide an explanation as to why the model generates impulse responses that are undistinguishable from those under RE.

7.2 *Euler Equation Learning*

There is an alternative form of learning widely used in the literature, for example, see Bullard and Mitra (2002) and Evans and Honkapohja (2006), where agents make one-step ahead forecasts. This is in contrast to the IH learning approach, assumed so far in this paper, where agents make forecasts infinite steps ahead. This approach would take equation (11) as the behavioral rule of firms. Under EE learning, in addition to making one-period ahead wage forecasts, firms make forecasts about labor market tightness ($\tilde{\theta}_t$) by following simple autoregressive learning rules. We find that EE learning delivers results that are not too far off from RE.

Thus, it turns out that the type of learning agents engage in is important for matching features observed in the data. Models featuring EE learning tend to generate greater persistence in $\tilde{\theta}_t$, which is self-defeating for amplification. This is partly because, under EE learning, the persistence of $\tilde{\theta}_t$ is assumed in the autoregression. Models with IH learning, on the other hand, tend to generate a great deal of amplification because agents over (under)estimate the impact the TFP shock on future profits associated with long-term employment relationships. We note that agents that engage in long-term decisions, such as hiring, are more interested in forecasting over longer time horizons (e.g., most Central Banks or International Organizations forecast a large set of variables over 3 years).

7.3 *Informational Assumptions*

We have so far assumed that agents do not observe period wages at the time they make their forecasts, that is, they only have information up until the last period and forecasts are conditioned on previous allocations. We now relax this assumption, that is, agents observe period wages when making their forecasts. To explain the intuition behind this result we refer to Section 3.4 and confine ourselves to AR(1) beliefs. Firms estimate the parameter a_1 by running the regression $\tilde{w}_t = a_1 \tilde{w}_{t-1} + \eta_{wt}$, to make wage forecasts. With wages observed at time t, forecasts will be given by $\mathbb{E}_t^* \tilde{w}_{t+j} = a_1^j \tilde{w}_t$. Combining this expression with equation (29) yields

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\beta \bar{q}}{\kappa}(\phi_1 \tilde{z}_t - \phi_2 \tilde{w}_t), \tag{35}$$

where $\phi_1 = \frac{\varrho}{1-\beta(1-\rho)\varrho}$ and $\phi_2 = \frac{a_1}{1-\beta(1-\rho)a_1}$. A TFP innovation increases the discounted value of future marginal products of employment, the term $\phi_1 \tilde{z}_t$ in equation (35), but has now a direct impact on the present discounted value of future wage costs, the term $\phi_2 \tilde{w}_t$. Note that wages depend on $\tilde{z}_t$ and $\tilde{\theta}_t$; see equation (19). Wages therefore contain information about productivity at time t, much like under RE. After a productivity innovation, the present discounted value of marginal revenues increases only slightly relative to the present discounted value of wage payments, leading to little vacancy creation. Therefore, amplification, impulse responses and forecast errors are not too dissimilar from RE.

8. CONCLUSION

In the standard search and matching model the vacancy posting decision depends crucially on what firms expect the present discounted value of profits per hire to be; this motivates the study of expectation formation at the firm level. In this paper, we relax the assumption of RE to study the role of AL on job creation in the standard search and matching model. We show that the combination of AL with simple forecasting models can match the volatility of U.S. labor market very well, outperforming the standard RE model. In addition, AL is able to match the properties of forecast data on annual wage growth.

LITERATURE CITED

Adam, Klaus, Pei Kuang, and Albert Marcet. (2012) "House Price Booms and the Current Account." *NBER Macroeconomics Annual*, 26, 77–122.

Adam, Klaus, Albert Marcet, and Johannes Beutel. (2017) "Stock Price Booms and Expected Capital Gains." *American Economic Review*, 107, 2352–2408.

Lago Alves, Sergio A. (2018) "Monetary Policy, Trend Inflation, and Unemployment Volatility." *Journal of Money, Credit and Banking*, 50, 637–73.

Barnichon, Regis. (2010) "Building a Composite Help-Wanted Index." *Economics Letters*, 109, 175–78.

Blanchard, Olivier, and Jordi Gali. (2010) "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment." *American Economic Journal: Macroeconomics*, 2, 1–30.

Bowlus, Audra, Haoming Liu, and Chris Robinson. (2002) "Business Cycle Models, Aggregation, and Real Wage Cyclicality." *Journal of Labor Economics*, 20, 308–35.

Box, George, Gwilym Jenkins, Gregory Reinsel, and Greta Ljung. (2015) *Time Series Analysis: Forecasting and Control*, 5th ed. New York: Wiley.

Branch, W. (2006). "Restricted Perceptions Equilibria and Learning in Macroeconomics." In *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, edited by D. Colander, pp. 135–60. Cambridge: Cambridge University Press.

Branch, William A., and George W. Evans. (2006a) "Intrinsic Heterogeneity in Expectation Formation." *Journal of Economic Theory*, 127, 264–95.

Branch, William A., and George W. Evans. (2006b) "A Simple Recursive Forecasting Model." *Economics Letters*, 91, 158–66.

