Unemployment and econometric learning

Singleton, Carl and Schaefer, Daniel

University of Edinburgh, University of Edinburgh

19 February 2015

Unemployment and econometric learning*

Daniel Schäfer[†]
University of Edinburgh

Carl A. Singleton[‡]
University of Edinburgh

Latest version: 19th February 2015

Abstract

We apply well-known results of the econometric learning literature to a standard RBC model with unemployment. The unique REE is always expectationally stable with decreasing gain learning, and this result is robust to over-parametrisation of the econometric model relative to the minimum state variable form used by agents (Strong E-stability). And so, from this perspective, the assumption of rational expectations in the Mortensen-Pissarides is not unreasonable. Using a parametrisation with UK data, simulations suggest that the implied rate of convergence to the rational expectations equilibrium (REE) with least squares learning is however slow. The cyclical response of unemployment to structural shocks is muted under learning, and a parametrisation which guarantees root-t convergence is generally not consistent with attempts to match the observed volatility of labour market data using the standard model.

Key words: real business cycle, unemployment, adaptive learning, expectational stability

JEL Codes: E24, E32, J64, D83

1 Introduction

The Mortensen & Pissarides (1994) model of search frictions has become the foundation for the investigation of the functioning of labour markets (see Rogerson & Shimer, 2011 for a survey). The existence of search frictions in labour markets is usually motivated by the observation that the market is decentralised due to geographical and sectoral differences in firms, and the fact that each worker has distinct features which make them more or less suitable for these jobs. That is, there is fundamental heterogeneity present in labour markets. Whilst searching for a match, unemployed workers and firms have to form expectations regarding future variables, which are relevant for their decision making process, such as the unemployment rate and the quantity of open vacancies. The decentralised nature of labour markets makes it a priori *not* obvious that workers and firms are able to correctly quantify the values of these

*We would like to thank participants at the 2014 SGPE Residential Methodology conference for comments on an earlier version.

[†]daniel.schaefer@.ed.ac.uk; add specific thanks to ESRC/ department etc.

[‡]c.a.singleton@sms.ed.ac.uk; with thanks to the Economic and Social Research Council for funding support

relevant variables at all times. Nevertheless, the assumption that firms and unemployed workers have rational expectations is usually made in search and matching models, which implies that expectations of the unemployment rate and vacancies reflect true values. This is likely to pose overly strong requirements on the cognitive abilities of economic agents and the implications of relaxing the assumption of rational expectations should be investigated.

Here we analyse the equilibrium properties and dynamics of the Mortensen-Pissarides model whilst representing agents as 'good econometricians,' who form forecasts according to their estimates of structural model parameters (Evans & Honkapohja, 2001). We follow the literature on econometric learning and assume that agents employ a recursive least squares algorithm to update their parameter estimates each period when new data becomes available. It can be argued that for agents in a labour market, a priori knowledge at all times of the true law of motion of the economy, determined by general equilibrium effects, is unrealistic in the presence of structural shocks, and small departures from the rational expectation assumption might alter qualitative or quantitative predictions. Econometric learning is able to provide a behavioural foundation for the assumption of rational expectations if the rational expectations equilibrium (REE) is shown to be learnable or E-stable (Evans & Honkapohja, 2001); i.e. small deviations from this equilibrium are reversed over time, and the economy and agents' parameter estimates converge asymptotically to the true equilibrium values.

Restricting our attention here to a discrete time treatment of the textbook RBC Mortensen-Pissarides model (see Hagedorn & Manovskii, 2008), we derive the conditions under which its REE is E-stable. We find that there are no parameter restrictions required beyond those made in the standard model to guarantee this. Furthermore, we confirm that the model is in fact globally stable and satisfies the properties of Strong E-stability (i.e. is robust to over-parametrisation of the econometric relationship by the agents). And so from this viewpoint, the assumption of rational expectations when studying or applying this model is not unreasonable.

There exists a significant literature on the representation of the good econometrician in macroeconomic models. Mankiw et al. (2004) offer evidence against the rational expectations hypothesis. Their analysis of surveys of professional forecasters and households containing expected inflation in one year's time shows significant autocorrelation in the forecast errors, which is compatible with econometric learning, but not with rational expectations. Milani (2007, 2011) also argues for the presence of adaptive learning in DSGE models, using Bayesian methods to estimate a New Keynesian model that nests learning with habit formation, indexation and other 'mechanical' ways of generating persistence. His results show that learning is capable of replacing the other sources of persistence whilst at the same time increasing the fit to the data as compared to rational expectations. Pfajfar & Santoro (2010) examine a survey of households' inflation forecasts during 1978-2005 and conclude that the hypothesis of the presence of rational expectations can be rejected, and that there is evidence in support of adaptive learning dynamics. This view is further supported by Berardi & Galimberti (2012), who examine post-WWII data of US inflation and output growth. Comparing the performance of different adaptive learning algorithms to match survey forecasts, their results suggest that economic agents form these according to recursive least squares.

Our formulation of learning here is such that agents need only make one-step-ahead forecasts of the labour market tightness using their estimated model parameters and observed productivity. This is not the first paper to apply the principles of econometric learning to this class of model. From the same model starting point, Di Pace et al. (2014) consider an alternate formulation in which agents must make infinite horizon forecasts about the future path of wages, unemployment and profits. This type of learning fits into the anticipated utility approach (Kreps, 1998), and has notably been applied to the RBC model by Eusepi & Preston (2011). Di Pace et al. (2014) focus on results with constant gain learning, for which there are no equivalent analytical results to the E-stability conditions we consider. The authors use the model to address the so-called 'unemployment volatility puzzle' (Shimer, 2005). Under infinite horizon, they not only match US forecast errors for unemployment, but also generate significant amplification relative to the baseline model. This is driven by persistence or inertia in agents' expectations of the future path of wages, which thus implies firms are over-optimistic about future profits, post more vacancies, and thus unemployment is more volatile relative to the REE baseline case. They also find some, but significantly less, propagation of unemployment when they reformulate the model in a one-step-ahead forecast guise.

We present illustrative simulations and analyse the dynamics of the unemployment model with econometric learning, and show that the REE could be a poor approximation to an economy in which agents play the role of good econometricians in terms of levels. Convergence to the REE is slow when we give agents only 50 quarters of historical data, even after we allow for weighted least squares learning. We also show that structural shocks, when agents are learning but initially begin at the REE, generate a more gradual 'cyclical' adjustment of wages, and thus unemployment volatility would be reduced relative to under rational expectations. Therefore, we find conflicting results to Di Pace et al. (2014). Given that we derive a minimum state variable form of the model, in which agents need only estimate the relationship between labour market tightness and productivity to form their expectations and close the model, the 'wage curve' is already built-in. Thus, econometric learning generates inertia in expectations of the labour market state and wages following shocks.[1] And hence the same amplification mechanism described in Di Pace et al. (2014) is not present. Given the only true choice variable in the model is the number of vacancies that firms open or close each period, our application of learning appears to be a natural way to extend the baseline model.

We consider in more detail the speed of convergence to the REE of the model. The same parameters and magnitudes which reduce the speed of convergence substantially are the same as those which in the REE model also effectively increase the model's ability to match unemployment volatility: a relatively high flow value of unemployment, low worker bargaining power and low matching function elasticity

[1] Although we do not expand on this point later, it is straightforward to see that the minimum state variable solution of the model, which agents learn and use to form expectations, could be re-written in terms of wages and productivity by substituting for the standard 'wage curve' which derives from the Nash Bargaining solution.

w.r.t. unemployment (see Hagedorn & Manovskii (2008) for a detailed discussion of these). Therefore, in order to match the observed volatility of unemployment etc., more extreme values of these parameters would be required when we relax rational expectations in this limited way.

The rest of the paper is structured as follows: in section 2 we add learning to the textbook search and match model of unemployment and derive key expectational stability results; in section 3 we discuss a simple parametrisation of the model and results of stochastic recursive simulations with learning relative to the REE; and in section 4 we summarise.

2 The Model

We briefly outline a standard textbook search and match model of the labour market (Pissarides, 2000) and the resultant discrete time dynamic equation for the state of the labour market, measured by its tightness, thus analogous to the treatment in Hagedorn & Manovskii (2008).

