BANK OF ENGLAND

Staff Working Paper No. 633
Adaptive learning and labour market dynamics

Federico Di Pace, Kaushik Mitra and Shoujian Zhang

December 2016

BANK OF ENGLAND

Staff Working Paper No. 633

Adaptive learning and labour market dynamics

Federico Di Pace,[(1)] Kaushik Mitra[(2)] and Shoujian Zhang[(3)]

Abstract

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 it is unable to replicate: properties of survey forecasts of unemployment in the near term. We present a parsimonious model with adaptive learning and simple autoregressive forecasting rules which 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.

Key words: Adaptive learning, bounded-rationality, search and matching frictions.

JEL classification: E24, E32, J64.

(1) Bank of England. Email: federico.dipace@bankofengland.co.uk
(2) University of Birmingham. Email: k.mitra@bham.ac.uk
(3) Addiko Bank.

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We would like to thank participants of the conference on 'Expectations in dynamic macroeconomic models' at the Bank of Finland (Helsinki), the 1st Birkbeck Centre of Applied Macroeconomics conference at Birkbeck College (London), the 20th Annual Computing in Economics and Finance Conference at the BI Norwegian Business School (Oslo), the 7th Conference of the Centre of Economic Growth and Business Cycles at the University of Manchester, the 2014 Asian Meeting of Econometric Society at the Academia Sinca (Taipei), 46th Annual Conference of the MMF at the University of Durham and seminar participants at the Universities of Exeter, Cardiff, the Technical University of Vienna and the Bank of England. We would especially like to thank William Branch, George Evans, Lien Laureys, Bruce Preston and Joseph Pearlman for useful comments and suggestions.

Information on the Bank's working paper series can be found at
www.bankofengland.co.uk/research/Pages/workingpapers/default.aspx

Publications Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 Fax +44 (0)20 7601 3298 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

The Diamond-Mortensen-Pissarides (DMP) search and matching model has become the standard theory of equilibrium unemployment. 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 labour market variables, such as unemployment, vacancies and the measure of labour market tightness present in the US data. 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. This feature relates to survey forecasts of the unemployment rate in the near term. One of the central variables in the search and matching model is the unemployment rate, which households and firms have to forecast to make consumption and vacancy posting decisions. Figure 1 shows the one and four step-ahead forecast errors in unemployment rates obtained from the Survey of Professional Forecasters (SPF). Note the systematic over/under-prediction of unemployment rates over the business cycle made in the surveys by professional forecasters. Most importantly, the forecast errors are positively correlated with the business cycle. and increase with the forecast horizon. Forecast errors in unemployment rates are much more volatile relative to output over the 4-quarter horizon. For instance, while the one quarter ahead forecast error in the unemployment rate is 4.56 times as volatile as output in the US data, this measure under RE is close to zero.

Since standard models assume rational expectations, they are inconsistent with the survey data because agents do not make systematic errors in these models. The search and matching model with RE is unable to match the properties of the forecast errors. This paper provides a solution to this puzzle by examining the role of expectation formation for hiring decisions. We replace the Rational Expectation assumption by a set of simple autoregressive subjective beliefs that have gained recent popularity in 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* as well as match the statistical properties of forecast errors in unemployment found in US data. Our paper is the first one to highlight the role of expectation formation in

[1] Mortensen and Nagypal (2007) argue that the performance of the standard model featuring RE beliefs depends on the variability of per hire rather than on the asusmption about the cyclicality of wages.

the study of hiring decisions and its capacity to match important features of labor market data.

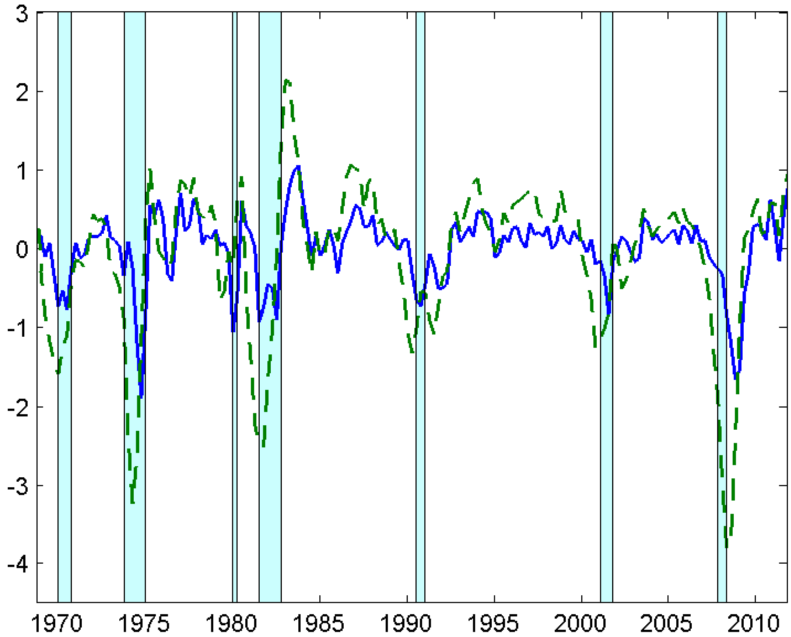

Figure 1: Forecast errors of unemployment from SPF. *Notes. Forecast errors are expresed in percentage points i.e. forecast unemployment rates minus actual unemployment rates for the period 1968Q4-2015Q2. The solid and dotted lines denote the one-step and four-step ahead forecast errors respectively. The figure shows the forecast errors increase as the time horizon of the forecast increases. Bands show the NBER recession dates. The forecast errors are negative during recessions and positive during expansions.*

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 and households form forecasts of unemployment rates, wages, profits and interest rates up to the infinite future to make consumption and hiring decisions using simple autoregressive models to form expectations. Firms in our model face a dynamic problem due to long-lasting employment relations modelled through search frictions. We assume that agents have incomplete knowledge about the structure of the economy in that they do not know the preferences and technologies of other agents in the economy. Hiring decisions at the firm level depend crucially upon the perception of future profits per hire.

We find that not all forms of adaptive learning 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 then the amplification results are not too dissimilar from RE. However, strikingly, assuming AL results in a much better fit to the data than assuming RE when agents use small forecasting models in their learning. In this sense, the results are consistent with Slobodyan and Wouters

(2012).[2] The AL models considered are in a sense minimal departures from the RE solution. As shown the RE solution involves employment, wages, profits and interest rates depending on the (only) state variable employment and the technology shock (assumed to be of autoregressive process of order one i.e. AR(1) process). This RE solution may in turn be written in an alternative way whereby wages, profits and interest rates are auto-regressive moving average processes i.e. ARMA(2,1) processes while employment is an autoregressive process of order two; i.e. an AR(2) process. We consider three variants of small forecasting models motivated by the nature of the RE solution. The first model we consider is a minimal departure from the RE solution in that agents assume the key four endogenous variables (employment, wages, profits and interest rates) are all AR(2) processes. The next model assumes these are all AR(1) processes. The final model we consider is where agents perceive these four variables to evolve as a VAR(1) process: this is a popular method of forecasting and is particularly appealing as it allows for possible interactions among the key endogenous variables since agents under incomplete knowledge are unaware of the general equilibrium restrictions among these variables.

We show that all of these models with adaptive learning generate much more amplification in labour market variables compared to their RE counterpart (to different degrees). For instance, for the AR(2) model, unemployment and vacancies are 7 and 10 times more volatile than the corresponding measure of output, which is in line with US data; these numbers are in fact 16 and 18 times higher than those in the corresponding RE model. Our model matches well the statistical properties of forecast errors on unemployment data taken from the SPF. In particular, the volatility and cyclical behavior of unemployment forecast errors generated by these models match the data closely. This finding is in sharp contrast with the RE model and supports the type of perceived beliefs assumed in the learning model.

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.[3] 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 carry out infinite horizon forecasts to choose how many vacancies to posts, firms' *discounting* of

[2] Note, however, that we use the infinite horizon learning approach advocated by Preston (2005) unlike the Euler equation learning approach of this paper.

[3] 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.

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.[4] The second popular approach by Hagedorn and Manovskii (2008) (henceforth, HM) carries out a simple calibration exercise that sets the value of non-market activity close to the value of search of the worker.[5] These characterisations of the labour market have been criticised on the following grounds: a) microeconometric evidence by Pissarides (2009) and Haefke, Sonntag, and van Rens (2013) are suggestive that wages for newly hired workers are cyclical and b) Costain and Reiter (2008) show that the implied elasticity of unemployment benefits with respect to the unemployment rate arising from the calibration in HM is implausibly large relative to the data.

The main assumption in the paper that economic agents engage in "learning" behavior has been incorporated into macroeconomic 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 adaptive learning approach treats economic agents like econometricians who estimate forecast rules, updating the parameter estimates over time as new data become available. Agents update their forecasting model, their forecasts of future variables and re-solve their dynamic optimization problem in order to make their decisions. In the context of infinite-horizon learning, agents solve dynamic optimization problems. This learning approach can be viewed as a version of the anticipated utility approach formulated by Kreps (1998) used by Eusepi and Preston (2011) and Kuang and Mitra (2016) within the context of the RBC model.[6]

The point that there is an important divergence between the implied expectations in macroeconomic models with RE and the expectations drawn from surveys (eg. SPF) has

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

[5]The list of solutions to the unemployment volatility puzzle is not exhaustive; see Menzio and Shi (2011), Colciago and Rossi (2011), Alves (2012), Quadrini and Trigari (2008), Gomes (2011), Reiter (2007), Guerrieri (2008), Robin (2011), Petrosky-Nadeau (2013) and Petrosky-Nadeau and Wasmer (2013).

[6]In the anticipated utility approach recommended by Kreps, agents update their forecasts over time but do not take into account the fact that their forecasting model will be revised in future periods. This is a bounded rationality approach, since a full Bayesian approach would recognize the uncertainty in the parameters of the estimated forecasting model. However, as noted by Cogley and Sargent (2008), a full Bayesian approach in macroeconomic settings is typically "too complicated to be implemented," and thus the anticipated utility approach is an appealing implementation of bounded rationality. An alternative approach by Adam and Marcet (2011) and Adam, Marcet, and Beutel (2016) derives jointly the optimal decisions and belief updating rules from the utility maximisation problem in the context of an asset pricing model.

been made in the context of other models; see e.g. Adam, Marcet, and Beutel (2016), Kuang and Mitra (2016), Slobodyan and Wouters (2012), Milani (2011) and Ormeno 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.

The remainder of the paper is organised as follows. Section 2 describes the model environment. Section 3 describes the RE solution of the model together with the learning models considered in the paper. Section 4 presents the results of the model for the baseline calibration. Section 5 evaluates the quantitative performance of the model in terms of the forecast errors of unemployment and highlights the similarities and differences between our approach and wage rigidities under RE beliefs. Section 6 analyses the robustness of the results to alternative parameterisations and values of the gain parameter. Section 7 concludes.

2 Model

We propose a simple model featuring labour market search and matching frictions as in Mortensen and Pissarides (1994) and a form of adaptive learning following Preston (2005) and Mitra, Evans, and Honkapohja (2013). Our model economy is inhabited by two types of agents: households and firms. There is a representative household consisting of a continuum of workers that search for jobs if unemployed and work for firms if employed. Following Andolfatto (1996), we make the assumption of perfect risk sharing at the household level so that both employed and unemployed members of the household consume equal amounts. Firms post job vacancies and employ workers with a lag so as to produce final goods using labour as the only input of production. Households consume the final goods supplied by firms. We assume that agents form their expectations by updating their beliefs as new information becomes available. Agents make infinite horizon (IH) forecasts about the future path of wages, interest rates, unemployment and profits by running simple autoregressive models in order to make current decisions about consumption and vacancy posting.

2.1 Labour Market

The labour 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 continuously 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 a matching function that exhibits constant returns to scale

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

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

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

We define the measure of labour market tightness as

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

Due to the assumption of constant returns in the matching technology, we define the job finding rate as $m(v_t, u_t)/u_t = m(v_t/u_t, 1) = p(\theta_t)$ and the job filling rate as

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

Note that the job finding rate, $p(\theta_t)$, is increasing in the measure of labour market tightness and the job filling rate, $q(\theta_t)$, decreasing.

2.2 Households

In our model economy there is a large representative household consisting of a continuum of members in the unit interval that maximises its life-time utility

$$\mathcal{H}(s_t, n_t) = E_t^* \sum_{s=t}^{\infty} \beta^{s-t} \mathcal{U}(c_s, n_s), \tag{5}$$

where E_t^* denotes the subjective expectation of the household at time t. $\mathcal{H}(s_t, n_t)$ denotes the value function of the household, which depends on last period savings and the period employment rate. The period utility of the representative household depends positively on consumption, c_t, and negatively on employment, n_t, and is defined by

$$\mathcal{U}(c_t, n_t) = \log c_t - \chi n_t,$$

where $\chi > 0$ is a parameter that captures the disutility of employment at the level of the household. The flow budget constraint is given by

$$\frac{s_{t+1}}{r_t} = w_t n_t + \pi_t + s_t - c_t, \tag{6}$$

where s_t denotes household savings at the end of period t, r_t the real interest rate and π_t aggregate profits, which are rebated to the household at the end of each period. There is a continuum of firms indexed by $f \in [0,1]$; the aggregate employment rate is the sum of employment across firms, $n_t = \int_0^1 n_{ft} df$, while $w_t n_t = \int_0^1 w_{ft} n_{ft} df$ and $\pi_t = \int_0^1 \pi_{ft} df$ are the pooled wage bill and profits respectively. In addition, the employment rate at the household level evolves according to

$$n_{t+1} = (1-\rho) n_t + u_t p(\theta_t). \tag{7}$$

The representative household chooses the level of consumption to maximise its life-time utility, (5), subject to the flow budget constraint, (6), and law of motion of employment, (7). The household's problem can be written in terms of the following Bellman equation

$$\mathcal{H}(s_t, n_t) = \max_{c_t, s_{t+1}} \left\{ \log c_t - \chi n_t + \beta E_t^* \mathcal{H}(s_{t+1}, n_{t+1}) \right\},$$

together with the constraints (6) and (7). By combining the first order conditions with respect to c_t and s_t, we obtain the standard Euler equation

$$\frac{1}{r_t} = \beta E_t^* \frac{c_t}{c_{t+1}}. \tag{8}$$

This condition states that the household's marginal utility derived from consumption at time t must equal the marginal utility of consumption derived at time $t+1$ expressed in terms of time t.

The envelope condition with respect to n_{ft} gives

$$\frac{\partial \mathcal{H}(s_t, n_t)}{\partial n_{ft}} = \frac{w_{ft}}{c_t} - \chi + \beta \left[1 - \rho - p(\theta_t)\right] E_t^* \frac{\partial \mathcal{H}(s_{t+1}, n_{t+1})}{\partial n_{ft+1}}. \tag{9}$$

This condition states that the net marginal value of having a member of the household employed at firm f is equal to the net flow value of employment, the difference between the wage, expressed in terms of utility, and the utility cost of working, plus the net continuation value of employment.