Bullard, James, and Kaushik Mitra. (2002) "Learning about Monetary Policy Rules." *Journal of Monetary Economics*, 49, 1105–29.

J Caballero, Ricardo, and Mohamad L Hammour. (1994) "The Cleansing Effect of Recessions." *American Economic Review*, 84, 1350–68.

Campbell, John Y. (1994) "Inspecting the Mechanism: An Analytical Approach to the Stochastic Growth Model." *Journal of Monetary Economics*, 33, 463–506.

Cho, In-Koo, Noah Williams, and Thomas J. Sargent. (2002) "Escaping Nash Inflation." *Review of Economic Studies*, 69, 1–40.

Christiano, Lawrence, Martin Eichenbaum, and Matthias Trabandt. (2016) "Unemployment and Business Cycles." *Econometrica*, 84, 1523–69.

Christoffel, Kai, and Keith Kuester. (2008) "Resuscitating the Wage Channel in Models with Unemployment Fluctuations." *Journal of Monetary Economics*, 55, 865–87.

Davis, Steven, John Haltiwanger, and Scott Schuh. (1996) *Job creation and destruction*. MIT Press.

den Haan, Wouter J., Garey Ramey, and Joel Watson. (2000) "Job Destruction and the Experiences of Displaced Workers." *Carnegie-Rochester Conference Series on Public Policy*, 52, 87–128.

Di Pace, Federico, Kaushik Mitra, and Shoujian Zhang. (2016) "Adaptive Learning and Labour Market Dynamics." Working papers 633, Bank of England.

Eusepi, Stefano, and Bruce Preston. (2011) "Expectations, Learning, and Business Cycle Fluctuations." *American Economic Review*, 101, 2844–72.

Evans, George, and Seppo Honkapohja. (2001) *Learning and Expectations in Macroeconomics (Frontiers of Economic Research)*. Princeton, NJ: Princeton University Press.

Evans, George W., and Seppo Honkapohja. (2003) "Expectations and the Stability Problem for Optimal Monetary Policies." *Review of Economic Studies*, 70, 807–24.

Evans, George W., and Seppo Honkapohja. (2006) "Monetary Policy, Expectations and Commitment." *Scandinavian Journal of Economics*, 108, 15–38.

Evans, George W., Seppo Honkapohja, and Kaushik Mitra. (2012) "Does Ricardian Equivalence Hold When Expectations Are Not Rational?" *Journal of Money, Credit and Banking*, 44, 1259–83.

Gennaioli Nicola, Yueran Ma, Andrei Shleifer. (2016) "Expectations and Investment." *NBER Macroeconomics Annual*, 30, 379–431.

Gertler, Mark, Luca Sala, and Antonella Trigari. (2008) "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining." *Journal of Money, Credit and Banking*, 40, 1713–64.

Gertler, Mark, and Antonella Trigari. (2009) "Unemployment Fluctuations with Staggered Nash Wage Bargaining." *Journal of Political Economy*, 117, 38–86.

Gomes, Pedro. (2011) "Fiscal Policy and the Labour Market: The Effects of Public Sector Employment and Wages." European Economy - Economic Papers 439, Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.

Granger, Clive, and Paul Newbold. (1986) *Forecasting Economic Time Series*. 2nd ed. Amsterdam: Elsevier.

Greenwood, Robin, and Andrei Shleifer. (2014) "Expectations of Returns and Expected Returns." *Review of Financial Studies*, 27, 714–46.

Guerrieri, Veronica. (2008) "Heterogeneity, Job Creation and Unemployment Volatility." *Scandinavian Journal of Economics*, 109, 667–93.

Haefke, Christian, Marcus Sonntag, and Thijs van Rens. (2013) "Wage Rigidity and Job Creation." *Journal of Monetary Economics*, 60, 887–99.

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

Hall, Robert E. (2005) "Employment Fluctuations with Equilibrium Wage Stickiness." *American Economic Review*, 95, 50–65.

Harvey, Andrew. (2008) *Forecasting, Structural Time Series and the Kalman Filter*. revised edition. Cambridge, MA: Cambridge University Press.

Heemeijer, Peter, Cars Hommes, Joep Sonnemans, and Jan Tuinstra. (2009) "Price Stability and Volatility in Markets with Positive and Negative Expectations Feedback: An Experimental Investigation." *Journal of Economic Dynamics and Control*, 33, 1052–72.

Hertweck, Matthias Sebastian. (2013) "Strategic Wage Bargaining, Labor Market Volatility, and Persistence." *The B.E. Journal of Macroeconomics*, 13, 27.

Hommes, Cars, Joep Sonnemans, Jan Tuinstra, and Henk van de Velden. (2005) "Coordination of Expectations in Asset Pricing Experiments." *Review of Financial Studies*, 18, 955–80.

Huang, K.X.D., Z. Liu, and T. Zha. (2009) "Learning, Adaptive Expectations and Technology Shocks." *Economic Journal*, 119, 377–405.