2.1 The labour market

There is a continuum of identical, risk neutral workers with total measure of one, and an infinite horizon. The matching function $M(u_t, v_t)$ gives the number of successful matches in a given period. It is increasing and concave in both of its arguments u_t and v_t, which represent the share of the total workforce currently unemployed and the level of vacancies relative to the size of the workforce respectively. Matches and separations occur after agents in the economy have made decisions, at the end of period t. A Cobb-Douglas, constant returns to scale matching function is chosen due to its simplicity, well-known features, and being commonplace in the literature;

$$M(u_t, v_t) = \mu u_t^{\alpha} v_t^{1-\alpha} \qquad \mu > 0 \qquad \alpha \in (0,1). \tag{1}$$

For the following exposition we define the key parameter, the measure of labour market tightness, as

$$\theta_t = \frac{v_t}{u_t}. \tag{2}$$

The probability of a firm filling an open vacancy is

$$q(\theta_t) = \frac{M(u_t, v_t)}{v_t} = \mu \theta_t^{-\alpha}, \tag{3}$$

and the corresponding probability that an unemployed worker gets matched to an open vacancy is thus

$$\theta_t q(\theta_t) = \frac{M(u_t, v_t)}{u_t} = \mu \theta_t^{1-\alpha}. \tag{4}$$

The law of motion for the fraction of workers who are unemployed at the beginning of period $t+1$ is given by

$$u_{t+1} = u_t + (1-u_t)\lambda - \theta_t q(\theta_t) u_t. \tag{5}$$

2.2 The household

For ex-positional simplicity we consider an economy comprising a single representative household of size one, in which all workers are identical, risk neutral and there is perfect consumption insurance across members. The household can therefore be represented by the Bellman equation

$$W_t(n,y) = w_t n_t + D_t + b(1-n_t) + \delta W^e_{t+1}(n,y), \tag{6}$$

with the law of motion for employment given by

$$n_{t+1} = (1-\lambda)n_t + \theta_t q(\theta_t)(1-n_t), \tag{7}$$

where $W_t(n,y)$ represents its current value function, labour productivity is denoted by y_t, and e shows time t subjective expectations. The household takes as given wages w_t, dividends from the representative firm D_t and labour market tightness θ_t. The period utility value from non-employment is given by b, and δ is the discount factor. Applying the envelope theorem, the surplus from an additional member of the household being employed is given by

$$\frac{\partial W_t(n,y)}{\partial n_t} = w_t - b + \delta(1-\lambda-\theta_t q(\theta_t))\frac{\partial W^e_{t+1}(n,y)}{\partial n_{t+1}}, \tag{8}$$

i.e. the net value of employment plus the expected continuation value.

2.3 The firm

A representative firm with linear production function maximises profits by choosing the quantity of vacancies to post in each period, at constant ongoing cost c, subject to the law of motion for the labour market, and taking w_t and θ_t as given. This problem can be represented as the Bellman equation

$$\Pi_t(v;,n,y) = \max_{v_t \geq 0} \quad y_t n_t - w_t n_t - c v_t + \delta \Pi^e_{t+1}(v;,n,y), \tag{9}$$

with the constraint

$$n_{t+1} = (1-\lambda)n_t + q(\theta_t)v_t, \tag{10}$$

$\Pi_t(v;,n,y)$ represents its current value function. The first order condition is given by

$$\frac{\partial \Pi^e_{t+1}(v;n,y)}{\partial n_{t+1}} = \frac{c}{\delta q(\theta_t)}. \tag{11}$$

Applying the envelope theorem, and using the first order condition gives the surplus to the firm from employing an additional worker

$$\frac{\partial \Pi_t(v;n,y)}{\partial n_t} = y_t - w_t + \frac{(1-\lambda)c}{q(\theta_t)}, \tag{12}$$

i.e. the net profit from an additional employed worker plus the expected continuation value.

2.4 Wage determination

Wages are determined by generalised Nash bargaining between firm and workers over the additional surpluses (8) & (12), with worker bargaining power $\beta \in (0,1)$. Combining the surplus sharing rules which form the solution of this problem, iterating forwards, and using (8) (12) & (11) gives what is referred to in textbook model as the 'wage curve',

$$w_t = (1-\beta)b + \beta(y_t + c\theta_t). \tag{13}$$

2.5 REE and dynamics

The equilibrium of the model can be characterised and determined uniquely by the value of labour market tightness at which point the representative firm is indifferent between opening an additional vacancy or not, due to the assumption of free entry. The non-linear equation determining this value of θ_t can be obtained by iterating forwards and taking expectations of the envelope theorem condition with respect to employment for the firm, and then using the firm's first order condition (11) and the wage curve equation (13);

$$\frac{c}{\delta q(\theta_t)} = (1-\beta)(y^e_{t+1} - b) + \frac{(1-\lambda)c}{q(\theta^e_{t+1})} - \theta^e_{t+1}\beta c. \tag{14}$$

To provide intuition for this expression, we stress the similarities to (11). Aggregate labour market tightness θ_t will adjust immediately to deviations from this equality via the representative firm instantaneously opening or closing vacancies. Thus, today's labour market tightness is determined by expectations of the value of a filled vacancy in the next period. However, this value is itself a function of the labour market tightness (and other variables) in the next period due to the infinite nature of the Bellman equations, and thus θ_t is a function of θ^e_{t+1}.

We assume that the process of worker productivity takes the form of a stationary AR(1) in logs

$$\log(y_t) = \rho \log(y_{t-1}) + \varepsilon_t, \qquad \rho \in [0,1) \tag{15}$$

with ε_t being an i.i.d. normal shock with variance $\sigma^2_\varepsilon > 0$.

Once agents have formed expectations for θ_{t+1} and y_{t+1}, the remaining endogenous variables of interest such as wages and unemployment follow mechanically using the model equations derived above. To solve the system consisting of (14) and (15) we linearise around the steady state values $\bar{\theta}$ and $\bar{y} = 1$ (see appendix for derivation);

$$\theta_t = \psi_0 + \psi_1 y^e_{t+1} + \psi_2 \theta^e_{t+1} \tag{16}$$

$$y_t = (1-\rho) + \rho y_{t-1} + \varepsilon_t. \tag{17}$$

In the following, the operator $E_t x$ denotes mathematical expectations formed at period t regarding variable x. The linearised dynamics of output (17) can be substituted into (16) for y^e_{t+1} by noting that under rational expectations $y^e_{t+1} = E_t y_{t+1} = (1-\rho) + \rho y_t$ and hence

$$\theta_t = \tilde{\psi}_0 + \tilde{\psi}_1 y_{t-1} + \tilde{\psi}_2 \theta^e_{t+1} + \tilde{\psi}_1 \rho^{-1} \varepsilon_t \tag{18}$$

with

$$\begin{aligned}
\tilde{\psi}_0 &= \psi_0 + \psi_1(1-\rho)(1+\rho) \\
\tilde{\psi}_1 &= \psi_1 \rho^2 \\
\tilde{\psi}_2 &= \psi_2.
\end{aligned}$$

A rational expectation equilibrium (REE) of the system (17) and (18) is a stochastic process θ_t that satisfies this system with $E_t \theta_{t+1} = \theta^e_{t+1}$. As is well known, linear RE models in which agents form expectations regarding an endogenous variable can possess multiple equilibria (or *bubble* solutions, i.e. which are not related to economic fundamentals)[2]. If there are multiple *stable* rational expectations equilibria then a model is said to be indeterminate. To see the possibility of this, note that (18) can be written in ARMA(1,1) form as

$$\theta_t = \tilde{\psi}_2^{-1}\left(\rho^{-1}\tilde{\psi}_1(1-\rho) - \tilde{\psi}_0\right) - \tilde{\psi}_1 \tilde{\psi}_2^{-1} \rho^{-1} y_{t-1} + \tilde{\psi}_2^{-1} \theta_{t-1} + d_1 \varepsilon_t + d_2 \eta_t + \tilde{\psi}_1 \rho^{-1} \left(\tilde{\psi}_2^{-1} - 1\right) \varepsilon_{t-1} \tag{19}$$

with d_1 and d_2 arbitrary parameters and $\eta_t := E_{t+1}[\theta_{t+2}] - E_t[\theta_{t+2}]$ being a martingale difference sequence such that $E_t[\eta_{t+1}] = 0$. No restrictions are imposed on d_1 or d_2 since rational expectations formed according to (19) always result in an identity regardless of these parameters. Therefore, there is a continuum of possible solutions to (19). Evans & Honkapohja (1986) have shown that any finite degree ARMA solution of an equation of the form (18) can at most be ARMA(1,1) and the particular form of (19) nests all possible ARMA solutions of finite degree. However, the ARMA class of solutions is stable if $|\tilde{\psi}_2| > 1$ and is unstable for $|\tilde{\psi}_2| < 1$, in which case the unique stationary solution to (17) and (18) is the fundamental or minimal state variable (MSV) solution. The MSV solution depends on a minimal set of state variables such that it is impossible to delete any of the variables from the set and still obtain solutions to (18) and (15) for all permitted parameter values (McCallum, 1983). The MSV solution can be guessed to be of the form

$$\theta_t = A + B y_{t-1} + C \varepsilon_t \tag{20}$$

since the only predetermined variable in the structural equation (18) is y_{t-1}. In the Appendix it is shown how (20) can be obtained from (19). We first focus on this solution, since inter alia McCallum (1983) and Evans & Honkapohja (2001) argue that the MSV solution is the natural benchmark for learning and the stability of equilibria. If agents are not able to learn the simplest representation (as few state variables as possible), they cannot be expected to be able to learn equilibria containing more state variables and to coordinate their behaviour towards them.