By iterating forward the budget constraint, (6), and substituting for future values of savings, we can find an expression for consumption

$$c_t + E_t^* \sum_{j=1}^{\infty} D_{t,t+j}^{-1} c_{t+j} = s_t + w_t n_t + \pi_t + E_t^* \sum_{j=1}^{\infty} D_{t,t+j}^{-1} (w_{t+j} n_{t+j} + \pi_{t+j}), \tag{10}$$

where $D_{t,t+j} = \prod_{i=0}^{j-1} r_{t+i}$, $j \geq 1$. The household's perceived transversality condition,

$$\lim_{j\to\infty} E_t^* D_{t,t+j}^{-1} s_{t+j} = 0,$$

holds. This expression states that the present discounted value of consumption is equal to the sum of perceived human and non-human wealth.[7]

2.3 Firms

Our model economy features a continuum of large firms of measure $f \in [0,1]$. Each firm f employs labour to produce consumption goods using the following technology

$$y_{ft} = z_t n_{ft}^{\alpha}, \tag{11}$$

where y_{ft} denotes output at the firm level, n_{ft} employment and $\alpha \in (0,1]$ the aggregate elasticity of output with respect to labour. In the baseline calibration, we assume constant returns to scale, i.e. α equal to 1, but later on we also examine the case of decreasing returns to scale in Section 6. The productivity innovation z_t follows an exogenous process given by

$$\ln z_{t+1} = \varrho \ln z_t + \epsilon_{t+1} \quad \text{with} \quad \epsilon_t \sim N(0, \varsigma), \tag{12}$$

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.

Since posting vacancies is costly, period profits, π_{ft}, at time t may be written as

$$\pi_{ft} = z_t n_{ft}^{\alpha} - w_{ft} n_{ft} - \mathcal{C}(v_{ft}), \tag{13}$$

where v_{ft} the number of job openings at the firm level and $\mathcal{C}(\cdot)$ is a convex and increasing vacancy cost function. The problem of each firm is to choose v_{ft} so as to maximise the present discounted value of expected profits, which may be written as

$$\max_{v_{ft+j}} \quad \pi_{ft} + \sum_{j=1}^{\infty} E_{ft}^* D_{t,t+j}^{-1} \pi_{ft+j} \quad \text{for} \quad j \geq 0, \tag{14}$$

subject to the law of motion of employment at the firm level

$$n_{ft+1} = (1-\rho) n_{ft} + v_{ft} q_t. \tag{15}$$

The Bellman equation of this problem may be written as

$$\mathcal{J}(n_{ft}) = \max_{v_{ft}} \left\{ \pi_{ft} + E_{ft}^* r_t^{-1} \mathcal{J}(n_{ft+1}) \right\},$$

[7] Note that profits are large under decreasing return but small under constant returns.

so that the problem of firm f is to maximise the above equation subject to (11) and (15). The first order condition with respect to v_{ft} is

$$\mathcal{C}'\left(v_{ft}\right) = q\left(\theta_t\right) E^*_{ft} r_t^{-1} \mathcal{V}_{ft+1}, \tag{16}$$

where $\mathcal{V}_{ft} = \mathcal{J}'(n_{ft})$ denotes the marginal value of having an additional worker employed at the firm. Equation (16) states that the marginal cost and benefit of posting a vacancy must be equal. The envelope condition with respect to n_{ft} is

$$\mathcal{V}_{ft} = \alpha z_t n_{ft}^{\alpha-1} - w_{ft} + (1-\rho) E^*_{ft} r_t^{-1} \mathcal{V}_{ft+1}. \tag{17}$$

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

By leading equation (17) one period forward, multiplying both sides of the expression by the stochastic discount factor, $\frac{1}{r_t}$, taking the expectation at time t and combining the resulting expression with equation (16), we obtain the following job creation condition

$$\frac{\mathcal{C}'\left(v_{ft}\right)}{q\left(\theta_t\right)} = E^*_{ft} r_t^{-1} \left[\alpha z_{t+1} n_{ft+1}^{\alpha-1} - w_{ft+1} + (1-\rho) \frac{\mathcal{C}'\left(v_{ft+1}\right)}{q\left(\theta_{t+1}\right)}\right]. \tag{18}$$

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 labour market conditions. The expected marginal cost of opening a vacancy can be re-written in terms of the perceived sum of future profits per additional hire,

$$\frac{\mathcal{C}'\left(v_{ft}\right)}{q\left(\theta_t\right)} = \sum_{j=1}^{\infty} (1-\rho)^{j-1} E^*_{ft} D^{-1}_{t,t+j} \left[\alpha z_{t+j} n_{ft+j}^{\alpha-1} - w_{ft+j}\right]. \tag{19}$$

Note that this expression states that the value of an additional job must be equal to the sum the future stream of profits that this job is expected to generate. In order to post the optimal number of vacancies, firm f must make forecasts up to the infinite future. The firm perceives the following transversality condition holds:

$$\lim_{j\to\infty} (1-\rho)^{j-1} E^*_{ft} \left[D^{-1}_{t,t+j} \frac{\mathcal{C}'\left(v_{ft}\right)}{q\left(\theta_{t+j}\right)}\right] = 0.$$

2.4 Wage Negotiation

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

$$\arg\max_{w_{ft}} \left[\frac{\partial \mathcal{H}(s_t, n_t)}{\partial n_{ft}} c_t\right]^{\xi} (\mathcal{V}_{ft})^{1-\xi},$$

where $\xi \in (0,1)$ denotes the workers' bargaining power or, alternatively, the share of surplus taken by the worker.[8] Note that the household's surplus is the product of the marginal value of having a member of the household employed at firm f, defined in equation (9), and c_t (the inverse of the marginal utility of consumption). Dividing $\frac{\partial \mathcal{H}(s_t,n_t)}{\partial n_{ft}}$ by the marginal utility of consumption translates utility units of $\mathcal{H}$ in terms of goods. The first order condition of this problem then yields the standard sharing rule that characterises the optimal split of the aggregate surplus between workers and firms,

$$(1-\xi)\left[\frac{\partial \mathcal{H}(s_t, n_t)}{\partial n_{ft}} c_t\right] = \xi \mathcal{V}_{ft}. \tag{20}$$

To derive an expression for the bargained wage, w_{ft}, we assume that expression (20) holds for subjective expectations (see equation (49) in Appendix A.1 and the details therein).[9] Furthermore, we assume that $\mathcal{C}(v_{ft})$ takes a linear form, $\mathcal{C}(v_{ft}) = \kappa v_{ft}$, as is standard in the literature, where κ denotes the unitary vacancy cost. Thus, we obtain

$$w_{ft} = \xi \alpha z_t n_{ft}^{\alpha-1} + \xi \kappa \theta_t + (1-\xi)\chi c_t. \tag{21}$$

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.[10]

2.5 Aggregation and Market Clearing

We make the assumption that households and firms share the same set of beliefs about the future. This assumption is reasonable because a) firms are owned by their employees, b) household members and firms coordinate on expectations during the wage negotiation process. This assumption of homogeneity in expectations across households and firms, $E_t^* = E_{ft}^*$, implies a symmetric equilibrium i.e., $n_{ft} = n_t$ and $v_{ft} = v_t$ for all f and t.

[8] The match values of the employment is specific to each member of the household and each employee of the firm.

[9] In line with what it is assumed under RE, this assumption makes it easy for the worker and the firm to agree on the bargained wage. This is a convenient simplifying assumption for our purpose. There may be alternative (more complex) ways of agreeing the bargained wage which we leave for future work.

[10] As shown by Krause and Lubik (2007), the aggregate effects of intra-firm bargaining are negligible in a standard search and matching framework with concave production functions. Thus, we abstract from intra-firm bargaining in the analysis.

The market clearing condition in the goods market can be obtained by summing up the period budget constraints and period profits (over f)

$$c_t + \kappa v_t = y_t. \tag{22}$$

Imposing symmetry and following Mitra, Evans, and Honkapohja (2013), we assume households use a consumption rule based on a linearisation of (10) around the steady state values $(\bar{c}, \bar{w}, \bar{n}, \bar{\pi}, \bar{r})$. 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). The linearised household's behavior rule is given by

$$\begin{aligned}\frac{\tilde{c}_t}{1-\beta} &= \tilde{s}_t + \bar{n}\tilde{w}_t + \bar{w}\tilde{n}_t + \beta\bar{w}\tilde{n}_{t+1} + \tilde{\pi}_t - \frac{(\bar{w}\bar{n}+\bar{\pi})\,\beta^2}{1-\beta}\tilde{r}_t + \\ & \sum_{j=2}^{\infty}\beta^j\bar{w}\tilde{n}_{t+j} + \sum_{j=1}^{\infty}\beta^j E_t^*\left[\bar{n}\tilde{w}_{t+j} + \tilde{\pi}_{t+j} - (\bar{w}\bar{n}+\bar{\pi})\,\beta\sum_{i=1}^{j-1}\tilde{r}_{t+i}\right].\end{aligned} \tag{23}$$

We make the assumption that initial financial wealth is zero (and by symmetry it is zero in all future periods). It is then convenient to re-write (23) as

$$\frac{\tilde{c}_t}{1-\beta} = \bar{n}\tilde{w}_t + \bar{w}\tilde{n}_t + \beta\bar{w}\tilde{n}_{t+1} + \tilde{\pi}_t - \frac{(\bar{w}\bar{n}+\bar{\pi})\,\beta^2}{1-\beta}\tilde{r}_t + \mathcal{S}_{ht}^w + \mathcal{S}_{ht}^n + \mathcal{S}_{ht}^\pi - \mathcal{S}_{ht}^r, \tag{24}$$

where

$$\begin{aligned}\mathcal{S}_{ht}^w &= \sum_{j=1}^{\infty}\beta^j\bar{n}E_t^*\tilde{w}_{t+j}, \quad \mathcal{S}_{ht}^n = \sum_{j=2}^{\infty}\beta^j\bar{w}E_t^*\tilde{n}_{t+j}, \\ \mathcal{S}_{ht}^\pi &= \sum_{j=1}^{\infty}\beta^j E_t^*\tilde{\pi}_{t+j} \quad \text{and} \quad \mathcal{S}_{ht}^r = (\bar{w}\bar{n}+\bar{\pi})\sum_{j=1}^{\infty}\left[\beta^j E_t^*\sum_{i=1}^{j-1}\tilde{r}_{t+i}\right].\end{aligned}$$

The $\mathcal{S}_{ht}^w, \mathcal{S}_{ht}^n$, etc variables denote the discounted sums of future forecasts and are key for understanding the dynamic properties of the model that we explain at a later stage.

In line with the household problem, firms use a vacancy posting rule based on linearisation of equation (19) around the steady state values of $\bar{w}$, $\bar{n}$, $\bar{q}$, $\bar{r}$ and $\bar{v}$. The firms' behavioral rule is therefore

$$\begin{aligned}&\bar{\lambda}_1\bar{q}\tilde{v}_t - \left[\frac{\kappa}{\bar{q}^2} + \beta\bar{\lambda}_1\bar{v}\right]\tilde{q}_t = \beta\bar{\lambda}_1(1-\rho)\,\tilde{n}_t - \frac{(\bar{z}\bar{\lambda}_2-\bar{w})\beta^2}{1-\beta(1-\rho)}\tilde{r}_t + \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^j E_t^*\bar{\lambda}_1\tilde{n}_{t+j} + \\ &\sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j E_{f,t}^*\left[\bar{\lambda}_2\tilde{z}_{t+j} - \tilde{w}_{t+j} - \left(\bar{z}\bar{\lambda}_2 - \bar{w}\right)\beta\sum_{i=1}^{j-1}\tilde{r}_{t+i}\right],\end{aligned} \tag{25}$$

where $\bar{\lambda}_1 = \alpha(\alpha-1)\bar{z}\bar{n}^{\alpha-2}$ and $\bar{\lambda}_2 = \alpha\bar{n}^{\alpha-1}$.[11] We re-write (25) as

$$\bar{\lambda}_1\bar{q}\tilde{v}_t - \left(\frac{\kappa}{\bar{q}^2} + \beta\bar{\lambda}_1\bar{v}\right)\tilde{q}_t = \beta\bar{\lambda}_1(1-\rho)\tilde{n}_t - \frac{(\bar{z}\bar{\lambda}_2 - \bar{w})\beta^2}{1-\beta(1-\rho)}\tilde{r}_t + \mathcal{S}^z_{ft} + \mathcal{S}^n_{ft} - \mathcal{S}^w_{ft} - \mathcal{S}^r_{ft}, \quad (26)$$

where

$$\begin{aligned}
\mathcal{S}^z_{ft} &= \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j\bar{\lambda}_2 E^*_t\tilde{z}_{t+j}, \quad \mathcal{S}^n_{ft} = \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^j\bar{\lambda}_1 E^*_t\tilde{n}_{t+j}, \\
\mathcal{S}^w_{ft} &= \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j E^*_t\tilde{w}_{t+j} \quad \text{and } \mathcal{S}^r_{ft} = (\bar{z}\bar{\lambda}_2 - \bar{w})\sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j+1}\sum_{i=1}^{j-1}E^*_t\tilde{r}_{t+i}.
\end{aligned}$$

The informational assumptions used in the derivation of equation (26) are detailed in Appendix A.2. In particular, households and firms are assumed not to observe $\tilde{w}_t, \tilde{n}_{t+1}, \tilde{\pi}_t$ and $\tilde{r}_t$ when they form their forecasts at time t. The standard assumption in the learning literature is to assume that agents' forecasts at t from period $t+1$ onwards are based on past endogenous variables (here n_t, w_{t-1}, r_{t-1}, and π_{t-1}). This approach conveniently avoids the simultaneous determination of forecasts and endogenous variables.[12] This informational assumption is particularly appealing here since the standard search literature is calibrated to monthly frequency and the data is generally available with a lag; viewing households and firms as having knowledge of contemporaneous monthly aggregate variables when forming their forecasts is therefore implausible. Wage determination in the search and matching literature is different from the standard neoclassical model and it entails a more complex decision problem involving expectations. Under RE, households and firms know that their expectations are the same. 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 $\tilde{c}_t$; see equations (21) or (27) below. In addition, the assumption of time-to-hire is such that workers are matched to employeers in period t but can only start working in period $t+1$. For all these reasons, the assumption of pre-determined forecasts in this model is particularly plausible and we maintain this assumption in the paper. An interpretation, owing to Evans and Honkapohja (2006), is that these forecasts are obtained by households and firms from an econometric forecasting firm before going to the market place. Armed with these forecasts, equations (24) and (26) provide a consumption schedule for the representative household and a vacancy posting rule for firms which they take to the market place. These (along with other equations outlined below) determines a temporary equilibrium for the economy where contemporaneous monthly

[11] Note that $\bar{\lambda}_1$ is 0 and $\bar{\lambda}_2$ is 1 when the production function exhibits constant returns to scale. For brevity, we denote by $\widehat{q}_t$ the absolute deviation of $q(\theta_t)$.

[12] The alternative approach would be to assume that current dated (here monthly) endogenous variables are included in the forecasts of agents so that these endogenous variables and their forecasts are determined simultaneously. As explained this approach is less realistic given the monthly frequency of the model.

values of endogenous variables eg. $\tilde{c}_t, \tilde{v}_t, \tilde{q}_t, \tilde{w}_t, \tilde{\pi}_t$ and $\tilde{r}_t$ and $\tilde{n}_{t+1}$ are all simultaneously determined.