Krause, Michael, and Thomas Lubik. (2007a) "The (Ir)relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions." *Journal of Monetary Economics*, 54, 706–27.

Krause, Michael, and Thomas A. Lubik. (2007b) "Does Intra-Firm Bargaining Matter for Business Cycle Dynamics?" Discussion Paper Series 1: Economic Studies 2007,17, Deutsche Bundesbank, Research Centre.

Kreps, D. (1998) "Anticipated Utility and Dynamic Choice (1997)." In *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983–1997 (Econometric Society Monographs)*, edited by D. Jacobs, E. Kalai, M. Kamien, & N. Schwartz, pp. 242–274. Cambridge: Cambridge University Press.

Kuang, Pei, and Kaushik Mitra. (2016) "Long-Run Growth Uncertainty." *Journal of Monetary Economics*, 79, 67–80.

Malmendier, Ulrike, and Stefan Nagel. (2016) "Learning from Inflation Experiences." *The Quarterly Journal of Economics*, 131, 53–87.

Menzio, Guido. (2005) "High Frequency Wage Rigidity." Technical report, mimeo.

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

Milani, Fabio. (2007) "Expectations, Learning and Macroeconomic Persistence." *Journal of Monetary Economics*, 54, 2065–82.

Milani, Fabio. (2011) "Expectation Shocks and Learning as Drivers of the Business Cycle." *Economic Journal*, 121, 379–401.

Mitra, Kaushik, George W. Evans, and Seppo Honkapohja. (2013) "Policy Change and Learning in the RBC Model." *Journal of Economic Dynamics and Control*, 37, 1947–71.

Mitra, Kaushik, George W. Evans, and Seppo Honkapohja. (2017) "Fiscal Policy Multipliers in an RBC Model with Learning." *Macroeconomic Dynamics*, 23, 1–44.

Mortensen, Dale, and Eva Nagypal. (2007) "More on Unemployment and Vacancy Fluctuations." *Review of Economic Dynamics*, 10, 327–47.

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

Mukoyama, Toshihiro, Christina Patterson, and AyŞegül Şahin. (2018) "Job Search Behavior over the Business Cycle." *American Economic Journal: Macroeconomics*, 10, 190–215.

Ormeño, Arturo. (2015) "Using Survey Data of Inflation Expectations in the Estimation of Learning and Rational Expectations Models." *Journal of Money, Credit and Banking* 47, 673–99.

Orphanides, Athanasios, and John C. Williams. (2007) "Robust Monetary Policy with Imperfect Knowledge." *Journal of Monetary Economics*, 54, 1406–35.

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

Petrosky-Nadeau Nicolas. (2014) "Credit, vacancies and unemployment fluctuations." *Review of Economic Dynamics* 17, 191–205.

Petrosky-Nadeau, Nicolas, and Etienne Wasmer. (2013) "The Cyclical Volatility of Labor Markets under Frictional Financial Markets." *American Economic Journal: Macroeconomics*, 5, 193–221.

Pissarides, Christopher A. (2009) "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" *Econometrica*, 77, 1339–69.

Preston, Bruce. (2005) "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter." *International Journal of Central Banking*, 1, 81–126.

Preston, Bruce. (2006) "Adaptive Learning, Forecast-Based Instrument Rules and Monetary Policy." *Journal of Monetary Economics*, 53, 507–35.

Preston, Bruce. (2008) "Adaptive Learning and the Use of Forecasts in Monetary Policy." *Journal of Economic Dynamics and Control*, 32, 3661–81.

Quadrini, Vincenzo, and Antonella Trigari. (2008) "Public Employment and the Business Cycle." *Scandinavian Journal of Economics*, 109, 723–42.

Reiter, Michael. (2007) "Embodied Technical Change and the Fluctuations of Unemployment and Wages." *The Scandinavian Journal of Economics*, 109, 695–721.

Robin, Jean-Marc. (2011) "On the Dynamics of Unemployment and Wage Distributions." *Econometrica*, 79, 1327–55.

Sargent, Thomas J. (1999) *The Conquest of American Inflation*. Princeton NJ: Princeton University Press.

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

Shimer, Robert. (2010) *Labor Markets and Business Cycles (Crei Lectures in Macroeconomics)*. Princeton, NJ: Princeton University Press.

Slobodyan, Sergey, and Raf Wouters. (2012) "Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models." *American Economic Journal: Macroeconomics*, 4, 65–101.

Solon, Gary, Robert Barsky, and Jonathan A Parker. (1994) "Measuring the Cyclicality of Real Wages: How Important Is Composition Bias?" *The Quarterly Journal of Economics*, 109, 1–25.

Thomas, Carlos. (2006) "Firing Costs, Labour Market Search and the Business Cycle." Mimeo, London School of Economics.

Trigari Antonella. (2006) “The Role of Search Frictions and Bargaining for Inflation Dynamics.” *SSRN Electronic Journal*.

Trigari, Antonella. (2009) “Equilibrium Unemployment, Job Flows, and Inflation Dynamics.” *Journal of Money, Credit and Banking*, 41, 1–33.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.