[2]The literature on bubbles and the related concept of indeterminacy is reviewed in Benhabib & Farmer (1999). The classic reference on bubble solutions is Blanchard & Watson (1983); see also Bullard & Mitra (2002) for a recent analysis of indeterminacy in a New Keynesian framework.

2.6 E-Stability of the MSV Solution

We now relax the assumption of rational expectations by modelling agents as econometricians attempting to estimate the parameters underpinning the true motion of the economy under uncertainty. Therefore, agents are endowed with a perceived law of motion (PLM) of the economy of the MSV form, (20). The task for agents is then to estimate the values of the parameters A, B and C. They form estimates in period t given by $\hat{A}_t$, $\hat{B}_t$ and $\hat{C}_t$, and update these estimates every period when new data becomes available. We assume agents perform this updating and estimation task by employing the recursive least squares estimator. This is the most widely used estimation technique in the learning literature and Berardi & Galimberti (2012) provide evidence that this estimator matches the observed survey forecasts of US time series closely.[3] According to the PLM, the forecast of next period's labour market tightness is

$$\theta^e_{t+1} = \hat{A}_t + \hat{B}_t\left((1-\rho)+\rho y_{t-1}+\varepsilon_t\right). \tag{21}$$

There are potential problems arising due to possible simultaneity in forward looking models, (Evans & Honkapohja, 2001). Therefore, it is assumed that although agents forecast θ_{t+1} by using y_t, the variable y_t is *not* in the information set for the estimation of $\hat{A}_t$ and $\hat{B}_t$. However, as proved by Marcet & Sargent (1989), this does not alter the asymptotic stability results obtained in the following as compared to an algorithm allowing for simultaneity if agents are assumed to ignore outliers, defined as being an observation outside a predetermined range. Inserting the econometric forecast into (18) gives the actual law of motion (ALM) for labour market tightness

$$\theta_t = \tilde{\psi}_0 + \tilde{\psi}_2\hat{A}_t + \tilde{\psi}_2(1-\rho)\hat{B}_t + (\tilde{\psi}_1 + \tilde{\psi}_2\rho\hat{B}_t)y_{t-1} + (\tilde{\psi}_2\hat{B}_t + \tilde{\psi}_1\rho^{-1})\varepsilon_t. \tag{22}$$

This defines the following T-mapping from the PLM to the ALM

$$T\left(\hat{A}_t,\hat{B}_t,\hat{C}_t\right) = \left(\tilde{\psi}_0 + \tilde{\psi}_2\hat{A}_t + \tilde{\psi}_2(1-\rho)\hat{B}_t, \tilde{\psi}_1 + \tilde{\psi}_2\rho\hat{B}_t, \tilde{\psi}_2\hat{B}_t + \tilde{\psi}_1\rho^{-1}\right). \tag{23}$$

It is apparent that the T-mapping to $\hat{C}_t$ is determined by the other coefficients, and thus the estimate $\hat{C}_t$ is independent of C, and does not influence the stability results. Therefore, in what follows we refer to the mappings $T\left(\hat{A}_t,\hat{B}_t\right)$, and for $\hat{C}_t$, $V(\hat{B}_t)$. There is a self-referential feature inherent in all learning models which can be seen in equation (22). Although the parameters of their estimation (20) are non-stationary during the transition to their steady state values, agents neglect this fact whilst forming their estimates since a least squares method assumes the 'true' A, B and C to be constants. Intuitively, if the coefficient which determines the responsiveness to expectations is sufficiently small, this specification error becomes asymptotically negligible and the economy converges to the REE (Evans & Honkapohja, 2001).

The estimation takes place via recursive least squares (RLS). Let $z'_{t-1} = (1, y_{t-1})$ and $x'_{t-1} = (\hat{A}_t, \hat{B}_t)$ so that

$$\theta_t = z'_{t-1}x_{t-1} + \eta_t. \tag{24}$$

[3] The presented algorithm is comparable to a restricted form of the well-known Kalman filter. For further discussion of this issue see Berardi & Galimberti (2013).

The estimation error η_t is perceived by the agents to be *iid*, however due to the self-referential nature of the model there is an endogeneity bias which agents are unaware of and thus η_t is *not* truly *iid*. We now define $S_t = \sum_{i=1}^{t} z_{i-1}z'_{i-1}$ which allows us to write the RLS estimator as

$$S_t = S_{t-1} + z_{t-1}z'_{t-1} \tag{25}$$

$$x_t = x_{t-1} + S_{t-1}^{-1}x_{t-1}(\theta_t - z'_{t-1}x_{t-1}). \tag{26}$$

Furthermore, let $R_t \equiv t^{-1}S_t$:

$$R_t = R_{t-1} + t^{-1}(z_{t-1}z'_{t-1} - R_{t-1}) \tag{27}$$

$$x_t = x_{t-1} + t^{-1}R_t^{-1}z_{t-1}(\theta_t - z'_{t-1}x_{t-1}), \tag{28}$$

and thus

$$x_t = x_{t-1} + t^{-1}R_t^{-1}z_{t-1}\left\{z'_{t-1}\left[T\left(\hat{A}_t, \hat{B}_t\right) - x_{t-1}\right] + V(\hat{B}_t)\varepsilon_t\right\}, \tag{29}$$

with the gain sequence $1/t$, often referred to as decreasing gain learning.[4] This gain guarantees that asymptotically new information is disregarded by agents.

For the rational expectations equilibrium to be asymptotically stable under learning, the E-stability principle has to be fulfilled as shown by Evans & Honkapohja (2001). According to this principle, the stability of the system in (27) and (29) with decreasing gain is governed by the following ordinary differential equation (ODE), where τ denotes 'notional' time

$$\frac{d}{d\tau}\left(\hat{A}, \hat{B}\right) = T\left(\hat{A}, \hat{B}\right) - \left(\hat{A}, \hat{B}\right). \tag{30}$$

The REE is then the unique fixed point of these differential equations. With this specification, it can be shown that the REE of (18) is stable under the learning, i.e. small deviations from the steady state value $\bar{\theta}$ are reversed, if the sufficient condition

$$\tilde{\psi}_2 < 1 \tag{31}$$

holds. The necessary condition for E-stability is $\tilde{\psi}_2\rho < 1$. In the Appendix we show that the sufficient condition always holds for this model, and we have global convergence to the REE. As was explained in the previous section, the model is determinate if $|\tilde{\psi}_2| < 1$. We therefore can state the following:

Proposition 2.1. *If the economy described by the system* (17) *and* (18) *exhibits determinacy and the PLM is of the MSV form, and if agents learn using least squares updating, then so long as* $\tilde{\psi}_2 < 1$ *the unique REE is E-stable.*

Corollary 2.2. *The textbook equilibrium model of labour market search and match frictions, with homogeneous agents and no on the job search (Pissarides, 2000: Chapter 1), is E-stable.*

[4] In the case of constant gain learning the weight given each observation is geometrically declining in the time since it was observed, and the gain sequence would be $0 < \gamma < 1$.