We turn to the other equations in the model. First, we linearise equation (21) around the steady state and integrate it over f to find an expression for the aggregate wage

$$\tilde{w}_t = \xi\bar{\lambda}_2\tilde{z}_t + \xi\bar{\lambda}_1\tilde{n}_t + \xi\kappa\tilde{\theta}_t + (1-\xi)\,\chi\tilde{c}_t. \tag{27}$$

Similarly, we linearise the unemployment rate (2), the production function (11), the profit function (13), the law of motion of employment (15) and the goods market clearing condition (22) to get the following conditions

$$\tilde{n}_t + \tilde{u}_t = 0, \tag{28}$$

$$\tilde{y}_t = \bar{n}^{\alpha}\tilde{z}_t + \bar{z}\alpha\bar{n}^{\alpha-1}\tilde{n}_t, \tag{29}$$

$$\tilde{\pi}_t = \bar{z}\alpha\bar{n}^{\alpha-1}\tilde{n}_t + \bar{n}^{\alpha}\tilde{z}_t - \bar{n}\tilde{w}_t - \bar{w}\tilde{n}_t - \kappa\tilde{v}_t, \tag{30}$$

$$\tilde{n}_{t+1} = (1-\rho)\,\tilde{n}_t + \bar{v}\tilde{q}_t + \bar{q}\tilde{v}_t, \tag{31}$$

$$\tilde{c}_t + \kappa\tilde{v}_t = \tilde{y}_t. \tag{32}$$

We also linearise equations (3), (4) and (12) to get

$$\bar{\theta}\tilde{u}_t + \bar{u}\tilde{\theta}_t = \tilde{v}_t, \tag{33}$$

$$\tilde{q}_t = (\sigma-1)\,m\bar{\theta}^{\sigma-2}\tilde{\theta}_t, \tag{34}$$

$$\tilde{z}_{t+1} = \varrho\tilde{z}_t + \epsilon_{t+1}. \tag{35}$$

We now define a temporary equilibrium for this economy. A temporary equilibrium is a set of values for the variables $\tilde{c}_t$, $\tilde{y}_t$, $\tilde{v}_t$, $\tilde{w}_t$, $\tilde{r}_t$, $\tilde{\pi}_t$, $\tilde{n}_{t+1}$, $\tilde{u}_t$, $\tilde{\theta}_t$, $\tilde{q}_t$ that, given the exogenous stochastic process $\{\tilde{z}_j\}_{j=t}^{\infty}$ and the initial condition $\tilde{n}_t$, satisfy the system of equations consisting of (24) and (26)-(35). This equilibrium is determined as follows. In every period t, given their forecasts, households enter the market with their consumption schedule (24). Similarly, given their forecasts, firms enter the market with their vacancy posting rule (26). These equations, together with the other equations outlined earlier (which do not depend on forecasts), determine simultaneously the temporary equilibrium values of $\tilde{c}_t$, $\tilde{y}_t$, $\tilde{v}_t$, $\tilde{w}_t$, $\tilde{r}_t$, $\tilde{\pi}_t$, $\tilde{n}_{t+1}$, $\tilde{u}_t$, $\tilde{\theta}_t$, $\tilde{q}_t$ at time t. To complete the description of temporary equilibrium one specifies how forecasts are formed. As argued before, in this model it is plausible to assume that forecasts of firms and households are pre-determined when they are brought to the market. The temporary equilibrium for the current period provides a new data point for the agents. Given these new data, the forecasts of agents are updated at the start of the following period using versions of recursive least squares algorithm (which we explain later).

3 RE and Learning with Correctly Specified Beliefs

The RE solution of the model is of the following form

$$
\begin{aligned}
n_{t+1} &= \bar{b}_n + \bar{a}_{nn} n_t + \bar{a}_{nz} \tilde{z}_t, \\
w_t &= \bar{b}_w + \bar{a}_{wn} n_t + \bar{a}_{wz} \tilde{z}_t, \\
r_t &= \bar{b}_r + \bar{a}_{rn} n_t + \bar{a}_{rz} \tilde{z}_t, \\
\pi_t &= \bar{b}_\pi + \bar{a}_{\pi n} n_t + \bar{a}_{\pi z} \tilde{z}_t,
\end{aligned}
\tag{36}
$$

since 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 and also the dispersion of that innovation. Rational agents have complete information about the structure of the economy and know with certainty the true parameter values of policy functions (denoted with a bar over the parameter). The solution of the RE model is then used to forecast future values of the endogeneous variables. Therefore, rational agents make no systematic mistakes. As is well known, the RE model is not able to match the amplification present in the labour market data since productivity shocks are unable to increase by much profits per hire at the margin. More detailed explanations are provided in Section 4.4.1.

Under adaptive learning (AL), agents are assumed to have perceived laws of motion (PLMs) of the same form as equation (36) but do not have knowledge of the RE parameters ($\bar{b}_x$, $\bar{a}_{xn}$ and $\bar{a}_{xz}$), learning them as new information becomes available. In particular, the correctly specified beliefs of agents consist of the following set of equations

$$
x_t = b_x + a_{xn} n_t + a_{xz} \tilde{z}_t + \eta_{xt}, \quad \text{for} \quad x_t = \{n_{t+1}, \pi_t, w_t, r_t\}, \tag{37}
$$

where η_{xt} are white noise processes. Agents are assumed to behave as econometricians and estimate parameters b_x, a_{xn} and a_{xz} by running regressions of variables like employment, profits, wages and interest rates on past employment and TFP innovations. This belief system entails only a mild deviation from RE. Tables 8 and 9 in Appendix B show that this belief system suffers from the same problem as the RE model i.e. it is not able to provide a solution to the unemployment volatility puzzle and unable to match the forecast error properties in the data.[13] This seems to be the case since the correctly specified beliefs converge to the RE solution. We, therefore, turn to a situation where

[13] The computations in these tables use the same definitions used in Tables 3, 4 and 5; please refer to sections 4 and 5.

agents use smaller forecasting models and consider if this can provide a resolution to these problems.

Before doing so we first note that, following Campbell (1994), the RE solution (36) may be written in an equivalent way. This RE solution involves n_t as an AR(2) process and π_t, w_t, r_t as ARMA (2,1) processes. In particular, this solution can be shown to be of the following form

$$\begin{aligned} n_{t+1} &= \vartheta_1 \bar{b}_n + \vartheta_2 n_t - \vartheta_3 n_{t-1} + \bar{a}_{nz}\epsilon_t, \qquad (38) \\ w_t &= \vartheta_1 \bar{b}_w + \vartheta_2 w_{t-1} - \vartheta_3 w_{t-2} + \bar{a}_{wz}\epsilon_t + \left(\bar{a}_{wn}\bar{a}_{nz} - \bar{a}_{wz}\bar{a}_{nn}\right)\epsilon_{t-1}, \\ r_t &= \vartheta_1 \bar{b}_r + \vartheta_2 r_{t-1} - \vartheta_3 r_{t-2} + \bar{a}_{rz}\epsilon_t + \left(\bar{a}_{rn}\bar{a}_{nz} - \bar{a}_{rz}\bar{a}_{nn}\right)\epsilon_{t-1}, \\ \pi_t &= \vartheta_1 \bar{b}_\pi + \vartheta_2 \pi_{t-1} - \vartheta_3 \pi_{t-2} + \bar{a}_{\pi z}\epsilon_t + \left(\bar{a}_{\pi n}\bar{a}_{nz} - \bar{a}_{\pi z}\bar{a}_{nn}\right)\epsilon_{t-1}, \end{aligned}$$

where $\vartheta_1 = (1-\bar{\varrho})(1-\bar{a}_{nn})$, $\vartheta_2 = (\bar{\varrho} + \bar{a}_{nn})$ and $\vartheta_3 = \bar{\varrho}\bar{a}_{nn}$.

Using correctly specified beliefs of the form (38) (or (36)) 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. Recent experimental evidence by Hommes, Sonnemans, Tuinstra, and van de Velden (2005) and Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) suggests agents estimate simple univariate autoregressive models to make forecasts about future variables. Slobodyan and Wouters (2012) too show that univariate autoregressive forecasting models can fit the data better than the RE model and are also closely related to the survey evidence. Using AR type PLMs are hence very appealing.[14] For generality, we also allow beliefs where agents estimate vector-autoregressive (VAR) models. This set of beliefs is appealing since it allows interactions among the key endogenous variables n_t, π_t, w_t, r_t whilst being easy to estimate; it is also a popular forecasting tool in many policy institutions. We deviate from RE in assuming that agents forecast the values of the variables of interest based on simple AR or VAR belief specifications and examine whether these belief specifications have the ability to get the search and matching model closer 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' preferences and beliefs. Thus, they do not know that their decisions are identical to those of other agents.

[14] Using AR (or VAR) beliefs obviates the need for agents to use productivity in their regression equations (37). 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.

3.1 Learning with Autoregressive Beliefs

In this model we propose alternative beliefs systems of autoregressive form.[15] 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 two i.e. AR(2) processes. This serves as our benchmark for agents' beliefs. In particular, we assume that agents have Perceived Law of Motions (PLMs) of the form

$$x_t = a_{x0} + a_{x1}x_{t-1} + a_{x2}x_{t-2} + \mu_{xt} \quad \text{for} \quad x_t = \{n_{t+1}, \pi_t, w_t, r_t\},$$

where μ_{xt} are white noise processes. Like RE, the belief system is expressed in levels. This belief specification represents only a modest departure from RE: the key difference with the RE solution is that the moving average (MA) terms are dropped from r_t, w_t, and π_t (the employment equation is correctly specified).[16] This seems like a reasonable assumption since wages, interest rates and profits are determined simultaneously by general equilibrium considerations and depend on aggregate variables like $\tilde{c}_t$: eg. firms do not know the equilibrium mapping from these endogenous variables to market outcomes.

We now explain parameter updating under this set of beliefs. Let $\Phi_{x,t} = \begin{pmatrix} a_{x0} & a_{x1} & a_{x2} \end{pmatrix}'$ for $x_t = \{n_{t+1}, \pi_t, w_t, r_t\}$ and $\Psi_{x,t} = \begin{pmatrix} 1 & x_{t-1} & x_{t-2} \end{pmatrix}'$. We assume that agents use a constant gain Recursive Least Squares (RLS) algorithm to update their beliefs

$$\begin{aligned} \Phi_{x,t} &= \Phi_{x,t-1} + \gamma R_{x,t}^{-1}\Psi_{x,t-1}\left(x_{t-1} - \Phi'_{x,t-1}\Psi_{x,t-1}\right)', \\ R_{x,t} &= R_{x,t-1} + \gamma\left(\Psi_{x,t-1}\Psi'_{x,t-1} - R_{x,t-1}\right), \end{aligned} \tag{39}$$

where $\gamma \in (0, 1)$ denotes the constant gain learning parameter and $R_{x,t}$ is the precision (3 x 3) matrix associated with each equation. Constant-gain least squares is widely used in the adaptive learning literature because it weighs recent data more heavily. See for example Sargent (1999), Cho, Williams, and Sargent (2002), McGough (2006), Orphanides and Williams (2007), Ellison and Yates (2007), Huang, Liu, and Zha (2009), Eusepi and Preston (2011) and Milani (2011). Evans, Honkapohja, and Williams (2010) provide a Bayesian justification for the constant gain RLS algorithm. In particular, they show that constant gain RLS learning algorithm asymptotically approximates the Bayesian optimal estimator when agents allow for drifting coefficients models. In addition, we assume that

[15] 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.

[16] 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 (2004) 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 survey forecasts, we do not study the nature of this RPE.

agents estimate the persistence parameter ϱ of the exogenous productivity process (12) using constant gain RLS; this captures their uncertainty about the persistence of the productivity process perhaps changing over time.[17]

Agents update their beliefs over time by revising the value of parameters using a Recursive Least Squares (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 consumption and vacancy posting decisions at every point in time. At the end of each period, agents are informed about factor prices, (un)employment and profits 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 $t-1$; eg. n_t, w_{t-1}, r_{t-1}, and π_{t-1} which is what we do here.

We investigate an alternative set of beliefs that entails further deviations from RE. A natural choice as a belief system is the univariate auto-regressive processes of order one, i.e. AR(1) beliefs. In this case the Perceived Law of Motions (PLMs) are of the form

$$x_t = a_{x0} + a_{x1}x_{t-1} + \mu_{xt}, \quad \text{for} \quad x_t = \{n_{t+1}, \pi_t, w_t, r_t\},$$

where μ_{xt} are white noise processes. Agents use RLS algorithms in their estimation and form their forecasts based on the AR(1) model using the informational assumptions stated before.

A final set of beliefs we consider is one that allows for interactions between endogenous variables. A typical way to allow for such interactions is to estimate vector auto-regressive (VAR) models. Adopting multivariate autoregressive models to compute forecasts tends to reduce the informational constraints that are present in simple univariate models. As a result, agents can potentially capture (unknown) general equilibrium effects and reduce forecast errors. In one sense, the AR specifications endow agents with more specific knowledge about the form of the RE solution in that these PLMs do not allow for direct dependence of any variable on the lags of other variables. The VAR specification, on the other hand, allows agents to be relatively agnostic of the RE beliefs. Unlike the univariate AR(2) and AR(1) belief specifications, the VAR(1) forecasting model allows possible interactions among variables (employment, profits, wages and interest rates). The VAR(1) nests the univariate AR(1) model but not necessarily the AR(2) specification. Since VARs have become standard forecasting tools in many macroeconomic institutions, whilst being simple to estimate, we view this as as a plausible and attractive way for agents

[17] In the AL literature, typically it is assumed that ϱ is known since it can be estimated consistently from the exognenous process by RLS. Here we prefer to be realistic about the uncertainty firms may face.

to form their forecasts in the learning set-up. For VAR(1), the PLM is of the form

$$X_t = A_0 + A_1 X_{t-1} + \nu_t \quad \text{for} \quad X_t = (n_{t+1}\ \pi_t\ w_t\ r_t)'$$

where A_0 is a 4x1 matrix, A_1 a 4x4 matrix and ν_t a vector of white noise processes. Similarly, constant gain RLS algorithms are used to estimate the elements of matrices A_0 and A_1 in order for agents to form their forecasts. For economy of space, we do not present the learning algorithms here.

4 Numerical Results

4.1 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 remaining parameters ensures that the stationary equilibrium of the model matches a number of stylised facts as observed in the post-WWII US 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 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 (the choice of the values of σ and ξ ensures that Hosios (1990) condition is met.). The elasticity of output with respect to employment, α, is set to 1. Following Shimer (2010), we choose the persistence of the technology shock, ϱ, to be 0.98 and the standard deviation of the productivity shocks to be 0.005.

The remaining three labour market parameters, namely κ, $\bar{m}$ and χ, are set to match:

i) a vacancy cost to output ratio of 0.01 in line with Andolfatto (1996), Gertler and Trigari (2009) and Blanchard and Gali (2010);

ii) 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);

Description	Parameter	Value
Discount Factor	β	0.996
Parameter in the Utility Function	χ	0.882
Efficiency of the Matching Technology	$\bar{m}$	0.379
Elasticity of the Matching Function	σ	0.5
Bargaining Power	ξ	0.5
Unitary Vacancy Posting Cost	κ	0.084
Separation Rate	ρ	0.033
Elasticity of Output w.r.t. Employment	α	1
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

Table 1: Calibrated Parameters - Monthly

iii) an unemployment rate of 6%, which corresponds with the standard ILO definition of unemployment for the post-WWII US average.