2.7 Strong E-stability of the MSV Solution

Strong E-stability of a system is defined if the previous result is robust to over-parametrisation of the PLM (Evans & Honkapohja, 2001). Assume that agents are forming their expectations of θ_{t+1} according to the general ARMA representation (19), and are not endowed with a PLM of the MSV form. Moreover, due to econometric considerations they start with an arbitrarily over-parametrised version

$$\theta_t = a + \sum_{j=1}^{s} b_j y_{t-j} + \sum_{j=1}^{r} c_j \theta_{t-j} + \sum_{j=1}^{q} d_j \varepsilon_{t-j} + \sum_{j=1}^{l} f_j \eta_{t-j} + d_0 \varepsilon_t + f_0 \eta_t. \tag{32}$$

Accordingly, expectations of θ_{t+1} take the form

$$\theta^e_{t+1} = a + \sum_{j=1}^{s} b_j y_{t+1-j} + \sum_{j=1}^{r} c_j \theta_{t+1-j} + \sum_{j=1}^{q} d_j \varepsilon_{t+1-j} + \sum_{j=1}^{l} f_j \eta_{t+1-j}, \tag{33}$$

which can be substituted into equation (18) to obtain the new ALM and a corresponding T-mapping in exactly the same way as before (see Appendix). Let $\mathbf{b}' = (b_1, ..., b_s)$, $\mathbf{c}' = (c_1, ..., c_r)$, $\mathbf{d}' = (d_0, ..., d_q)$, and also $\mathbf{f}' = (f_0, ..., f_l)$. Further, define $\phi' = (a, \mathbf{b}', \mathbf{c}', \mathbf{d}', \mathbf{f}')$. According to the E-stability principle, the ODE governing the stability of the above system is again given by

$$\frac{d\phi}{d\tau} = T(\phi) - \phi. \tag{34}$$

To investigate whether agents will detect the over-parametrisation and will converge towards the MSV solution, the stability of the ODE at the REE represented by the MSV solution, given in (47), must be compared. Notice that in this case $\bar{a} = A$, $\bar{b}_1 = B$ and $\bar{d}_0 = C$ is given, and all other coefficients of (32) are equal to zero. In the appendix we show the following:

Proposition 2.3. *If the economy described by the system* (17) *and* (18) *exhibits determinacy and the PLM is of the over-parametrised ARMA form, and if agents learn using least squares updating, then so long as* $\tilde{\psi}_2 < 1$ *the unique REE is E-stable.*

Corollary 2.4. *The textbook equilibrium model of labour market search and match frictions, with homogeneous agents and no on the job search (Pissarides, 2000: Chapter 1), is Strongly E-stable.*

3 Analysis

In the following section we present a brief analysis of the unemployment model with econometric learning described above. We consider two illustrative simulations to demonstrate the implied speed of convergence and dynamics of the endogenous variables in the standard search model with learning. First, we demonstrate e-stability when starting 'realistically' far away from the REE. Second, with agents initially having assumed to have learned the RE equilibrium, we consider the impact of a shift implied by an arbitrary change in some estimated parameter value. We then describe how learning affects the dynamics of the model.

3.1 Simulations

Since this is not a strict empirical exercise, we follow a relatively ad hoc and illustrative only parametrisation strategy using seasonally adjusted UK quarterly data for the period 1998-2013.[5]

We normalise average productivity to be 1. For the productivity process We estimate an AR(1) in log deviations from trend[6] of output per worker, and find an auto-regressive parameter ρ for the period of 0.84 and standard deviation of the shocks σ_ε to be 0.0063 (assuming them to be normally distributed). We estimate the aggregate matching function below using OLS as per Şahin et al. (2012) for 2002q1-2013q2.[7] As is common in the literature, we approximate a two state labour market simply by ignoring those who are inactive. We proxy the number of matches, or new hires, as the total inflow into employment m_t in any quarter, and use the total number of vacancies in the economy.

$$\log\left(\frac{m_t}{u_t}\right) = \log(\mu) + (1-\alpha)\log\left(\frac{v_t}{u_t}\right) + \zeta_t. \tag{35}$$

We find estimates of $\alpha = 0.39$ and $\mu = 1.08$. For the exogenous separation we use the UK Labour Force Survey two quarter hazard rate estimate for leaving employment, which for 2002q1-2013q2 gives a quarterly average rate of $\lambda = 0.035$. Brief summary statistics of the key model quarterly stock and rate variables for the UK consistent with the parametrisation method here, are described in table 1.[8] The discount factor is set as $\delta = 0.99$ and to restrict the number of free parameters, we let the bargaining power match the Hosios condition, $\beta = \alpha = 0.39$. We also set the flow value of unemployment to 0.8.[9] The remaining parameter, the flow vacancy cost c is set to match the observed level of average labour market tightness. The full list of parameters and implied values of the endogenous variables for the deterministic equilibrium are given below in table 2.

[5] All data used from ONS, downloaded 01/08/2014. Labour market data for those aged 16-64 only. For a more complete calibration of the unemployment model using UK data see Burgess & Turon (2010)

[6] dynamically detrended using Hodrick & Prescott (1997) filter with standard quarterly smoothing parameter

[7] Also following Borowczyk-Martins et al. (2013) we consider time trends in the estimation to account for the enodogeneity of unobserved shifts in the matching efficiency with the number of vacancies firms open, but these all drop out. We also carry out tests that the matching function is Cobb-Douglas, and reject the alternative. In line with the existing literature, we find that the data suggests the matching function has decreasing returns to scale, although we proceed as though it is constant; see Petrongolo & Pissarides (2001) for a thorough review of estimates of the aggregate matching function.

[8] All regressed on cubic trend to account for low frequency shifts for the short period in question

[9] How to select or estimate appropriate values of both the bargaining power and the flow value of unemployment are open to significant academic debate. Shimer (2005) and subsequently Hagedorn & Manovskii (2008) are often considered in the literature as the extreme examples for parametrisations, and highlight how this greatly affects the ability of the model to match the observed volatility of unemployment and vacancy creation. With the arbitrary parametrisation here we are somewhere in between these two extreme values.

Table 1: Level & quarterly rate summary statistics:
Consistent with model parametrisation, 2002q1-2013q2

	Mean	Std err.
Tightness - $\theta_t = \frac{v_t}{u_t}$	0.35	0.039
Job finding rate - $\theta_t q(\theta_t) = \frac{m_t}{u_t}$	0.59	0.045
Job separation rate - λ_t	0.035	0.0021
'Equilibrium unemployment' - $u_t^* = \frac{\lambda_t}{\lambda_t + \frac{m_t}{u_t}}$	0.058	0.0088
Actual unemployment rate, 16-64	0.058	0.0039

Source: authors' calculations using UK Labour Force Survey and Labour Market Statistics

Table 2: Assumed/estimated parameter values and deterministic equilibrium

Deterministic parameter	Assumed value
y	1
b	0.8
c	0.74
λ	0.035
μ	1.08
α	0.39
β	0.39
δ	0.99
ρ	0.84
σ	0.006
Endogenous variable	**Deterministic equilibrium**
θ	0.35
u	0.058
v	0.02
w	0.98

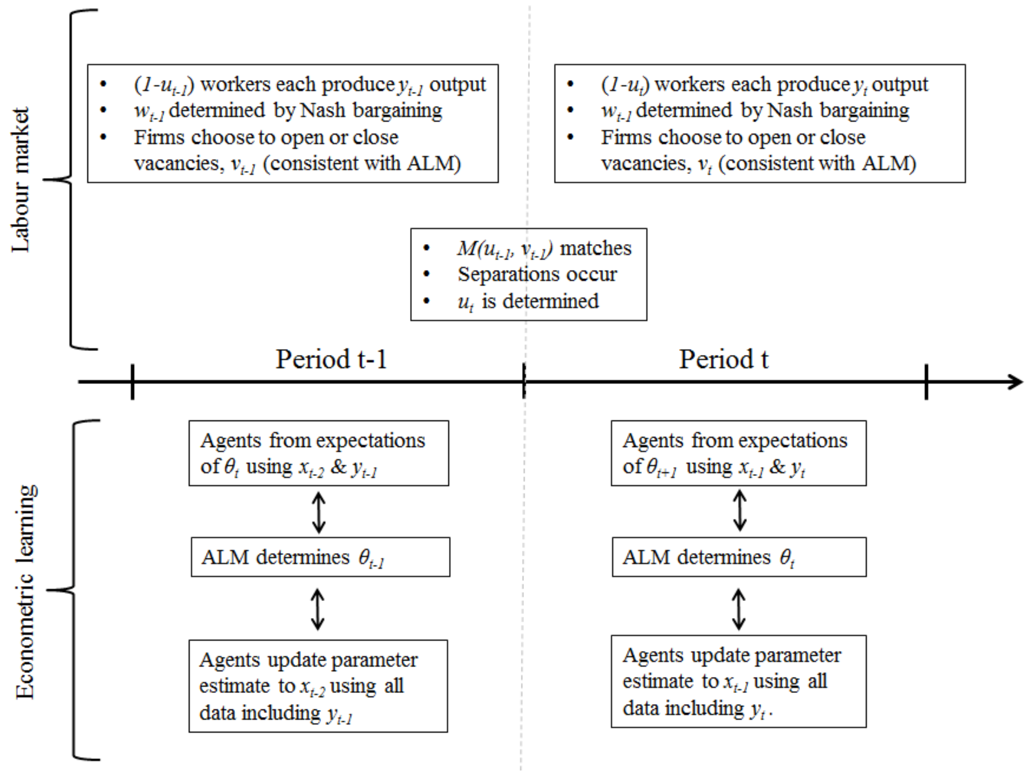
Source: authors' calculations

For completeness we write-out in full the stochastic recursive sequence that represents the model, stating from period t_0:

$$
\begin{aligned}
&(I) \quad u_{t+2} = \lambda(1-u_{t+1}) + \left[1-\mu\left(z_t'T(x_t)+V(x_t)\varepsilon_{t+1}\right)^{1-\alpha}\right]u_{t+1} \\
&(II) \quad R_{t+1} = R_t + \frac{1}{t+1}(z_t z_t' - R_t) \\
&(III) \quad x_{t+1} = x_t + \frac{1}{t+1}R_{t+1}^{-1}z_t\left\{z_t'[T(x_t)-x_t]+V(x_t)\varepsilon_{t+1}\right\} \\
&(IV) \quad y_{t+1} = (1-\rho)+\rho y_t + \varepsilon_{t+1} \\
&(V) \quad \varepsilon_{t+2} \sim i.i.d.\, N(0,\sigma^2).
\end{aligned}
$$

When written out in sequence order, the simultaneity which requires us to exclude y_t from the information set used to estimate x_t becomes clear. The adaptive learning process, which takes place at the beginning of each period, can also be represented by figure 1.

Figure 1: Timeline of the labour market & agents' learning



To initiate the sequence from t_0 we must choose initial values u_1, x_0, z_0 S_0 and ε_1. The asymptotic properties of decreasing or constant gains least squares recursion will hold irrespective of the initial conditions. As suggested by Carceles-Poveda & Giannitsarou (2007) the approach taken to set initial values z_0 and S_0 should depend on the particular model in question and the empirical purpose of the researchers. One approach we might take is to use historic or randomly generated data, with t_0 set sufficiently large such that S_0 is invertible; in this case $t_0 \geq 2$. This approach is most useful when attempting to compare the performance of models with learning and the assumption of agents as being 'good econometricians' against real data. However, this gives few clues as to how large t_0 should be, and the subsequent simulation is likely to be sensitive to this assumed level of memory of agents, particularly for decreasing gain least squares. Another attractive option is to choose initial values from an assumed distribution around the REE.

To set the initial conditions here we use the same data used to parametrise the model to estimate (36); where $cubtr_t$ represents a cubic time trend to account for the possibility that agents recognise low frequency structural breaks in the relationship; productivity is normalised but not de-trended; and we include significant MA terms to account for otherwise unaccounted for auto-correlation when the MSV is applied to real world data, which the good econometrician may in practice account for by what we have referred to before as an over-parametrised PLM.[10] $t_0 = 50$ is the maximum number of UK observations available to us.

$$\theta_t = A + cubtr_t + By_{t-1} + \kappa_1 \zeta_{t-1} + \kappa_2 \zeta_{t-2} + \zeta_t, \qquad t = 2001q2...2013q3. \tag{36}$$

Using this approach, we set $S_0 = \begin{pmatrix} 50 & 50 \\ 50 & 50.075 \end{pmatrix}$, $x_0' = (-1.22, 1.77)$, $z_0' = (1,1)$, $\varepsilon_1 = 0$ and

$$u_1 = \frac{\lambda}{\lambda + \mu \left(z_0' T(x_0)\right)^{1-\alpha}} \quad (= 0.052). \tag{37}$$

To analyse the impact of adding adaptive learning to the basic search frictions model of the labour market we should focus on the simulated time paths of wages and the tightness parameter, which are independent of u_1. With the parametrisation as described above the REE parameters of the MSV solution are given by $x_{REE}' = (-0.9, 1.25)$. That is to say the elasticity of θ to productivity at the long-run average level is around 3, which is significantly lower than observed in the data.

Figure 2 (see appendix) demonstrates a simulation for the baseline case of agents with rational expectations.[11] Unsurprisingly, as is common with this class of models, as described in table 3, the REE does not come close to matching the key moments for labour market tightness and unemployment. Figure 3 shows results from a simulation with decreasing gain learning taking initial values as described above. The key result is that when giving agents a relatively small amount of historical data (12.5 years), and initial estimates of the model parameters which are not unrealistically far from the true REE values, decreasing gain learning convergence is shown to be very slow. As shown by figure 4 this takes roughly 500 years despite being exponential. This indicates that under adaptive learning, an economy could be persistently away from its REE level of unemployment, either on the high or low side, even though agents are behaving rationally in the limited sense prescribed by the 'good econometrician.' In this sense, rational expectations is a poor approximation to a model with learning in terms of levels. One recommendation from this result is that when calibrating the Mortensen-Pissarides, targeting second

[10] In determining initial conditions, one could also consider the class of GARCH, error correction, or even VAR models, however we believe this would be an unnecessarily significant leap from the straightforward least squares updating we assume that a 'good econometrician' carries outs in practice, and which constitutes the learning algorithm we study here.

[11] Random number seed set to 42 in Python Numpy application for all simulations.

moments of the data should always be preferable, whereas not hitting levels of the endogenous variables should not be too concerning.

Table 3: Simulation results

	St.dev $t \leq t_0 + 20$	St.dev	Min $t \leq t_0 + 100$	Max
Rational expectations eq.				
w	0.01	0.008	0.96	0.99
u	0.0014	0.0012	0.056	0.061
θ	0.018	0.014	0.032	0.038
Learning (I) - least squares				
w	0.011	0.009	0.98	1.02
u	0.0014	0.0014	0.05	0.056
θ	0.023	0.019	0.37	0.45
Learning (IV) - constant gain $\gamma = 0.05$				
w	0.013	0.010	0.97	1.01
u	0.0018	0.0022	0.051	0.059
θ	0.027	0.026	0.33	0.44

Source: authors' calculations

As a further example in figure 5 we simulate the model with no memory, and allow the agents' to have guessed the correct initial parameter estimates $x_0' = x_{REE}'$, but suppose that there is an immediate negative 10% shock to the flow value of unemployment. Under rational expectations, due to the rise in the surplus of a match, firms immediately open more vacancies, and the unemployment rate falls. Under learning, the initial increase in θ is smaller. Therefore, unemployment falls more slowly as agents effectively attempt to disentangle the effects of the structural shock from the stochastic process. In this sense, the response to the shock leads to a less volatile response for unemployment. Thus, if actual labour market series feature the effects of frequent permanent shocks of this kind, econometric learning will not improve the ability of the model to match their cyclicality.

3.2 Speed of convergence

As shown theoretically in (Benveniste et al., 2012), the learning of the agents results in root-t convergence to the true REE parameter estimates if all the eigenvalues of the Jacobian of the system have real part $< 1/2$.[12] As shown, in the appendix when we derive the E-stability conditions, this requires that $\tilde{\psi} < 1/2$. In the example parametrisation above (figure 3) this is ensured, with $\tilde{\psi} = 0.39$. In fact, more generally, it can also be shown with simulations that the speed of convergence decreases substantially as the value of $\tilde{\psi}$ moves closer to 1, the threshold for E-Stability. To illustrate this, in figure 6 we consider a doubling of $\tilde{\psi}$ by decreasing worker bargaining power to $\beta = 0.1$, keeping all other parameters except c constant, which we again use to match our target value of θ. As expected, the rate of convergence is

[12] I.e. the rate at which in classical econometrics the mean of the least squares parameter estimate converges to the true value.

slowed substantially, and the economy remains more persistently away from the REE. A higher speed of convergence is one way in which the REE model could become an improved approximation of an alternative with econometric learning.