The resulting replacement ratio – computed as the disutility of work over the wage and the marginal utility of consumption ($\chi c/w$) – is equal to 83%, which is above the value of 70% suggested by Mortensen and Nagypal (2007). According to a study by Hagedorn and Manovskii (2008), a total replacement ratio of around 95% can generate labour 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 labour market variables. The replacement ratio in their study is equal to the value of non-market activity that includes both unemployment subsidies and home production. Costain and Reiter (2008) have criticised this calibration because it gives rise to excessive sensitivity of unemployment to unemployment benefits. The resulting replacement ratio ensures that our results are not driven by the Hagedorn and Manovskii effect. This will become clear once we examine the amplification properties of the model under RE beliefs. Table 1 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). 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 (for example as in Mitra, Evans, and Honkapohja (2016) or Mitra, Evans, and Honkapohja (2013) where a higher gain parameter is used) but lies within the range

		$\hat{y}_t$	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$	$\hat{c}_t$
$\sigma_{\widehat{x}_{1t}}/\sigma_{\hat{y}_t}$		1	0.57	9.40	8.84	17.82	0.79
$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t-1}\right)$		0.81	0.90	0.91	0.89	0.91	0.80
$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.79	0.83	-0.78	0.83	0.85
	$\hat{n}_t$	–	1	0.90	-0.97	0.95	0.63
	$\hat{v}_t$	–	–	1	-0.91	0.98	0.69
	$\hat{u}_t$	–	–	–	1	-0.98	-0.61
	$\hat{\theta}_t$	–	–	–	–	1	0.66
	$\hat{c}_t$	–	–	–	–	–	1

Table 2: Summary Statistics, Quarterly US Data

Notes. Relative standard deviations, autocorrelation and correlation coefficients in this Table correspond to the quarterly data series detrended using a Hodrick-Prescott filter with smoothing parameter 1600. Each data series x correspond to a variable in the model. *The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2.*

of parameters suggested in the literature. We conduct a robustness exercise with different values of the gain parameter in Section 6.

4.2 US Data

In this section we compare the main statistical properties of the simulated labour market series generated from the model with the corresponding series in the US data, in particular focusing on second moments. It is standard practice, in one-worker-one-firm models, to compute the standard deviations of labour market data relative to productivity, but, in large-firm models, the standard measures that capture the volatilities of data are typically divided by the corresponding measure of output. Since our model is a large-firm model, we take the latter approach to compute relative standard deviations.

The seasonally adjusted series of (un)employment is taken from the Bureau of Labour 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 $2014Q4$ (due to the data availability on vacancies). Table 2 summarises 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 labour market tightness as reported in Table 2. In particular, both vacancies and unemployment are about 9.40 and 8.84 times more volatile than the aggregate output respectively. Moreover, the measure of labour market tightness is around 17.82 times more volatile than output.[18]

[18] Another well known stylised fact in labour market dynamics is the negative relationship between vacancies and unemployment, also known as the Beveridge curve.

4.3 Simulation Results

We simulate a 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 and are 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 this procedure $1,000$ times and report the mean values of the variables of interest.[19] Since our model is calibrated for monthly frequencies and the US data is reported only in quarterly frequencies, we then convert the monthly simulated series into quarterly frequencies. Finally, we transform the simulated series from deviations into percentage deviations by dividing the simulated series by the corresponding steady state values.

Tables 3 and 4 report the statistical properties of the simulated series of interest under learning for all three belief specifications. In the tables and figures below, a hat over a variable denotes the percentage deviation of the variable from its steady state value (eg. $\hat{y} = \tilde{y}/\bar{y}$). Table 3 shows that the RE model generates very little amplification in labour market variables. The table shows that the search and matching model with learning can replicate the second moments of the US labour market remarkably well. All the 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 8.87 and 7.00 times more volatile than output respectively with the corresponding numbers in US data being 9.30 and 8.84. All the learning models do significantly better than the RE model in matching amplification in the data. The correlations and autocorrelations in the VAR(1) model are, however, worse than those in the AR(1) and AR(2) specifications.

4.4 Impulse Responses

In this sub-section, we study how labour 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 1% standard deviation

[19] We check the stability of each learning model at each point in time by examining the highest eigenvalue of the system and when this stability condition is not met we disregard the entire draw. We discard around 12% of the draws in the univariate AR models and around 23% in the VAR(1) model. 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. We have experimented with other initial values and our results do not change significantly.

Model	Statistics		$\hat{y}_t$	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$	$\hat{c}_t$
RE	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$		1	0.03	0.50	0.43	0.91	1.00
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.96	0.97	0.92	0.97	0.96	0.96
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.99	0.99	−0.99	1.00	1.00
		$\hat{n}_t$	–	1	0.96	−1.00	0.99	0.99
		$\hat{v}_t$	–	–	1	−0.96	0.99	0.99
		$\hat{u}_t$	–	–	–	1	−0.99	−0.99
		$\hat{\theta}_t$	–	–	–	–	1	1.00
		$\hat{c}_t$	–	–	–	–	–	1
AR(1)	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$		1	0.58	11.40	9.16	19.85	0.92
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.95	0.94	0.71	0.94	0.89	0.94
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.95	0.83	−0.96	0.92	1.00
		$\hat{n}_t$	–	1	0.94	−0.99	0.99	0.93
		$\hat{v}_t$	–	–	1	−0.86	0.97	0.79
		$\hat{u}_t$	–	–	–	1	−0.96	−0.95
		$\hat{\theta}_t$	–	–	–	–	1	0.89
		$\hat{c}_t$	–	–	–	–	–	1

Table 3: Summary Statistics: RE and AR(1) learning model

Notes. Relative standard deviations, autocorrelation and correlation coefficients in this Table correspond to the quaterly simulated series expressed in percentage deviations from steady state value. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2.

Model	Statistics		$\hat{y}_t$	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$	$\hat{c}_t$
AR(2)	$\sigma_{\widehat{x}_{1t}}/\sigma_{\widehat{y}_t}$		1	0.44	8.87	7.00	15.21	0.94
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.95	0.94	0.70	0.94	0.89	0.94
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.94	0.80	−0.95	0.90	1.00
		$\hat{n}_t$	–	1	0.92	−0.98	0.99	0.92
		$\hat{v}_t$	–	–	1	−0.83	0.96	0.76
		$\hat{u}_t$	–	–	–	1	−0.95	−0.94
		$\hat{\theta}_t$	–	–	–	–	1	0.88
		$\hat{c}_t$	–	–	–	–	–	1
VAR(1)	$\sigma_{\widehat{x}_{1t}}/\sigma_{\widehat{y}_t}$		1	0.31	8.67	4.83	11.60	0.99
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.89	0.59	−0.07	0.61	0.35	0.88
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.64	0.30	−0.71	0.52	1.00
		$\hat{n}_t$	–	1	0.77	−0.91	0.95	0.59
		$\hat{v}_t$	–	–	1	−0.43	0.93	0.22
		$\hat{u}_t$	–	–	–	1	−0.74	−0.69
		$\hat{\theta}_t$	–	–	–	–	1	0.45
		$\hat{c}_t$	–	–	–	–	–	1

Table 4: Summary Statistics: AR(2) and VAR(1) learning models

Notes. Relative standard deviations, autocorrelation and correlation coefficients in this Table correspond to the quaterly simulated series expressed in percentage deviations from steady state value. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2.

(productivity) shock in period $10,001$ and compute the impulse responses as the difference between the two resulting set of impulse responses from period $10,001$ onwards. This experiment is then repeated $1,000$ 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 2 shows the impulse responses of aggregate output, (un)employment, vacancies, labour market tightness and the job filling 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 the general equilibrium restrictions that determine future wages. For this reason, expected wages tend to absorb most of the productivity increase and, as a result, unemployment, vacancies and the measure of labour market tightness respond only marginally to a TFP innovation.

Very distinct dynamic responses are observed under learning to a positive TFP innovation for all three 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. A productivity innovation thus leads to more vacancy creation under AL. The response of employment is largest for the AR(1) beliefs followed by AR(2) and VAR(1) beliefs; the amplification generated in the learning models is, however, much larger than that of the RE specification. Quantitatively, the magnitude of the responses is preserved for the next four years, i.e. it is greater under AR(1) than under AR(2) and VAR(1). After four years, the magnitude of these responses is reversed i.e. it is smaller under AR(1) than under AR(2) and VAR(1). We observe over and under-shooting of the variables under AL compared to RE as they converge to the steady state.

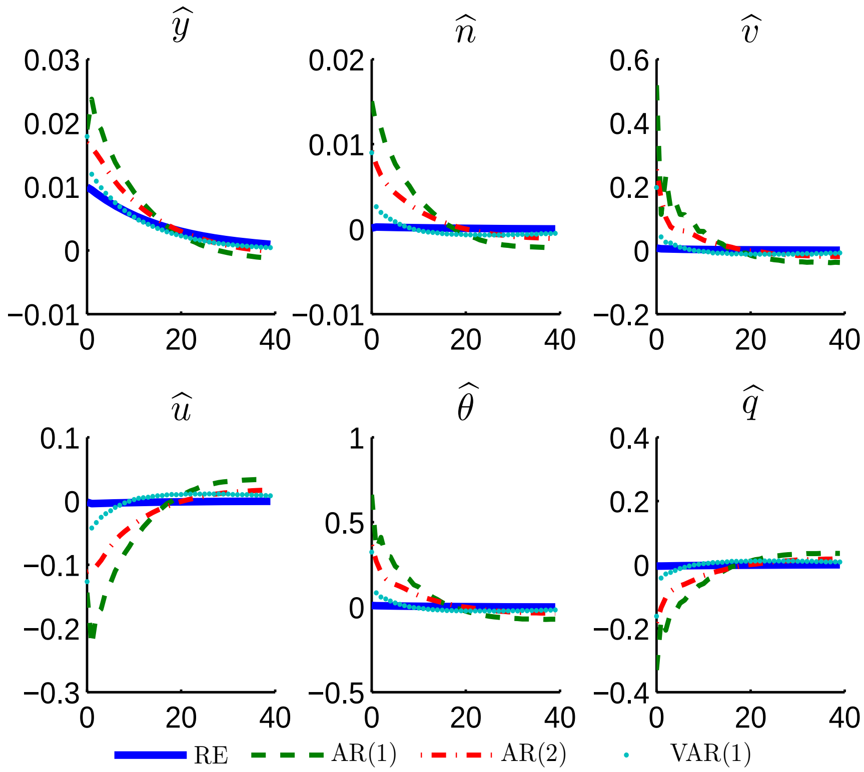

Figure 2: Impulse responses to labour market variables. *Notes. Impulse responses to a 1% standard deviation (productivity) shock. The solid and dotted lines show the impulse responses for the key labour variables under RE and AL. Percentage deviations from steady state values reported along the vertical axis. The horizontal axis display the number of quarters after the shock.*

Figure 2 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 response of output is approximately the same as the TFP innovation, which is a feature of the *unemployment volatility puzzle*. The reason behind the greater response in output under learning has to do with a larger increase in employment. The dynamics of consumption (not plotted in the figure) is qualitatively the same as that of output.

Intuitively, as the labour market tightens and the job finding probability falls, both the marginal costs and the benefits associated with filling an additional vacancy tend to increase. The response in unemployment is, however, more sluggish relative to the response in vacancies given that unemployment is a pre-determined variable.

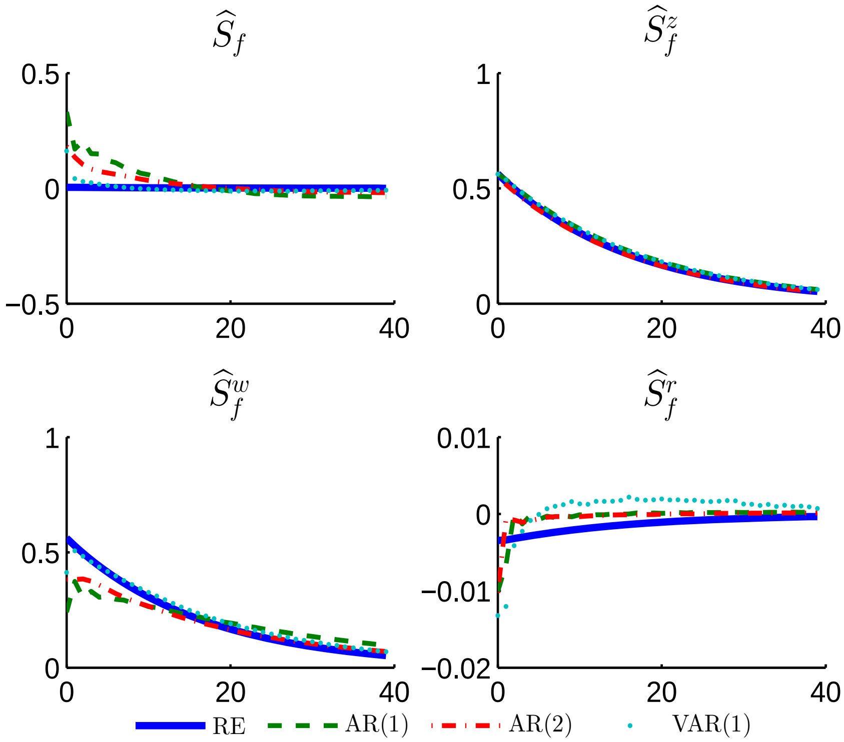

Figure 3: Impulse responses to infinite sums in job creation condition. *Notes. Impulse responses to a 1% standard deviation (productivity) shock. The solid and dotted lines show the impulse responses for the infinite sums in the job creation condition under RE and AL. Percentage deviations from steady state values reported along the vertical axis. The horizontal axis display the number of quarters after the shock.*

To disentangle further these effects under learning and gain intuition, we rearrange the job creation condition, equation (26), in linearised form. For simplicity, we set α equal to 1, in which case $\bar{\lambda}_1 = 0$ and $\bar{\lambda}_2 = 1$ and, hence, $\mathcal{S}^n_{ft} = 0$.[20] Thus, we have

$$-\frac{\kappa}{\bar{q}^2}\tilde{q}_t = \mathcal{S}_{ft} \equiv \mathcal{S}^z_{ft} - \mathcal{S}^w_{ft} - \mathcal{S}^r_{ft}. \tag{40}$$

and we plot the impulse responses of the infinite sums in (40). As shown earlier, the infinite sum $\mathcal{S}_{ft}$, which denotes the present discounted value of profits per hire, responds much more strongly under AL relative to RE. Figure 3 decomposes this infinite sum, $\mathcal{S}_{ft}$, into the three components, $\mathcal{S}^z_{ft}$, $\mathcal{S}^w_{ft}$ and $\mathcal{S}^r_{ft}$. Note that $\mathcal{S}^z_{ft}$ is the same under RE and learning and, hence, the lines overlap. This figure shows that the order of magnitude of $\mathcal{S}^r_{ft}$ is much smaller relative to that of $\mathcal{S}^z_{ft}$ and $\mathcal{S}^w_{ft}$ (note the vertical access of $\mathcal{S}^r_{ft}$); the sum of stochastic discount factors displays little variation relative to the other two infinite sums.[21] Note that the size of the responses of both $\mathcal{S}^z_{ft}$ and $\mathcal{S}^w_{ft}$ under RE are of equal magnitudes, which explains the negligible amplification generated by the standard RE search and matching model. In sharp contrast, the magnitude of the response of $\mathcal{S}^z_{ft}$ is much higher

[20] Note that wage bargaining turns out to be the same for small and large firms when $\alpha = 1$.

[21] This result is suggestive that a model where households' preferences are linear in consumption is likely to generate similar results.

than that of $\mathcal{S}^w_{ft}$ after a TFP innovation under learning for all belief specifications. The difference between $\mathcal{S}^z_{ft}$ and $\mathcal{S}^w_{ft}$ explains the strong amplification mechanism of the learning models and the high variation in vacancy posting on impact. The extent of the difference in these two magnitudes, and hence amplification, depends on the forecasting models used by agents.