3.3 Constant gain learning

In figure 7, for completeness we also compare the results of our first simulation with decreasing gain learning to an equivalent example with a gain parameter $\gamma_t = 0.05$.[13] As expected, when agents weight recent data more highly, i.e. with weighted least squares learning, convergence to the REE is faster, and agents' parameter estimates are more volatile. This faster convergence also results in more volatile series of labour market tightness, wages and unemployment. However, the gain parameter we use here to guarantee faster convergence implies agents roughly only use data over the past 20 quarters to update their beliefs, and is notably outside the range suggested by the adaptive learning literature (see for example Di Pace et al. (2014) for a discussion). Therefore, the simulation results with constant gain learning, with more reasonable levels of memory weighting, are not dissimilar to those with decreasing gain.

4 Summary and discussion

We take the textbook RBC version of the Mortensen-Pissarides model of search and match frictions for the labour market and show that the unique REE is not only always E-stable, but this result is robust to over-parametrisation of the PLM used by agents (Strong E-stability) with decreasing gain learning. This PLM is assumed to be in the minimum state variable form, and as such agents only use productivity and estimated model parameters to form expectations of future labour market conditions. These local convergence conditions also extend trivially to global convergence. Hence, from the perspective by which we think agents may in practice apply something that looks like econometric learning, the rational expectations assumption in this class of model is not unreasonable. We use recent UK data to parametrise the model, and show that although the model is E-stable, implied convergence can be very slow. Therefore, the rational expectations model of unemployment could be a poor approximation to one in which agents learn econometrically, particularly in the presence of frequent structural or permanent shocks.

The so-called 'Shimer puzzle' addresses the extent to which the RBC version of the Mortensen-Pissarides can match key moments of economic time series, in particular the volatility of unemployment, and lack thereof for wages. Hagedorn & Manovskii (2008) discuss at length how the model can be calibrated to match these moments. They focus on the role of a high flow utility from non-employment (b), and low worker bargaining power (β) and matching elasticity (α) to match these volatilities. In our linearised system, the sensitivity of θ to productivity shocks is determined by the true REE parameter B, which becomes large as $\tilde{\psi}$ goes to 1. It is straightforward to show that the same combination of high

[13] For constant gain learning there is no analytic solution for expectational stability and so we must select a reasonably small gain parameters to ensure convergence.

b, low β & α not only implies higher volatility of θ and unemployment, but also slower convergence to the REE in the model with learning.[14] Thus, in a world with least squares learning, more 'extreme' values of these parameters would be implied to match the observed cyclicality of unemployment, wages etc.. And so the result of relaxing rational expectations in this limited way is to move us further from addressing the unemployment volatility puzzle.

References

Benhabib, J., & Farmer, R. E. A. (1999). Indeterminacy and sunspots in macroeconomics. In J. B. Taylor, & M. Woodford (Eds.) *Handbook of Macroeconomics*, vol. 1, Part A, chap. 06, (pp. 387–448). Elsevier, 1 ed.

Benveniste, A., Métivier, M., & Priouret, P. (2012). *Adaptive algorithms and stochastic approximations*. Springer Publishing Company, Incorporated.

Berardi, M., & Galimberti, J. K. (2012). On the plausibility of adaptive learning in macroeconomics: A puzzling conflict in the choice of the representative algorithm. Centre for Growth and Business Cycle Research Discussion Paper Series 177, The Univeristy of Manchester.

Berardi, M., & Galimberti, J. K. (2013). A note on exact correspondences between adaptive learning algorithms and the Kalman filter. *Economics Letters*, *118*(1), 139 – 142.

Blanchard, O., & Watson, M. W. (1983). Bubbles, rational expectations and financial markets. NBER Working Papers 0945, National Bureau of Economic Research.

Borowczyk-Martins, D., Jolivet, G., & Postel-Vinay, F. (2013). Accounting for endogeneity in matching function estimation. *Review of Economic Dynamics*, *16*(3), 440 – 451.

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

Burgess, S., & Turon, H. (2010). Worker flows, job flows and unemployment in a matching model. *European Economic Review*, *54*(3), 393 – 408.

Carceles-Poveda, E., & Giannitsarou, C. (2007). Adaptive learning in practice. *Journal of economic dynamics & control*, *31*(8), 2659 – 2697.

Di Pace, F., Mitra, K., & Zhang, S. (2014). Adaptive learning and labour market dynamics. Tech. rep., Centre for Dynamic Macroeconomic Analysis.

Eusepi, S., & Preston, B. (2011). Expectations, learning, and business cycle fluctutations. *The American Economic Review*, *101*, pp. 2844 – 2872.

[14]Note, for $\tilde{\psi}$ to be decreasing in α we require $\bar{\theta} < \frac{1-\alpha}{\alpha}$ which will generally be true for reasonable parameter values and long-run equilibrium value of labour market tightness.

Evans, G., & Honkapohja, S. (1986). A complete characterization of ARMA solutions to linear rational expectations models. *The Review of Economic Studies*, *53*(2), 227–239.

Evans, G. W., & Honkapohja, S. (1998). Economic dynamics with learning: New stability results. *The Review of Economic Studies*, (1), 23.

Evans, G. W., & Honkapohja, S. (2001). *Learning and expectations in macroeconomics*. Frontiers of economic research. Princeton University Press, Princeton, NJ.

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

Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit & Banking*, *29*(1), 1 – 16.

Kreps, D. M. (1998). Anticipated utility and dynamic choice. *Econometric Society Monographs*, *29*, 242–274.

Mankiw, N. G., Reis, R., & Wolfers, J. (2004). Disagreement about inflation expectations. In *NBER Macroeconomics Annual 2003, Volume 18*, (pp. 209–270). The MIT Press, Cambridge, MA.

Marcet, A., & Sargent, T. J. (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, *48*(2), 337 – 368.

McCallum, B. T. (1983). On non-uniqueness in rational expectations models: An attempt at perspective. *Journal of Monetary Economics*, *11*(2), 139 – 168.

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

Milani, F. (2011). Expectation shocks and learning as drivers of the business cycle*. *The Economic Journal*, *121*(552), 379–401.

Mortensen, D. T., & Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, *61*(3), 397–415.

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

Pfajfar, D., & Santoro, E. (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior & Organization*, *75*(3), 426–444.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. The MIT Press, Cambridge, MA.

Rogerson, R., & Shimer, R. (2011). Search in macroeconomic models of the labor market. In O. Ashenfelter, & D. Card (Eds.) *Handbook of Labor Economics*, vol. 4, Part A, (pp. 619 – 700). Elsevier, North Holland.

Şahin, A., Song, J., Topa, G., & Violante, G. L. (2012). Mismatch unemployment. Tech. rep., National Bureau of Economic Research.

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

A Appendix

A.1 Linearisation

We take a first order Taylor approximation around the deterministic steady state values of θ and y_t, $\bar{\theta}$ and $\bar{y} = 1$ respectively, approximating the right and the left hand side of equation (14) which is stated here again for convenience

$$\frac{c}{\delta q(\theta_t)} = (1-\beta)(y^e_{t+1} - b) + \frac{(1-\lambda)c}{q(\theta^e_{t+1})} - \theta^e_{t+1}\beta c. \tag{38}$$

This results in

$$\begin{aligned}\frac{c}{\delta q(\bar{\theta})} - \frac{cq'(\bar{\theta})}{\delta[q(\bar{\theta})]^2}(\theta_t - \bar{\theta}) =& (1-\beta)(\bar{y} - b) + (1-\beta)(y^e_{t+1} - \bar{y}) \\ & - \bar{\theta}\beta c - \beta c(\theta^e_{t+1} - \bar{\theta}) + \frac{(1-\lambda)c}{q(\bar{\theta})} \\ & - \frac{(1-\lambda)cq'(\bar{\theta})}{[q(\bar{\theta})]^2}(\theta^e_{t+1} - \bar{\theta}).\end{aligned} \tag{39}$$

By noting that

$$\frac{c}{\delta q(\bar{\theta})} = (1-\beta)(\bar{y} - b) + \frac{(1-\lambda)c}{q(\bar{\theta})} - \bar{\theta}\beta c \tag{40}$$

must hold in equilibrium according to (14), this steady state condition can be subtracted from both sides of the approximated equation. Then solving explicitly for θ_t and defining the functional form $q(\theta) = \mu\theta^{-\alpha}$, (39) becomes

$$\begin{aligned}\theta_t = & \left\{1 + \frac{\beta\delta\mu^2(\bar{\theta})^{-2\alpha}}{\alpha\mu(\bar{\theta})^{-\alpha-1}} - (1-\lambda)\delta\right\}\bar{\theta} \\ & - \frac{(1-\beta)\delta\mu^2(\bar{\theta})^{-2\alpha}}{c\alpha\mu(\bar{\theta})^{-\alpha-1}}\bar{y} \\ & + \frac{(1-\beta)\delta\mu^2(\bar{\theta})^{-2\alpha}}{c\alpha\mu(\bar{\theta})^{-\alpha-1}}y^e_{t+1} \\ & + \left\{-\frac{\beta\delta\mu^2(\bar{\theta})^{-2\alpha}}{\alpha\mu(\bar{\theta})^{-\alpha-1}} + (1-\lambda)\delta\right\}\theta^e_{t+1}.\end{aligned} \tag{41}$$

which can be simplified to

$$
\begin{aligned}
\theta_t = & \left\{1+\frac{\beta\delta\mu(\bar{\theta})^{1-\alpha}}{\alpha}-(1-\lambda)\delta\right\}\bar{\theta} \\
& -\frac{(1-\beta)\delta\mu(\bar{\theta})^{1-\alpha}}{c\alpha}\bar{y} \\
& +\frac{(1-\beta)\delta\mu(\bar{\theta})^{1-\alpha}}{c\alpha}y_{t+1}^{e} \\
& +\left\{-\frac{\beta\delta\mu(\bar{\theta})^{1-\alpha}}{\alpha}+(1-\lambda)\delta\right\}\theta_{t+1}^{e}.
\end{aligned}
\tag{42}
$$

The form given in the text, (16), $\theta_t = \psi_0 + \psi_1 y_{t+1}^e + \psi_2 \theta_{t+1}^e$ thus has coefficients

$$
\psi_0 = \left\{1+\frac{\beta\delta\mu(\bar{\theta})^{1-\alpha}}{\alpha}-(1-\lambda)\delta\right\}\bar{\theta}-\frac{(1-\beta)\delta\mu(\bar{\theta})^{1-\alpha}}{c\alpha}\bar{y}, \tag{43}
$$

$$
\psi_1 = \frac{(1-\beta)\delta\mu(\bar{\theta})^{1-\alpha}}{c\alpha}, \tag{44}
$$

$$
\psi_2 = \left\{-\frac{\beta\delta\mu(\bar{\theta})^{1-\alpha}}{\alpha}+(1-\lambda)\delta\right\}, \tag{45}
$$

with the steady state value for labour market tightness the solution to

$$
(1-\beta)(\bar{y}-b)-\frac{(\frac{1-\delta}{\delta}+\lambda)c\bar{\theta}^{\alpha}}{\mu}-\beta c\bar{\theta}=0. \tag{46}
$$

A.2 REE Values

The REE values of the parameters A, B, and C are found using the method of undetermined coefficients

$$
A = \frac{\tilde{\psi}_0}{1-\tilde{\psi}_2}+\frac{\tilde{\psi}_1\tilde{\psi}_2(1-\rho)}{(1-\tilde{\psi}_2)(1-\tilde{\psi}_2\rho)}, \tag{47}
$$

$$
B = \frac{\tilde{\psi}_1}{1-\tilde{\psi}_2\rho}, \tag{48}
$$

$$
C = B\rho^{-1}, \tag{49}
$$

where we have assumed that $\tilde{\psi}_2 \neq 1$ and $\tilde{\psi}_2\rho \neq 1$.

A.3 E-Stability

According to Evans & Honkapohja (2001) an REE is E-stable if the associated ordinary differential equation given in (30) is asymptotically locally stable under learning. This is the case, if all the eigenvalues of the Jacobian of

$$
T(\hat{A},\hat{B})-(\hat{A},\hat{B}) \tag{50}
$$

have negative real parts, see Evans & Honkapohja (2001). Thus, we must have

$$
\tilde{\psi}_2\rho-1<0 \tag{51}
$$

and

$$\tilde{\psi}_2 - 1 < 0, \tag{52}$$

whereby the second condition implies the validity of the first. Therefore, we need to check that the second condition is true. Writing out the term $\tilde{\psi}_2$ and rearranging, we see that

$$\delta((1-\lambda) - \beta\mu(\bar{\theta})^{1-\alpha}\alpha^{-1}) < 1 \tag{53}$$

is always true for the widest range of sensible assumed parameter values ($0 \leq \delta < 1, \quad \lambda > 0, \quad 0 < \beta < 1, \quad \mu > 0, \quad \alpha > 0;$ and which all imply $\bar{\theta} > 0$) and hence the REE is E-stable.

A.4 Global convergence

Given the model discussed here has a unique equilibrium, and satisfies the assumptions of Evans & Honkapohja (1998) that guarantee global convergence, we simply apply their *Theorem 2* to the recursive learning algorithm given by (27) and (29).

For R_t, using $Ez_tz_t^{'} = M_z$, where M_z is some positive definite matrix, taking expectations we have the ODE

$$\frac{dR}{d\tau} = M_z - S, \tag{54}$$

which is clearly globally asymptotically stable and independent of x_t.

It is possible that for some t R_t may not be invertible, though this will happen only a finite number of times with probability 1. We modify the algorithm for x_t to

$$x_t = x_{t-1} + t^{-1}u(R_{t-1})z_{t-1}\left\{z_{t-1}'\left[T\left(\hat{A}_t, \hat{B}_t\right) - x_{t-1}\right] + \eta_t\right\}, \tag{55}$$

where $u(R)$ is a bounded regular function from the space of $2x2$ matrices to the subspace of positive definite matrices such that $u(R) = R^{-1}$ in the neighbourhood of M_z. Then taking expectations the ODE is given by

$$\frac{dx}{d\tau} = u(s)M_z(T(\hat{A},\hat{B}) - (A,B))' \tag{56}$$

$$= u(s)M_z(\tilde{\psi}_2 - 1)((\hat{A},\hat{B}) - (A,B))'. \tag{57}$$

Given that the other requirements of the theorem are trivially satisfied, then it applies, and this differential equation is clearly globally asymptotically stable for $\tilde{\psi}_2 < 1$, and this stability is exponential; $(\hat{A},\hat{B}) \rightarrow (A,B)$ globally almost surely.

A.5 ALM and T-mapping ARMA solution and E-stability

$$
\begin{aligned}
\theta_t = {} & \frac{\tilde{\psi}_0+\tilde{\psi}_2(a+b_1(1-\rho))}{1-\tilde{\psi}_2 c_1} + \frac{\tilde{\psi}_1+\tilde{\psi}_2(b_2+b_1\rho)}{1-\tilde{\psi}_2 c_1} y_{t-1} + \frac{\tilde{\psi}_2(b_1+d_1)+\tilde{\psi}_1\rho^{-1}}{1-\tilde{\psi}_2 c_1}\varepsilon_t + \frac{\tilde{\psi}_2 f_1}{1-\tilde{\psi}_2 c_1}\eta_t \\
& + \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1}\sum_{j=3}^{s} b_j y_{t+1-j} + \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1}\sum_{j=2}^{r} c_j \theta_{t+1-j} \\
& + \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1}\sum_{j=2}^{q} d_j \varepsilon_{t+1-j} + \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1}\sum_{j=2}^{l} f_j \eta_{t+1-j}.
\end{aligned}
\tag{58}
$$

This defines again a T-mapping from the PLM to the ALM with corresponding elements

$$a = \frac{\tilde{\psi}_0+\tilde{\psi}_2(a+b_1(1-\rho))}{1-\tilde{\psi}_2 c_1}, \tag{59}$$

$$b_1 = \frac{\tilde{\psi}_1+\tilde{\psi}_2(b_1\rho+b_2)}{1-\tilde{\psi}_2 c_1}, \tag{60}$$

$$d_0 = \frac{\tilde{\psi}_1\rho^{-1}+\tilde{\psi}_2(b_1+d_1)}{1-\tilde{\psi}_2 c_1}, \tag{61}$$

$$b_j = \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1} b_{j+1}, \quad j=2,\ldots,s-1, \quad b_s = 0, \tag{62}$$

$$c_j = \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1} c_{j+1}, \quad j=1,\ldots,r-1, \quad c_r = 0, \tag{63}$$

$$d_j = \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1} d_{j+1}, \quad j=1,\ldots,q-1, \quad d_q = 0, \tag{64}$$

$$f_j = \frac{\tilde{\psi}_2}{1-\tilde{\psi}_2 c_1} f_{j+1}, \quad j=0,\ldots,l-1, \quad f_l = 0. \tag{65}$$

Since (59) - (65) describes a non-linear system of differential equations, we first have to linearise (34) to study stability properties. However, the subsystem (63) is independent of the other equations and can be analysed separately. The eigenvalues of the Jacobian of $T(\mathbf{c})-\mathbf{c}$ at the REE values $c_j = 0$ for $j = 1t,\ldots,r$ are found to be r times repeatedly equal to -1 and therefore the subsystem (63) will converge towards the REE values. Due to the convergence of $\mathbf{c}$ it is apparent that $\mathbf{d}$ (apart from d_0) and $\mathbf{f}$ will also converge to their REE values being vectors of zeros. Moreover, $b_j = 0$ for $j = 2,\ldots,s$ is easily verified to be the values towards which the economy under learning converges. Finally, convergence of a, b_1 and d_0 are studied by analysing the Jacobian of the system (59)-(61). If this Jacobian has eigenvalues strictly less than unity, then the whole system is E-stable. It can easily be verified that the eigenvalues are $\tilde{\psi}_2$ and $\tilde{\psi}_2\rho$.