4.4.1 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 amplification in labour market variables and then provide an intuition as to why the learning models are able to provide a solution for the unemployment volatility puzzle.

A TFP shock has well-understood implications in the standard search and matching model under RE beliefs. A positive innovation, ϵ_{t+1}, shifts the production frontier and the labour demand schedule, by increasing labour 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 more vacancy creation. However, the extent to which the model is able to generate amplification depends on the way in which wages are determined.

Nash bargained wages tend to absorb most of the increase in productivity under RE, leading to little incentive for vacancy creation. To get further intution, we start from equation (40). We ignore the term involving $\mathcal{S}^r_{ft}$ in equation (40), since it plays a minor role, and re-write equation (40) approximately as

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\bar{q}}{\kappa} E_t \sum_{j=1}^{\infty} (1-\rho)^{j-1} \beta^j \left[\tilde{z}_{t+j} - \tilde{w}_{t+j}\right] \equiv \frac{\bar{q}}{\kappa} \left(\mathcal{S}^z_{ft} - \mathcal{S}^w_{ft}\right). \quad (41)$$

We replace E^*_t by E_t, the expectation operator under RE. Since unemployment is a pre-determined 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

$$\begin{aligned} \tilde{n}_{t+1} &= \bar{a}_{nn}\tilde{n}_t + \bar{a}_{nz}\tilde{z}_t, \\ \tilde{w}_t &= \bar{a}_{wn}\tilde{n}_t + \bar{a}_{wz}\tilde{z}_t, \\ \tilde{z}_{t+1} &= \varrho\tilde{z}_t + \epsilon_{t+1}, \end{aligned}$$

where $\bar{a}_{x_1x_2}$ denotes the elasticities of variable x_1 with respect to variable x_2. For the baseline calibration, $0 < \bar{a}_{x_1x_2} < 1$ for all possible combinations of x_1 and x_2. Plugging the RE solution into (41) yields

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\bar{\beta q}}{\kappa}\left(\bar{\phi}_1\tilde{z}_t - \bar{\phi}_2\tilde{n}_t - \bar{\phi}_3\tilde{z}_t\right), \tag{42}$$

where

$$\bar{\phi}_1 = \frac{\bar{\varrho}}{1-\beta\left(1-\rho\right)\bar{\varrho}}, \quad \bar{\phi}_2 = \frac{\bar{a}_{wn}\bar{a}_{nn}}{1-\beta\left(1-\rho\right)\bar{a}_{nn}} \quad \text{and} \tag{43}$$

$$\bar{\phi}_3 = \frac{\bar{a}_{wn}\bar{a}_{nz}\left(1-\beta\left(1-\rho\right)\rho\right) + \varrho\bar{a}_{wz}\left(1-\beta\left(1-\rho\right)\bar{a}_{nn}\right)}{\left[1-\beta\left(1-\rho\right)\bar{a}_{nn}\right]\left[1-\beta\left(1-\rho\right)\varrho\right]}. \tag{44}$$

This expression provides the key insight into the lack of amplification under RE. At the outset, note that the vacancy posting decision does not respond to employment because $\tilde{n}_t$ is pre-determined; 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 $\bar{\phi}_1\tilde{z}_t$ in equation (42), which corresponds to the first infinite sum in (41)) and an indirect effect from productivity innovations through shifts in future wage costs (the terms $\bar{\phi}_2\tilde{n}_t$ and $\bar{\phi}_3\tilde{z}_t$ in equation (42), which correspond to the second infinite sum in (41)). For the baseline calibration, it can be shown that $\bar{\phi}_1$ and $\bar{\phi}_3$ in equation (42) are very close to each other, which implies that productivity innovations have little impact on the present discounted value of profits per hire.[22] In other words, when firms correctly forecast future profits per hire, the variability of profits per hire is dampened because wages are flexible and 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 general 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 general equilibrium restrictions, i.e. 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.

To keep the explanation simple, we restrict ourselves to a discussion of the results under learning with AR(1) beliefs. Similar logic applies to the AR(2) and VAR(1) belief specifications. Under AR(1) beliefs, firms estimate the parameter a_{w1} by running the following regression $\tilde{w}_t = a_{w1}\tilde{w}_{t-1}$, to make their forecasts of future wages. Combining

[22] The baseline calibration ensures that our results are not driven by the HM strategy; see Hagedorn and Manovskii (2008).

this expression with (41) yields

$$-\frac{\tilde{q}_t}{\bar{q}} \approx \frac{\beta\bar{q}}{\kappa}\left(\phi_1\tilde{z}_t - \phi_2\tilde{w}_{t-1}\right), \tag{45}$$

where $\phi_1 = \frac{\varrho}{1-\beta(1-\rho)\varrho}$and $\phi_2 = \frac{\beta a_{w1}^2}{1-\beta(1-\rho)a_{w1}}$. A TFP innovation increases the discounted value of future marginal products of employment, the term $\phi_1\tilde{z}_t$ in equation (45), but has no direct impact effect on the present discounted value of future wage costs, the term $\phi_2\tilde{w}_{t-1}$.[23] The high persistence of the TFP innovation is key for generating amplification under learning because it raises the rate at which future marginal products of employment are discounted. This greater persistence together with 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 labour market becomes tighter and the job filling rate falls after a TFP innovation, which then increases the expected costs of vacancy posting. After a technology innovation, expected marginal products of employment increase more than expected wage payments. More vacancies in equilibrium translate into higher employment and output. Figure 3 shows that the discounted value of future costs (wages) is lower than the discounted value of future marginal products under AL.

To provide further clarity, Figure 4 plots the path of simulated $\mathcal{S}_{ft}^z - \mathcal{S}_{ft}^w$ (i.e. the present discounted value of profits per hire in the model with constant returns to scale) 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 (note the magnitude on the vertical axis). The figure shows that the difference in the two sums is greatest under AR(1) beliefs followed by AR(2) and VAR(1) beliefs. Note, however, that the order of magnitude in the discounted value of profits per hire even under VAR(1) beliefs is significantly higher than under RE beliefs. This means that learning models can help better match the relative volatility of labour market variables with the US data. The persistence of the simulated series is highest under AR(1) and AR(2) beliefs, helping explain the greater propagation in vacancy creation.

[23]Recall that the elasticity $\bar{\phi}_1$ under RE differs from the elasticity ϕ_1 under learning because ϱ is estimated at each point in time.

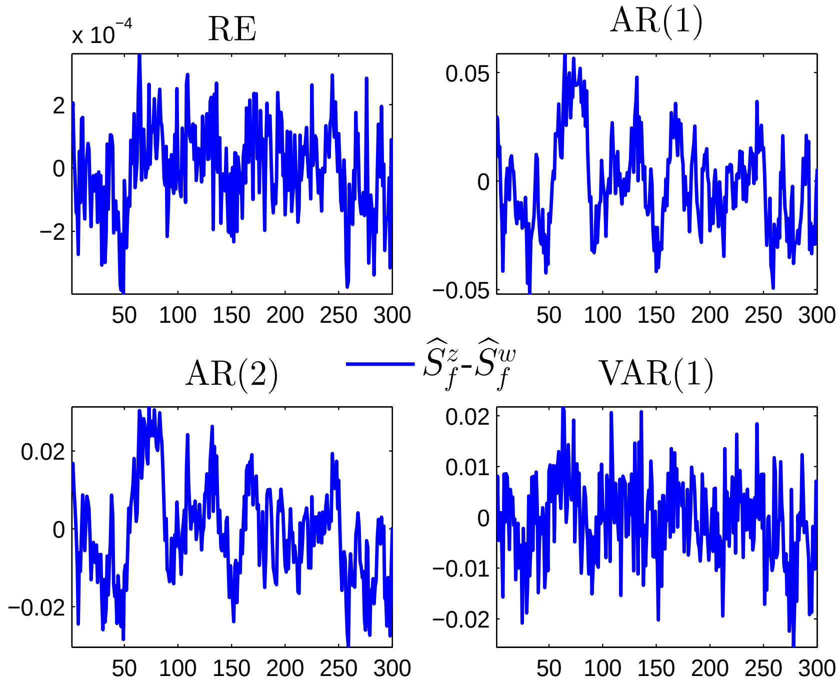

Figure 4: Simulated path of infinte sums in the job creation condition. *Notes. Infinite sums are expressed in percentage deviations from state state values. The first 10000 months are discarded and the remaining 900 months averaged over 500 replications under RE and AL. Monthly series are converted into quarterly series.*

5 Forecast Properties of the Model

Vacancy posting decisions in the search and matching models typically depend on the marginal profitability of long-term employment relationships. Thus, (un)employment and 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 this assumption can generate greater amplification in the value of a filled vacancy. Given that forecasts are central for explaining amplification in labour market data, a suitable test of the AL models would be to compare how well the statistical properties of forecasts generated by the AL models (and RE) compare with those in the data. As mentioned in Section 1, adaptive learning models have been used to successfully match macro and survey data (see e.g. Adam, Marcet, and Beutel (2016), Kuang and Mitra (2016), Slobodyan and Wouters (2012), Ormeno and Molnar (2015) and Eusepi and Preston (2011)). Since wage forecasts are not available in the Survey of Professional Forecasters (SPF), we focus only on unemployment forecasts. We show that our learning models are more successful than RE at matching survey data on unemployment forecasts.

As indicated by the PLMs, agents forecast all future variables using simple autoregressive models. Since agents in our model are required to forecast (un)employment rates, we can therefore compare the performance of forecast errors of unemployment rates relative to the data. We take the quarterly forecasts of unemployment rates from the SPF, which are collected by the Federal Reserve Bank of Philadelphia. Questionnaires are sent to the individual forecasters to carry out their expert forecasts at the end of the first month of each quarter and individual forecasts are released towards the end of the second month of each quarter. Forecast data are available from 1968Q4 to 2014Q4 and both mean and median of forecasts are reported up to four quarters ahead. We compute the forecast errors by removing the log of the actual realisation from the log of the forecast data of the unemployment rates up to four periods ahead using BLS data.

Table 5 reports the statistical properties of forecast errors, including their relative volatility with respect to detrended output, autocorrelation of the forecast errors and their correlation with the first differences of output and unemployment. We find that the forecast errors are much more volatile relative to output over the 4-quarter horizon. The table shows that the one quarter-ahead forecast errors is almost 4.56 times more volatile than output. In addition, the quarterly forecast errors display high serial correlation, which is stronger as the time horizon of the forecast increases. Furthermore, over the business cycle, the forecast errors of unemployment are all pro-cyclical, which means agents tend to over-predict unemployment during expansions and under-predict it in recessions.

The statistical properties of the simulated forecast errors, expressed in percentage deviations from the steady state unemployment rate under both RE and AL, are also reported in Table 5. The simulation methods used here are the same as described in Section 4.4. The quarterly and yearly forecast errors of unemployment are computed as the three and twelve months ahead forecasts. Consistent with SPF the quarterly and yearly forecasts are conducted in the second month of the quatter. It follows from the table that the performance of the RE model at matching the statistical properties of the forecast errors in the data is rather poor. The RE model predicts a very low relative volatility of the forecast errors, negligible autocorrelation and no pro-cyclicality.

The learning models perform better than the RE model in matching the statistical properties of forecast errors. In particular, the relative volatility of forecast errors *vis-a-vis* output are better matched by all the learning models; in sharp contrast RE does a poor job in this regard. AR(1) and AR(2) models match the first order autocorrelations of forecast errors two to four quarters ahead than the RE and VAR(1) models. The models have less success in matching the first order autocorrelation of the one-step ahead forecast errors; the AR(2) model is the best in this regard. The correlations between the forecast errors and the changes in unemployment and output are less well matched with the data; qualitatively, the AR(2) matches the signs of the correlations.[24] Arguably it

[24] In an early version Di Pace, Mitra, and Zhang (2014), we show that the forecast error properties in

		FE_t^{Q1}	FE_t^{Q2}	FE_t^{Q3}	FE_t^{Q4}
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	4.56	6.21	8.02	9.89
Data	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	0.70	0.79	0.85	0.89
	$\rho(\hat{x}_{1t}, \Delta\hat{u}_t)$	−0.69	−0.55	−0.47	−0.36
	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	0.44	0.42	0.37	0.32
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	0.05	0.12	0.18	0.21
RE	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	0.13	0.24	0.61	0.73
	$\rho(\hat{x}_{1t}, \Delta\hat{u}_t)$	−0.10	−0.04	−0.03	0.00
	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	0.27	0.00	0.04	0.0
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	3.38	5.11	6.50	7.42
AR(1)	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	0.07	0.73	0.84	0.90
	$\rho(\hat{x}_{1t}, \Delta\hat{u}_t)$	0.00	−0.14	−0.19	−0.18
	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	0.12	0.20	0.24	0.23
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	2.82	3.97	4.79	5.42
AR(2)	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	0.33	0.70	0.84	0.88
	$\rho(\hat{x}_{1t}, \Delta\hat{u}_t)$	−0.10	−0.11	−0.15	−0.17
	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	0.14	0.14	0.20	0.21
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	3.93	4.49	4.71	4.80
VAR(1)	$\rho(\hat{x}_{1t}, \hat{x}_{1t-1})$	0.13	0.53	0.59	0.61
	$\rho(\hat{x}_{1t}, \Delta\hat{u}_t)$	0.08	−0.02	0.03	−0.12
	$\rho(\hat{x}_{1t}, \Delta\hat{y}_t)$	0.02	0.10	0.13	0.20

Table 5: Forecast Properties

Notes. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2. Data from Survey of Professional Forecasters. Forecast errors are defined as $FE_t^{Qj} = \ln\left(u_{t+j}^F\right) - \ln(u_{t+j})$ for $j = 1, 2, 3$ and 4 in the data, where u_{t+j}^F denotes forecasted unemployment,.and as $FE_t^{Qj} = E_t\hat{u}_{t+j} - \hat{u}_{t+j}$ for $j = 1, 2, 3$ and 4 in the models.

is remarkable that our simple models with learning are able to match the data along so many dimensions especially in terms of amplification and forecast error properties.

5.1 Wage Rigidity under RE as an alternative?

As mentioned in Section 1, the introduction of wage rigidity in the standard search and matching model under RE is able to generate greater amplification in labour market variables. The rationale behind this result is simple: if quantities display little amplification after TFP innovations, fixing prices have the potential to generate the expected results. We now show that, while the introduction of wage rigidities leads to higher volatility under RE (as is well known in the literature), the properties of forecast errors of unemployment are less well matched to those of the SPF.

In accordance with Hall (2005) a wage norm or social consensus can be perceived as a rule to select an equilibrium for the wage within the bargaining set. We assume that the wage is given by a weighted average of the steady state wage ($\bar{w}$) and the negotiated

unemployment are better matched by an alternative set of beliefs where agents regress the key endogenous variables only on unemployment (but not on the productivity shock). This is perhaps due to the fact that agents *only* use unemployment rates in this alternative version unlike our current learning models.

Relative Volatilities					
	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$	$\hat{c}_t$
$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	0.38	6.97	6.00	12.84	0.94
Forecast Error Properties					
Statistics		FE_t^{Q1}	FE_t^{Q2}	FE_t^{Q3}	FE_t^{Q4}
$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$		0.58	1.73	2.48	3.05
$\rho\left(\hat{x}_{1t}, \hat{x}_{1t-1}\right)$		−0.04	0.25	0.62	0.76
$\rho\left(\hat{x}_{1t}, \Delta\hat{u}_t\right)$		0.00	−0.07	−0.08	−0.09
$\rho\left(\hat{x}_{1t}, \Delta\hat{y}_t\right)$		−0.12	0.03	0.07	0.08

Table 6: Wage Rigidities and RE

Notes. Relative standard deviations this table correspond to the quarterly simulated series expressed in percentage deviations from the steady state values in the RE model featuring wage rigidities. Forecast errors FE_t^{Qj} for $j = \{1, 2, 3, 4\}$ as defined in Table 5.

wage under the Nash protocol in equation (27),

$$\tilde{w}_t = \omega\bar{w} + (1-\omega)\,\tilde{w}_t^{Nash},$$

where $\tilde{w}_t^{Nash}$ is the bargained wage in (27) and $\omega \in (0,1)$ denotes the wage rigidity parameter. We set the value of ω to 0.7 in line with Blanchard and Gali (2010).

Table 6 reports the relevant statistics for this model. Amplification in labour market variables and the forecast error properties are much improved relative to the RE model with flexible wages. Nevertheless, learning models continue to perform better in terms of amplification in labour market variables (which are closer to the data) relative to the RE model with wage rigidity.[25] In addition, they provide a much better fit to the forecast error properties in the data than RE. For this reasons, the learning models are our preferred relative to sticky wages under RE. In one sense, the learning models maybe thought of a way to rationalise wage rigidity under RE with the distinction that, while under learning wages respond fully to TFP innovations, wage expectations are sluggish.

6 Robustness

In this subsection we look at the sensitivity of our results to alternative parameterisations of the learning models. We first consider a calibration of the model with decreasing returns to scale in labour and then consider two values of the constant gain parameter. For economy of space, we choose to focus on the AR(2) specification for our robustness exercise.

We consider two alternative values of the gain parameter to check the sensitivity

[25] We have experimented with higher values of ω and we find that relative volatilities of labour market variables are in line with the data but the correlations remain the same. In addition, wage norms that depend on past wages, rather than on steady state values, the RE model struggles to match the data along all dimensions.

	$\sigma_{\hat{n}_t}/\sigma_{\hat{y}_t}$	$\sigma_{\hat{v}t}/\sigma_{\hat{y}_t}$	$\sigma_{\hat{u}_t}/\sigma_{\hat{y}_t}$	$\sigma_{\hat{\theta}_t}/\sigma_{\hat{y}_t}$	$\sigma_{\hat{c}_t}/\sigma_{\hat{y}_t}$
Data	0.57	9.30	8.84	17.71	0.79
Alternative parameterisations					
AR(2) (baseline)	0.44	8.87	7.00	15.21	0.94
AR(2) ($\gamma = 0.001$)	0.44	8.94	6.95	15.19	0.94
AR(2) ($\gamma = 0.003$)	0.44	9.75	6.87	15.44	0.95
AR(2) ($\alpha = 0.7$)	0.43	9.78	6.81	15.25	0.94
AR(2) ($\sigma = 0.6$ and $\xi = 0.4$)	0.50	8.75	7.76	14.57	0.96

Table 7: Alternative Parameterisations

Notes. Relative standard deviations in this Table correspond to the quarterly simulated series expressed in percentage deviations from steady state values under different parameterisations.

of our results to the way in which agents discount past information. We choose the following two parameters of γ: 0.001 and 0.003. Table 7 reports the relative standard derivations of labour market variables under the two alternative parameterisations (along with the results for the baseline AR(2) model). Our simulations show that the values of the gain parameter have very little influence in terms of the statistical properties of the simulated data. A smaller gain parameter indicates that agents put more weight on past data relative to current data to update beliefs, which means that the time it takes for the effects of TFP innovations to vanish from agents' information set is longer. The amplification results are therefore not much affected by the choice of the parameter γ.

The calibration of our baseline model assumes constant returns to scale in labour. In line with models featuring physical capital, we set the value of α equal to 0.7 and simulate the model under this new parameterisation and check the sensitivity of the robustness of the results. The main conceptual difference with the baseline calibration is that the demand curve under decreasing returns to labour is downward sloping rather than perfectly elastic. We find that the results from this simulation are in line with the simulations of our baseline AR(2) model. Table 7 shows that the labour market variables still display a great deal of variation.

Finally we carry a robustness exercise with regards to the elasticity of the matching function with respect to vacancies (σ) and the bargaining power of workers (ξ). In carrying out this exercise we ensure that Hosios condition is met by increasing the elasticity of the matching function with respect to vacancy and reducing at the same time the bargaining power of workers by 0.1 relative to the baseline calibration. Table 7 shows that this alternative calibration delivers robust results under AR(2) beliefs. Table 10 in the Appendix shows that the forecast errors properties under learning are unchanged under these robustness exercises i.e. the learning models continue to provide a good match to the data.

7 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 rational expectations (RE) to study the role of adaptive learning (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 US labour market and survey data very well, outperforming the standard RE model.

In particular, under learning with autoregressive beliefs, such as AR(1), AR(2) and VAR(1), we find that Total Factor Productivity (TFP) innovations can generate greater incentive for vacancy creation than RE. A TFP innovation under learning increases the firms' present discounted value of profits per hire (i.e., firms become overoptimistic). Thus, expectation formation has a central role to play in providing a solution to *unemployment volatility puzzle*. Moreover, our simple learning set-up has the additional advantage of being able to match the properties of forecast data on unemployment, which is compatible with the expectation formation assumption in the paper.

References

Adam, K., and A. Marcet (2011): "Internal rationality, imperfect market knowledge and asset prices," *Journal of Economic Theory*, 146(3), 1224–1252.

Adam, K., A. Marcet, and J. Beutel (2016): "Stock Price Booms and Expected Capital Gains," mimeo., University of Mannheim.

Alves, S. A. L. (2012): "Trend Inflation and the Unemployment Volatility Puzzle," Working Papers Series 277, Central Bank of Brazil, Research Department.

Andolfatto, D. (1996): "Business Cycles and Labor-Market Search," *American Economic Review*, 86(1), 112–32.

Barnichon, R. (2010): "Building a composite Help-Wanted Index," *Economics Letters*, 109(3), 175–178.

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

Branch, W. (2004): "Restricted Perceptions Equilibria and Learning in Macroeconomics," in Colander (2004).

Branch, W. A., and G. W. Evans (2006a): "Intrinsic heterogeneity in expectation formation," *Journal of Economic Theory*, 127(1), 264–295.

——— (2006b): "A simple recursive forecasting model," *Economics Letters*, 91(2), 158–166.

Bullard, J., and K. Mitra (2002): "Learning about monetary policy rules," *Journal of Monetary Economics*, 49(6), 1105–1129.

Campbell, J. Y. (1994): "Inspecting the mechanism: An analytical approach to the stochastic growth model," *Journal of Monetary Economics*, 33(3), 463–506.

Cho, I.-K., N. Williams, and T. J. Sargent (2002): "Escaping Nash Inflation," *Review of Economic Studies*, 69(1), 1–40.

Christiano, L., M. Eichenbaum, and M. Trabandt (2016): "Unemployment and Business Cycles," *Econometrica*, 84(4), 1523–1569.

Christoffel, K., and K. Kuester (2008): "Resuscitating the wage channel in models with unemployment fluctuations," *Journal of Monetary Economics*, 55(5), 865–887.

Cogley, T., and T. J. Sargent (2008): "Anticipated Utility And Rational Expectations As Approximations Of Bayesian Decision Making," *International Economic Review*, 49(1), 185–221.

Colander, D. (ed.) (2004): *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*. Cambridge University Press, Cambridge.

Colciago, A., and L. Rossi (2011): "Endogenous Market Structures and Labor Market Dynamics (New version)," Quaderni di Dipartimento 155, University of Pavia, Department of Economics and Quantitative Methods.

Costain, J. S., and M. Reiter (2008): "Business cycles, unemployment insurance, and the calibration of matching models," *Journal of Economic Dynamics and Control*, 32(4), 1,120–55.

Davis, S., J. Haltiwanger, and S. Schuh (1996): *Job creation and destruction*. MIT Press.

den Haan, W. J., G. Ramey, and J. Watson (2000): "Job destruction and the experiences of displaced workers," *Carnegie-Rochester Conference Series on Public Policy*, 52(1), 87–128.

Di Pace, F., K. Mitra, and S. Zhang (2014): "Adaptive Learning and Labour Market Dynamics," CDMA Working Paper Series 201408, Centre for Dynamic Macroeconomic Analysis.

Ellison, M., and T. Yates (2007): "Escaping Volatile Inflation," *Journal of Money, Credit and Banking*, 39, 981–993.

Eusepi, S., and B. Preston (2011): "Expectations, Learning, and Business Cycle Fluctuations," *American Economic Review*, 101(6), 2844–72.

Evans, G., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics (Frontiers of Economic Research)*. Princeton.

Evans, G. W., and S. Honkapohja (2003): "Expectations and the Stability Problem for Optimal Monetary Policies," *Review of Economic Studies*, 70(4), 807–824.

——— (2006): "Monetary Policy, Expectations and Commitment," *Scandinavian Journal of Economics*, 108, 15–38.

Evans, G. W., S. Honkapohja, and N. Williams (2010): "Generalized Stochastic Gradient Learning," *International Economic Review*, 51(1), 237–262.

Gertler, M., L. Sala, and A. Trigari (2008): "An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining," *Journal of Money, Credit and Banking*, 40(8), 1,713–64.

Gertler, M., and A. Trigari (2009): "Unemployment fluctuations with staggered Nash wage bargaining," *Journal of Political Economy*, 117(1), 38–86.

Gomes, P. (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.

Guerrieri, V. (2008): "Heterogeneity, Job Creation and Unemployment Volatility," *Scandinavian Journal of Economics*, 109(4), 667–693.

Haefke, C., M. Sonntag, and T. van Rens (2013): "Wage Rigidity and Job Creation," *Journal of Monetary Economics*, 60(8).

Hagedorn, M., and I. Manovskii (2008): "The cyclical behavior of equilibrium unemployment and vacancies revisited," *American Economic Review*, 98(4), 1,692–706.

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

Heemeijer, P., C. Hommes, J. Sonnemans, and J. Tuinstra (2009): "Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation," *Journal of Economic Dynamics and Control*, 33(5), 1052–1072.

Hertweck, M. S. (2013): "Strategic wage bargaining, labor market volatility, and persistence," *The B.E. Journal of Macroeconomics*, 13(1), 27.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005): "Coordination of Expectations in Asset Pricing Experiments," *Review of Financial Studies*, 18(3), 955–980.

Hosios, A. J. (1990): "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57(2), 279–98.

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

Jacobs, D., E. Kalai, and M. Kamien (eds.) (1998): *Frontiers of Research in Economic Theory*. Cambridge University Press, Cambridge.

Krause, M., and T. A. Lubik (2007): “Does intra-firm bargaining matter for business cycle dynamics?,” Discussion paper.

Kreps, D. M. (1998): “Anticipated Utility and Dynamic Choice,” in Jacobs, Kalai, and Kamien (1998), pp. 242–274.

Kuang, P., and K. Mitra (2016): “Long-Run Growth Uncertainty,” *Journal of Monetary Economics*, 79, 67–80.

McGough, B. (2006): “Shocking Escapes,” *Economic Journal*, 116, 507–528.

Menzio, G. (2005): “High frequency wage rigidity,” Discussion paper, *mimeo.*

Menzio, G., and S. Shi (2011): “Efficient Search on the Job and the Business Cycle,” *Journal of Political Economy*, 119(3), 468 – 510.

Milani, F. (2007): “Expectations, learning and macroeconomic persistence,” *Journal of Monetary Economics*, 54(7), 2065–2082.

——— (2011): “Expectation Shocks and Learning as Drivers of the Business Cycle,” *Economic Journal*, 121, 379–401.

Mitra, K., G. W. Evans, and S. Honkapohja (2013): “Policy change and learning in the RBC model,” *Journal of Economic Dynamics and Control*, 37(10), 1947–1971.

——— (2016): “Fiscal Policy Multipliers in an RBC Model with Learning,” *Macroeconomic Dynamics,* forthcoming.

Mortensen, D., and E. Nagypal (2007): “More on Unemployment and Vacancy Fluctuations,” *Review of Economic Dynamics*, 10(3), 327–347.

Mortensen, D. T., and C. A. Pissarides (1994): “Job Creation and Job Destruction in the Theory of Unemployment,” *Review of Economic Studies*, 61(3), 397–415.

Ormeno, A., and K. Molnar (2015): “Using Survey Data of Inflation Expectations in the Estimation of Learning and Rational Expectations Models,” *Journal of Money, Credit and Banking,* forthcoming.

Orphanides, A., and J. C. Williams (2007): “Robust monetary policy with imperfect knowledge,” *Journal of Monetary Economics*, 54(5), 1406–1435.

Petrongolo, B., and C. A. Pissarides (2001): “Looking into the black box: a survey of the matching function,” *Journal of Economic Literature*, 39(2), 390–431.

Petrosky-Nadeau, N. (2013): “Credit, Vacancies and Unemployment Fluctuations,” *Review of Economic Dynamics*, Forthcoming.

Petrosky-Nadeau, N., and E. Wasmer (2013): "The Cyclical Volatility of Labor Markets under Frictional Financial Markets," *American Economic Journal: Macroeconomics*, 5(1), 193–221.

Pissarides, C. A. (2009): "The unemployment volatility puzzle: is wage stickiness the answer?," *Econometrica*, 77, 1,339–369.

Preston, B. (2005): "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter," *International Journal of Central Banking*, 1(2).

——— (2006): "Adaptive learning, forecast-based instrument rules and monetary policy," *Journal of Monetary Economics*, 53(3), 507–535.

——— (2008): "Adaptive learning and the use of forecasts in monetary policy," *Journal of Economic Dynamics and Control*, 32(11), 3661–3681.

Quadrini, V., and A. Trigari (2008): "Public Employment and the Business Cycle," *Scandinavian Journal of Economics*, 109(4), 723–742.

Reiter, M. (2007): "Embodied technical change and the fluctuations of unemployment and wages," *The Scandinavian Journal of Economics*, 109(4), 695–721.

Robin, J.-M. (2011): "On the Dynamics of Unemployment and Wage Distributions," *Econometrica*, 79(5), 1327–1355.

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

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

——— (2010): *Labor Markets and Business Cycles (Crei Lectures in Macroeconomics)*. Princeton University Press.

Slobodyan, S., and R. Wouters (2012): "Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models," *American Economic Journal: Macroeconomics*, 4(2), 65–101.

Trigari, A. (2006): "The Role of Search Frictions and Bargaining for Inflation Dynamics," Working Papers 304, Innocenzo Gasparini Institute for Economic Research, Bocconi University.

A Appendix

A.1 Wage Bargaining

The surplus of workers in firm f, $\mathcal{W}_{ft}$, is given by the marginal value of employment, equation (46), expressed in terms of goods

$$\mathcal{W}_{ft} = \frac{\partial \mathcal{H}(s_t, n_t)}{\partial n_{ft}} c_t. \tag{46}$$

Thus,

$$\mathcal{W}_{ft} = w_{ft} - \chi c_t + [1 - \rho - p(\theta_t)] r_t^{-1} E_t^* \mathcal{W}_{ft+1}. \tag{47}$$

The negotiated wage w_{ft} is set to maximise the joint surplus of a match between workers and firm j,

$$\arg\max_{w_{ft}} (\mathcal{W}_{ft})^{\xi} (\mathcal{V}_{ft})^{1-\xi},$$

where $\xi \in (0,1)$ denotes the workers' bargaining power or the share of the surplus the worker is able to take. The first order condition of this problem is given by

$$(1-\xi)\mathcal{W}_{ft} = \xi \mathcal{V}_{ft}. \tag{48}$$

Substituting for equation (17) and (46) gives the negotiated wage w_{ft}

$$\begin{aligned} w_{ft} &= (1-\xi)\left\{\chi c_t - [1-\rho-p(\theta_t)] r_t^{-1} E_t^* \mathcal{W}_{ft+1}\right\} + \\ &+ \xi\left[\alpha z_t n_{ft}^{\alpha-1} + (1-\rho) r_t^{-1} E_{ft}^* \mathcal{V}_{ft+1}\right] \end{aligned}$$

Because of continuous re-negotiation, we assume that the first order condition also holds for subjective expectations

$$(1-\xi) E_t^* \mathcal{W}_{ft+1} = \xi E_{ft}^* \mathcal{V}_{ft+1} \tag{49}$$

to get the following expression

$$w_{ft} = (1-\xi)\left[\chi c_t + p(\theta_t) r_t^{-1} E_t^* \mathcal{W}_{ft+1}\right] + \xi \alpha z_t n_{ft}^{\alpha-1}. \tag{50}$$

Using (49) yields

$$w_{ft} = (1-\xi)\chi c_t + \xi\left[\alpha z_t n_{ft}^{\alpha-1} + p(\theta_t) r_t^{-1} E_{ft}^* \mathcal{V}_{ft+1}\right] \tag{51}$$

and substituting (16) gives

$$w_{ft} = (1-\xi)\chi c_t + \xi\left[\alpha z_t n_{ft}^{\alpha-1} + p(\theta_t)\frac{\kappa}{q(\theta_t)}\right]. \tag{52}$$

Rearranging gives equation (21) in the main text.

A.2 Behavioural Rules

In this appendix derive the linearised consumption and vacancy decision rules, equations (24) and (25)

A.2.1 Job Creation Equation

We first linearise equations (15) and (19)

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

$$\begin{aligned} -\frac{\kappa}{\bar{q}^2}\tilde{q}_t \quad = \quad & -\frac{\alpha\bar{z}\bar{n}^{\alpha-1}-\bar{w}}{\bar{r}^2}\tilde{r}_t + \\ & \frac{1}{\bar{r}}\alpha\bar{z}(\alpha-1)\bar{n}^{\alpha-2}\tilde{n}_{t+1} + \frac{1}{\bar{r}}\alpha\bar{n}^{\alpha-1}\tilde{z}_{t+1} - \frac{1}{\bar{r}}\tilde{w}_{t+1} + \\ & \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^j E_t^*\left[\alpha\bar{z}(\alpha-1)\bar{n}^{\alpha-2}\tilde{n}_{t+j} + \alpha\bar{n}^{\alpha-1}\tilde{z}_{t+j} - \tilde{w}_{t+j}\right] - \\ & (\alpha\bar{z}\bar{n}^{\alpha-1}-\bar{w})\sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j+1}\sum_{i=1}^{j}E_t^*\tilde{r}_{t+i-1}. \end{aligned} \tag{53}$$

By replacing $\tilde{n}_{t+1}$ into (53) and using $\beta = 1/\bar{r}$, we obtain

$$\begin{aligned} \bar{\lambda}_1\bar{q}\tilde{v}_t - \left[\frac{\kappa}{\bar{q}^2} + \beta\bar{\lambda}_1\bar{v}\right]\tilde{q}_t \quad = \quad & \beta\bar{\lambda}_1(1-\rho)\,\tilde{n}_t - \frac{(\bar{z}\bar{\lambda}_2-\bar{w})\beta^2}{1-\beta(1-\rho)}\tilde{r}_t + \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^j E_t^*\bar{\lambda}_1\tilde{n}_{t+j} + \\ & \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j E_t^*\left[\bar{\lambda}_2\tilde{z}_{t+j} - \tilde{w}_{t+j} - \left(\bar{z}\bar{\lambda}_2-\bar{w}\right)\beta\sum_{i=1}^{j-1}\tilde{r}_{t+i}\right] \end{aligned} \tag{54}$$

where $\bar{\lambda}_1 = \alpha(\alpha-1)\bar{z}\bar{n}^{\alpha-2}$ and $\bar{\lambda}_2 = \alpha\bar{n}^{\alpha-1}$.A more compact representation gives (25) in the main text

$$\bar{\lambda}_1\bar{q}\tilde{v}_t - \left(\frac{\kappa}{\bar{q}^2} + \beta\bar{\lambda}_1\bar{v}\right)\tilde{q}_t = \beta\bar{\lambda}_1(1-\rho)\,\tilde{n}_t - \frac{(\bar{z}\bar{\lambda}_2-\bar{w})\beta^2}{1-\beta(1-\rho)}\tilde{r}_t + \mathcal{S}_{ft}^z + \mathcal{S}_{ft}^n - \mathcal{S}_{ft}^w - \mathcal{S}_{ft}^r, \tag{55}$$

where

$$\begin{aligned} \mathcal{S}_{ft}^z \quad &= \quad \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^j\bar{\lambda}_2E_t^*\tilde{z}_{t+j}, \quad \mathcal{S}_{ft}^n = \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^j\bar{\lambda}_1E_t^*\tilde{n}_{t+j}, \\ \mathcal{S}_{ft}^w \quad &= \quad \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^jE_t^*\tilde{w}_{t+j} \quad \text{and } \mathcal{S}_{ft}^r = (\bar{z}\bar{\lambda}_2-\bar{w})\sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j+1}\sum_{i=1}^{j-1}E_t^*\tilde{r}_{t+i}. \end{aligned}$$

A.2.2 Consumption Function

By linearising (8) and substituting into the linearised version of equation (10), we get the following consumption rule

$$\begin{aligned}\frac{1}{1-\beta}\tilde{c}_t &= \tilde{\pi}_t + \bar{n}\tilde{w}_t + \bar{w}\tilde{n}_t + \beta\bar{w}\tilde{n}_{t+1} - \frac{(\bar{w}\bar{n}+\bar{\pi})\beta^2}{1-\beta}\tilde{r}_t \\ &\quad + \sum_{j=1}^{\infty}\beta^j E_t^*\bar{n}\tilde{w}_{t+j} + \sum_{j=2}^{\infty}\beta^j E_t^*\bar{w}\tilde{n}_{t+j} - (\bar{w}\bar{n}+\bar{\pi})\beta^2\sum_{j=1}^{\infty}\left[\beta^j E_t^*\sum_{i=1}^{j-1}\tilde{r}_{t+i}\right] + \sum_{j=1}^{\infty}\beta^j E_t^*\tilde{\pi}_{t+j} \\ &\equiv \tilde{\pi}_t + \bar{n}\tilde{w}_t + \bar{w}\tilde{n}_t + \beta\bar{w}\tilde{n}_{t+1} - \frac{(\bar{w}\bar{n}+\bar{\pi})\beta^2}{1-\beta}\tilde{r}_t + \mathcal{S}_{ht}^w + \mathcal{S}_{ht}^n + \mathcal{S}_{ht}^\pi - \mathcal{S}_{ht}^r,\end{aligned}$$

where

$$\begin{aligned}\mathcal{S}_{ht}^w &= \sum_{j=1}^{\infty}\beta^j\bar{n}E_t^*\tilde{w}_{t+j}, \quad \mathcal{S}_{ht}^n = \sum_{j=2}^{\infty}\beta^j\bar{w}E_t^*\tilde{n}_{t+j}, \\ \mathcal{S}_{ht}^\pi &= \sum_{j=1}^{\infty}\beta^j E_t^*\tilde{\pi}_{t+j} \quad \text{and} \quad \mathcal{S}_{ht}^r = (\bar{w}\bar{n}+\bar{\pi})\sum_{j=1}^{\infty}\left[\beta^{j+1}E_t^*\sum_{i=1}^{j}\tilde{r}_{t+i}\right].\end{aligned}$$

A.3 Learning

For expositional purpuses, we derive the infinite sums under AR(2). The derivation of the infinite sums under AR(1) and VAR(1) belief specifications are available upon request. As noted in the main text, the AR(2) belief specification is given by

$$x_t = a_{x0} + a_{x1}x_{t-1} + a_{n2}x_{t-2} \quad \text{for} \quad x_t = \{n_{t+1}, \pi_t, w_t, r_t\}$$

Let $\Phi_{x,t} = \begin{pmatrix} a_{x0} & a_{x1} & a_{x2}\end{pmatrix}'$ and $\Psi_{x,t} = \begin{pmatrix} 1 & x_{t-1} & x_{t-2}\end{pmatrix}'$. Agents use a constant gain Recursive Least Squares (RLS) algorithm to update their beliefs

$$\begin{aligned}\Phi_{x,t} &= \Phi_{x,t-1} + \gamma R_{x,t}^{-1}\Psi_{x,t-1}\left(x_{t-1} - \Phi'_{x,t-1}\Psi_{x,t-1}\right)', \quad (56) \\ R_{x,t} &= R_{x,t-1} + \gamma\left(\Psi_{x,t-1}\Psi'_{x,t-1} - R_{x,t-1}\right),\end{aligned}$$

where $\gamma \in (0,1)$ denotes the constant gain learning parameter and $R_{x,t}$ is the precision (3 x 3) matrix associated with each equation. In addition, we assume that agents estimate the persistence parameter ϱ of the exogenous productivity process (12) using constant gain RLS; this captures their uncertainty about the persistence of the productivity process perhaps changing over time.

A.3.1 Computation of Infinite Sums

Since the AR(2) forecasting models are estimated in levels, agents' perceptions of the steady state changes with the arrival of new information. A bar over the variable together with the superscript e denotes the changes in the perception of the steady state at each point in time. Under AR(2), the perceived steady states are

$$\bar{x}^e = \frac{a_{x0}}{1 - a_{x1} - a_{x2}}. \quad \text{for} \quad x = \{n, \pi, w, r\}.$$

Let's denote the deviations of the variable x from the perceived steady state with a tilde and a superscript e.

$$\tilde{x}_t^e = x_t - \bar{x}^e \quad \text{for} \quad x_t = \{n_{t+1}, \pi_t, w_t, r_t\}.$$

Knowing the evolution of the perceived steady states, the belief system can then be written as

$$\tilde{x}_t^e = a_{x1}\tilde{x}_{t-1}^e + a_{x2}\tilde{x}_{t-2}^e \quad \text{for} \quad x_t^e = \left\{n_{t+1}^e, \pi_t^e, w_t^e, r_t^e\right\} \quad \text{and} \quad x = \{n, \pi, w, r\}.$$

Finally, we can express each equation in VAR(1) form, $\tilde{K}_{x,t}^e = \Lambda_x \tilde{K}_{x,t-1}^e$, where $\tilde{K}_{x,t}^e = \begin{pmatrix} \tilde{x}_t^e & \tilde{x}_{t-1}^e \end{pmatrix}'$ and $\Lambda_x = \begin{pmatrix} a_{x1} & a_{x2} \\ 1 & 0 \end{pmatrix}$. Thus, it follows that future forecasts can simply be computed as $\tilde{x}_{t+j}^e = \begin{pmatrix} 1 & 0 \end{pmatrix} \Lambda^{j+1} \tilde{K}_{x,t-1}^e$ for $x = \{n, \pi, w, r\}$ and $\tilde{x}_t^e = \left\{\tilde{n}_{t+1}^e, \tilde{\pi}_t^e, \tilde{w}_t^e, \tilde{r}_t^e\right\}$.

A.3.2 Computational Details of Sums

The infinite sums under AR(2) beliefs are derived using the decomposition $\tilde{x}_t = (x_t - \bar{x}^e) + (\bar{x}^e - \bar{x})$ and $\tilde{x}_{t+j}^e = \begin{pmatrix} 1 & 0 \end{pmatrix} \Lambda_x^{j+1} \tilde{K}_{x,t-1}^e$ for $x = \{n, \pi, w, r\}$ and $\tilde{x}_t^e = \left\{\tilde{n}_{t+1}^e, \tilde{\pi}_t^e, \tilde{w}_t^e, \tilde{r}_t^e\right\}$. $\mathcal{S}_{ft}^z$, $\mathcal{S}_{ft}^w$ and $\mathcal{S}_{ft}^n$ in equation (25) are computed as

$$\mathcal{S}_{ft}^z = \sum_{j=1}^{\infty} (1-\rho)^{j-1} \beta^j \bar{\lambda}_2 E_t^* \tilde{z}_{t+j} = \sum_{j=1}^{\infty} (1-\rho)^{j-1} \beta^j \bar{\lambda}_2 \hat{\varrho}^j \tilde{z}_t = \frac{\beta \hat{\varrho} \bar{\lambda}_2}{1 - (1-\rho)\beta\hat{\varrho}} \tilde{z}_t,$$

$$
\begin{aligned}
\mathcal{S}_{ft}^{w} &= \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j}E_t^{*}\tilde{w}_{t+j} = \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j}\tilde{w}_{t+j}^{e} + \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j}(\bar{w}^{e}-\bar{w}) \\
&= \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j}\tilde{w}_{t+j}^{e} + \frac{\beta(\bar{w}^{e}-\bar{w})}{1-\beta(1-\rho)} \\
&= \sum_{j=1}^{\infty}(1-\rho)^{j-1}\beta^{j}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_w^{j+1}\tilde{K}_{w,t-1}^{e} + \frac{\beta(\bar{w}^{e}-\bar{w})}{1-\beta(1-\rho)} \\
&= \beta\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_w^{2}\left[I-\beta\left(1-\rho\right)\Lambda_w\right]\tilde{K}_{w,t-1}^{e} + \frac{\beta(\bar{w}^{e}-\bar{w})}{1-\beta(1-\rho)},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{S}_{ft}^{n} &= \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j}\bar{\lambda}_1E_t^{*}\tilde{n}_{t+j} = \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j}\tilde{n}_{t+j}^{e} + \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j}(\bar{n}^{e}-\bar{n}) \\
&= \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j}\tilde{n}_{t+j}^{e} + \frac{\bar{\lambda}_1\beta^{2}(1-\rho)(\bar{n}^{e}-\bar{n})}{1-\beta(1-\rho)} \\
&= \sum_{j=2}^{\infty}(1-\rho)^{j-1}\beta^{j}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_n^{j}\tilde{K}_{n,t-1}^{e} + \frac{\bar{\lambda}_1\beta^{2}(1-\rho)(\bar{n}^{e}-\bar{n})}{1-\beta(1-\rho)} \\
&= \beta^{2}(1-\rho)\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_n^{2}\left[I-\beta\left(1-\rho\right)\Lambda_n\right]\tilde{K}_{n,t-1}^{e} + \frac{\bar{\lambda}_1\beta^{2}(1-\rho)(\bar{n}^{e}-\bar{n})}{1-\beta(1-\rho)}.
\end{aligned}
$$

The sum $\mathcal{S}_{ft}^{r}$ contains a double summation operator and it requires using $\sum_{j=1}^{\infty}(1-\rho)^{j}\beta^{j}j = \frac{\beta(1-\rho)}{[1-\beta(1-\rho)]^{2}}$. The computation of $\mathcal{S}_{ft}^{r}$ is thus

$$
\begin{aligned}
\mathcal{S}^{r}_{ft} &= (\bar{z}\bar{\lambda}_2-\bar{w})\beta^2\sum_{j=1}^{\infty}\left[(1-\rho)^j\,\beta^j\sum_{i=1}^{j-1}E^*_t\tilde{r}_{t+i}\right]\\
&= (\bar{z}\bar{\lambda}_2-\bar{w})\beta^2\left\{\sum_{j=1}^{\infty}\left[(1-\rho)^j\,\beta^j j\,(r^e-\bar{r})\right]+\sum_{j=1}^{\infty}\left[(1-\rho)^j\,\beta^j\sum_{i=1}^{j-1}E^*_t\tilde{r}^e_{t+i}\right]\right\}\\
&= (\bar{z}\bar{\lambda}_2-\bar{w})\beta^2\left\{\frac{\beta(1-\rho)\,(r^e-\bar{r})}{[1-\beta(1-\rho)]^2}+\sum_{j=1}^{\infty}\left[(1-\rho)^j\,\beta^j\sum_{i=1}^{j-1}E^*_t\tilde{r}^e_{t+i}\right]\right\}\\
&= (\bar{z}\bar{\lambda}_2-\bar{w})\beta^2\left\{\frac{\beta(1-\rho)\,(r^e-\bar{r})}{[1-\beta(1-\rho)]^2}+\sum_{j=1}^{\infty}\left[(1-\rho)^j\,\beta^j\sum_{i=1}^{j-1}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_r^{j+1}\tilde{K}^e_{r,t-1}\right]\right\}\\
&= (\bar{z}\bar{\lambda}_2-\bar{w})\frac{\beta^3(1-\rho)\,(r^e-\bar{r})}{[1-\beta(1-\rho)]^2}+\\
&\quad (\bar{z}\bar{\lambda}_2-\bar{w})\beta^3(1-\rho)\begin{pmatrix}1 & 0\end{pmatrix}(I-\Lambda_r)^{-1}\Lambda_r^2\sum_{j=1}^{\infty}\left[(1-\rho)^{j-1}\beta^{j-1}\left(I-\Lambda_r^{j-1}\right)\tilde{K}^e_{r,t-1}\right]\\
&= \frac{(\bar{z}\bar{\lambda}_2-\bar{w})\beta^3(1-\rho)\,(r^e-\bar{r})}{[1-\beta(1-\rho)]^2}+\\
&\quad (\bar{z}\bar{\lambda}_2-\bar{w})\beta^3(1-\rho)\begin{pmatrix}1 & 0\end{pmatrix}(I-\Lambda_r)^{-1}\\
&\quad \Lambda_r^2\left[\left(I\,(1-(1-\rho)\,\beta)^{-1}-\Lambda_r\,(I-(1-\rho)\,\beta\Lambda_r)^{-1}\right)\tilde{K}^e_{r,t-1}\right]
\end{aligned}
$$

A similar approach is used to compute the infinite sums in the household's consumption behavioural rule. $\mathcal{S}^z_{ht}$, $\mathcal{S}^w_{ht}$ and $\mathcal{S}^n_{ht}$ in equation (24) are calculated as follows

$$
\begin{aligned}
\mathcal{S}^{w}_{ht} &= \bar{n}\sum_{j=1}^{\infty}\beta^j E^*_t\tilde{w}_{t+j}=\bar{n}\left[\sum_{j=1}^{\infty}\beta^j\tilde{w}^e_{t+j}+\sum_{j=1}^{\infty}\beta^j(\bar{w}^e-\bar{w})\right]\\
&= \bar{n}\left[\sum_{j=1}^{\infty}\beta^j\tilde{w}^e_{t+j}+\frac{\beta(\bar{w}^e-\bar{w})}{1-\beta}\right]=\bar{n}\left[\sum_{j=1}^{\infty}\beta^j\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_w^{j+1}\tilde{K}^e_{w,t-1}+\frac{\beta(\bar{w}^e-\bar{w})}{1-\beta}\right]\\
&= \bar{n}\left[\beta\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_w^2\,(I-\beta\Lambda_w)^{-1}\tilde{K}^e_{w,t-1}+\frac{\beta(\bar{w}^e_t-\bar{w})}{1-\beta}\right],
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{S}^{\pi}_{ht} &= E^*_t\left\{\sum_{j=1}^{\infty}\beta^j\tilde{\pi}_{t+j}\right\}=E^*_t\left\{\sum_{j=1}^{\infty}\beta^j_t\tilde{\pi}^e_{t+j}+\sum_{j=1}^{\infty}\beta^j(\bar{\pi}^e-\bar{\pi})\right\}\\
&= \sum_{j=1}^{\infty}\beta^j\tilde{\pi}^e_{t+j}+\frac{\beta(\bar{\pi}^e-\bar{\pi})}{1-\beta}=\beta\begin{pmatrix}1 & 0\end{pmatrix}\sum_{j=1}^{\infty}\beta^{j-1}\Lambda_\pi^{j+1}\tilde{K}^e_{\pi,t-1}+\frac{\beta(\bar{\pi}^e-\bar{\pi})}{1-\beta}=\\
&= \beta\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_\pi^2\,(I-\beta\Lambda_\pi)^{-1}\tilde{K}^e_{\pi,t-1}+\frac{\beta(\bar{\pi}^e-\bar{\pi})}{1-\beta},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{S}_{ht}^{n} &= \bar{w}\sum_{j=2}^{\infty}\beta^{j}E_{t}^{*}\tilde{n}_{t+j} = \bar{w}\left[\sum_{j=2}^{\infty}\beta^{j}\tilde{n}_{t+j}^{e} + \sum_{j=2}^{\infty}\beta^{j}(\bar{n}^{e}-\bar{n})\right] \\
&= \bar{w}\left[\sum_{j=2}^{\infty}\beta^{j}\tilde{n}_{t+j}^{e} + \frac{\beta^{2}(\bar{n}^{e}-\bar{n})}{1-\beta}\right] \\
&= \bar{w}\left[\sum_{j=2}^{\infty}\beta^{j}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_{n}^{j}\tilde{K}_{n,t-1}^{e} + \frac{\beta^{2}(\bar{n}^{e}-\bar{n})}{1-\beta}\right] = \\
&= \bar{w}\left[\beta^{2}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_{n}^{2}\left(I-\beta\Lambda_{n}\right)^{-1}\tilde{K}_{n,t-1}^{e} + \frac{\beta^{2}(\bar{n}^{e}-\bar{n})}{1-\beta}\right],
\end{aligned}
$$

The sum $\mathcal{S}_{ht}^{r}$ contains a double summation operator and it involves using $\sum_{j=1}^{\infty}\beta^{j}j = \frac{\beta}{(1-\beta)^{2}}$. Thus,

$$
\begin{aligned}
\mathcal{S}_{ht}^{r} &= \left(\bar{w}\bar{n}+\bar{\pi}\right)\beta^{2}\sum_{j=1}^{\infty}\left[\beta^{j}\sum_{i=1}^{j-1}E_{t}^{*}\tilde{r}_{t+i}\right] \\
&= \left(\bar{w}\bar{n}+\bar{\pi}\right)\beta^{2}\left\{\sum_{j=1}^{\infty}\left[\beta^{j}j\left(r^{e}-\bar{r}\right)\right] + \sum_{j=1}^{\infty}\left[\beta^{j}\sum_{i=1}^{j-1}E_{t}^{*}\tilde{r}_{t+i}^{e}\right]\right\} \\
&= \left(\bar{w}\bar{n}+\bar{\pi}\right)\beta^{2}\left\{\frac{\beta\left(r^{e}-\bar{r}\right)}{1-\beta^{2}} + \sum_{j=1}^{\infty}\left[\beta^{j}\sum_{i=1}^{j-1}\begin{pmatrix}1 & 0\end{pmatrix}\Lambda_{r}^{j+1}\tilde{K}_{r,t-1}^{e}\right]\right\} \\
&= \left(\bar{w}\bar{n}+\bar{\pi}\right)\beta^{2}\left\{\frac{\beta\left(r^{e}-\bar{r}\right)}{1-\beta^{2}} + \beta\sum_{j=1}^{\infty}\left[\beta^{j-1}\begin{pmatrix}1 & 0\end{pmatrix}\left(\Lambda_{r}^{2}-\Lambda_{r}^{j+2}\right)\left(I-\Lambda_{r}\right)^{-1}\tilde{K}_{r,t-1}^{e}\right]\right\} \\
&= \left(\bar{w}\bar{n}+\bar{\pi}\right)\beta^{3}\left\{\frac{\left(r^{e}-\bar{r}\right)}{1-\beta^{2}} + \begin{pmatrix}1 & 0\end{pmatrix}\left(I-\Lambda_{r}\right)^{-1}\Lambda_{r}^{2}\left[I\left(1-\beta\right)^{-1}-\Lambda_{r}\left(I-\beta\Lambda_{r}\right)^{-1}\right]\tilde{K}_{r,t-1}^{e}\right\}.
\end{aligned}
$$

B Additional Tables

Recall that the difference between the RE model (36) and the AL model with correctly specified beliefs (37) is that, whilst under RE agents know the parameters of the RE solution, under learning agents do not know the value of the parameters and update them as new information becomes available. Tables 8 and 9 below show that this model suffers from the same problem as the RE model i.e. it is not able to provide a solution to the unemployment volatility puzzle and unable to match the forecast error properties in the data.

Model	Statistics		$\hat{y}_t$	$\hat{n}_t$	$\hat{v}_t$	$\hat{u}_t$	$\hat{\theta}_t$	$\hat{c}_t$
RE	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$		1	0.03	0.50	0.43	0.91	1.00
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.96	0.97	0.92	0.97	0.96	0.96
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	0.99	0.99	−0.99	1.00	1.00
		$\hat{n}_t$	–	1	0.96	−1.00	0.99	0.99
		$\hat{v}_t$	–	–	1	−0.96	0.99	0.99
		$\hat{u}_t$	–	–	–	1	−0.99	−0.99
		$\hat{\theta}_t$	–	–	–	–	1	1.00
		$\hat{c}_t$	–	–	–	–	–	1
AL with correctly specified beliefs	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$		1	0.03	0.50	0.43	0.92	1.00
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{1t-1}\right)$		0.96	0.97	0.92	0.97	0.96	0.96
	$\rho\left(\widehat{x}_{1t}, \widehat{x}_{2t}\right)$	$\hat{y}_t$	1	1.00	0.99	−0.99	1.00	1.00
		$\hat{n}_t$	–	1	0.98	−1.00	1.00	1.00
		$\hat{v}_t$	–	–	1	−0.96	0.99	0.99
		$\hat{u}_t$	–	–	–	1	−0.99	−0.99
		$\hat{\theta}_t$	–	–	–	–	1	1.00
		$\hat{c}_t$	–	–	–	–	–	1

Table 8: Summary Statistics: RE and Correctly Specified Beliefs

Notes. Relative standard deviations, autocorrelation and correlation coefficients in this Table correspond to the quaterly simulated series expressed in percentage deviations from steady state value. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2.

		FE_t^{Q1}	FE_t^{Q2}	FE_t^{Q3}	FE_t^{Q4}
RE	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	0.05	0.12	0.18	0.21
	$\rho\left(\hat{x}_{1t}, \hat{x}_{1t-1}\right)$	0.13	0.24	0.61	0.73
	$\rho\left(\hat{x}_{1t}, \Delta\hat{u}_t\right)$	−0.10	−0.04	−0.03	0.00
	$\rho\left(\hat{x}_{1t}, \Delta\hat{y}_t\right)$	0.27	0.00	0.04	0.00
AL with correctly specified beliefs	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	0.04	0.12	0.18	0.21
	$\rho\left(\hat{x}_{1t}, \hat{x}_{1t-1}\right)$	−0.06	0.24	0.61	0.73
	$\rho\left(\hat{x}_{1t}, \Delta\hat{u}_t\right)$	−0.01	−0.04	−0.04	−0.02
	$\rho\left(\hat{x}_{1t}, \Delta\hat{y}_t\right)$	−0.02	0.00	0.04	0.01

Table 9: Forecast Properties of RE and Correctly Specified Beliefs

Notes. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2. Data from Survey of Professional Forecasters. Forecast errors FE_t^{Qj} for $j = \{1, 2, 3, 4\}$ as defined in Table 5

The table below reports on the forecast error properties for the robustness exercise in Section 6..

		FE_t^{Q1}	FE_t^{Q2}	FE_t^{Q3}	FE_t^{Q4}
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	4.56	6.21	8.02	9.89
Data	$\rho\left(\hat{x}_{1t},\hat{x}_{1t-1}\right)$	0.70	0.79	0.85	0.89
	$\rho\left(\hat{x}_{1t},\Delta\hat{u}_t\right)$	-0.69	-0.55	-0.47	-0.36
	$\rho\left(\hat{x}_{t1},\Delta\hat{y}_t\right)$	0.44	0.42	0.37	0.32
Alternative parameterisations					
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	3.08	4.64	5.65	6.23
AR(2) ($\gamma = 0.001$)	$\rho\left(\hat{x}_{1t},\hat{x}_{1t-1}\right)$	0.48	0.82	0.89	0.92
	$\rho\left(\hat{x}_{1t},\Delta\hat{u}_t\right)$	-0.08	-0.16	-0.15	-0.17
	$\rho\left(\hat{x}_{t1},\Delta\hat{y}_t\right)$	0.15	0.21	0.20	0.22
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	3.33	4.65	5.40	5.96
AR(2) ($\gamma = 0.003$)	$\rho\left(\hat{x}_{1t},\hat{x}_{1t-1}\right)$	0.19	0.58	0.75	0.81
	$\rho\left(\hat{x}_{1t},\Delta\hat{u}_t\right)$	-0.15	-0.11	-0.12	-0.12
	$\rho\left(\hat{x}_{t1},\Delta\hat{y}_t\right)$	0.20	0.18	0.16	0.16
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	3.49	4.65	5.37	5.86
AR(2) ($\alpha = 0.7$)	$\rho\left(\hat{x}_{1t},\hat{x}_{1t-1}\right)$	0.18	0.57	0.72	0.79
	$\rho\left(\hat{x}_{1t},\Delta\hat{u}_t\right)$	-0.11	-0.10	-0.11	-0.08
	$\rho\left(\hat{x}_{t1},\Delta\hat{y}_t\right)$	0.15	0.13	0.13	0.11
	$\sigma_{\hat{x}_{1t}}/\sigma_{\hat{y}_t}$	4.53	5.78	6.57	7.06
AR(2) ($\sigma = 0.6$ and $\xi = 0.4$)	$\rho\left(\hat{x}_{1t},\hat{x}_{1t-1}\right)$	0.19	0.61	0.74	0.79
	$\rho\left(\hat{x}_{1t},\Delta\hat{u}_t\right)$	0.00	-0.07	-0.10	-0.11
	$\rho\left(\hat{x}_{t1},\Delta\hat{y}_t\right)$	0.04	0.10	0.12	0.14

Table 10: Forecast Properties under Alternative Parameterisations

Notes. Data from Survey of Professional Forecasters. Relative standard deviations in this Table correspond to the quarterly simulated series expressed in percentage deviations from steady state values under different parameterisations. The term $\rho(x_1, x_2)$ stands for the correlation coefficient between variables x_1 and x_2. Forecast errors FE_t^{Qj} for $j = \{1, 2, 3, 4\}$ as defined in Table 5.