A.6 ARMA(1,1) and the MSV solution

Derivation of MSV solution: (19) can be re-written as

$$\theta_t = \frac{\rho\tilde{\psi}_0-\tilde{\psi}_1(1-\rho)}{\rho(1-\tilde{\psi}_2)} + \frac{\tilde{\psi}_1}{\rho(L-\tilde{\psi}_2)} y_{t-1} - \frac{d_1\tilde{\psi}_2}{(L-\tilde{\psi}_2)}\varepsilon_t - \frac{d_2}{(L-\tilde{\psi}_2)}\eta_t, \tag{66}$$

with L denoting the lag operator such that $Lx_t = x_{t-1}$. The parameter d_1 and d_2 can be chosen arbitrarily. In particular, to obtain the MSV solution

$$\theta_t = A + By_{t-1} + C\varepsilon_t \tag{67}$$

one must first set $d_2 = 0$. (66) can be re-written as:

$$\theta_t = \frac{\rho\tilde{\psi}_0 - \tilde{\psi}_1(1-\rho)}{\rho(1-\tilde{\psi}_2)} - (\rho^{-1}\tilde{\psi}_1 y_{t-1} - d_1\tilde{\psi}_2\varepsilon_t)\sum_{i=1}^{\infty}\tilde{\psi}_2^{-i}L^{i-1}. \tag{68}$$

$$\begin{aligned}\theta_t =& \frac{\rho\tilde{\psi}_0 - \tilde{\psi}_1(1-\rho)}{\rho(1-\tilde{\psi}_2)} + \rho^{-1}\tilde{\psi}_1\tilde{\psi}_2^{-1}(1-\rho)\sum_{i=1}^{\infty}(\sum_{j=1}^{i}\rho^{-j})\tilde{\psi}_2^{-i} - \rho^{-1}\tilde{\psi}_1\tilde{\psi}_2^{-1}y_{t-1}\sum_{i=0}^{\infty}(\rho\tilde{\psi})^{-i} \\ &+ \varepsilon_{t-1}(\rho^{-1}\tilde{\psi}_1\tilde{\psi}_2^{-1}\sum_{i=1}^{\infty}(\sum_{j=1}^{i}\rho^{-j}L^{i-j})\tilde{\psi}_2^{-i} + d_1\sum_{i=1}^{\infty}\tilde{\psi}_2^{-i}L^{i-1}) + d_1\varepsilon_t.\end{aligned} \tag{69}$$

Therefore, to derive an MSV solution from a broader the class of ARMA(1,1) solutions, in which no lags of ε_t can remain, we therefore see from (69) that

$$d_1 = -\frac{\tilde{\psi}_1}{\rho\tilde{\psi}_2}\left(\frac{1}{\rho\tilde{\psi}_2} + (\frac{1}{\rho\tilde{\psi}_2})^2 + (\frac{1}{\rho\tilde{\psi}_2})^3 + \dots\right) \tag{70}$$

$$= \frac{\tilde{\psi}_1}{\rho\tilde{\psi}_2(1-\rho\tilde{\psi}_2)} \qquad if \quad \tilde{\psi}_2 > \frac{1}{\rho} > 1, \tag{71}$$

which corresponds to the condition for stable ARMA(1,1) solutions. Otherwise, the MSV solution cannot be derived from the class of unstable ARMA(1,1) solutions, and is instead the only stable solution.

A.7 Simulation Figures

Figure 2: Baseline – Rational expectations eq.

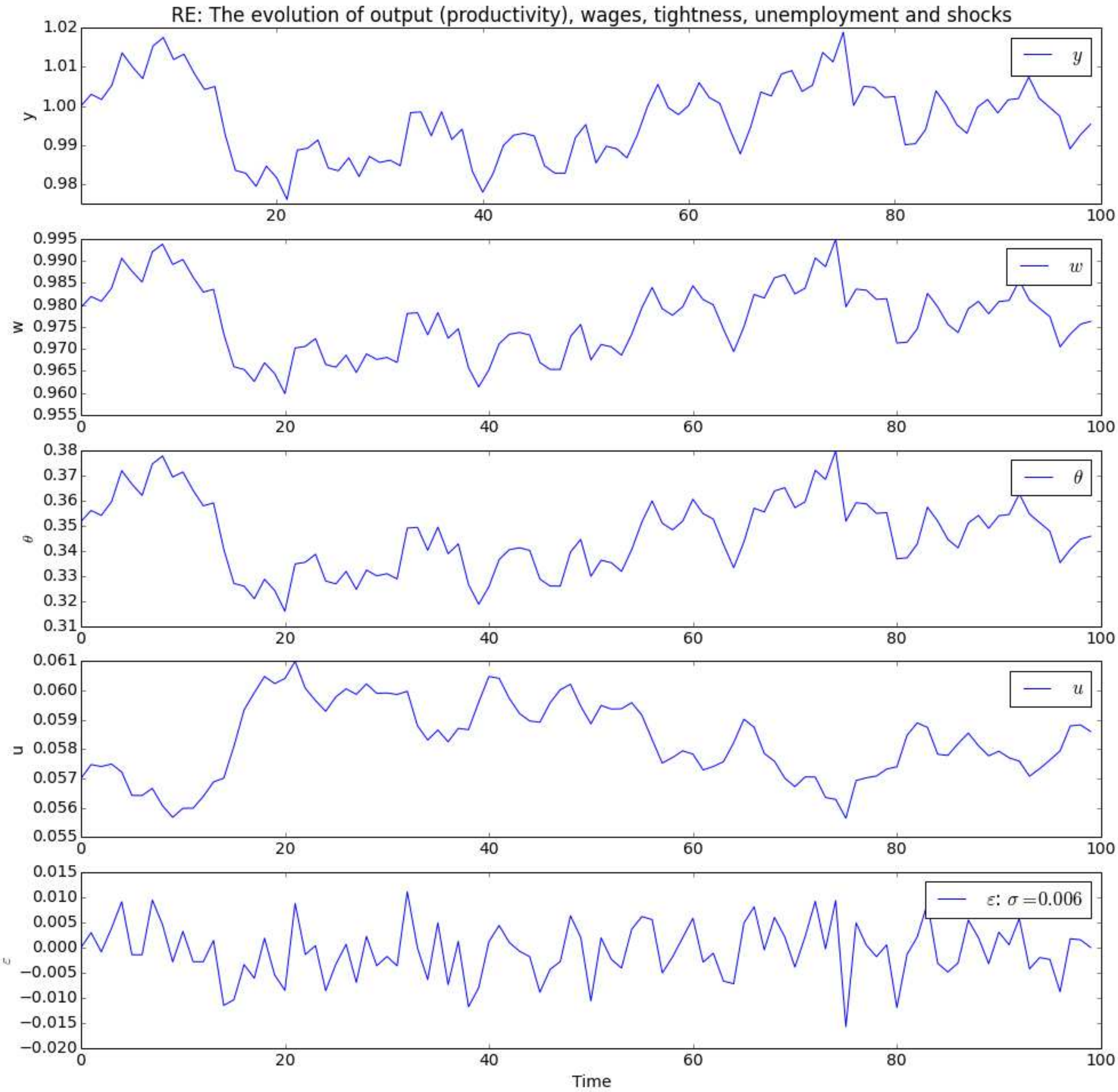
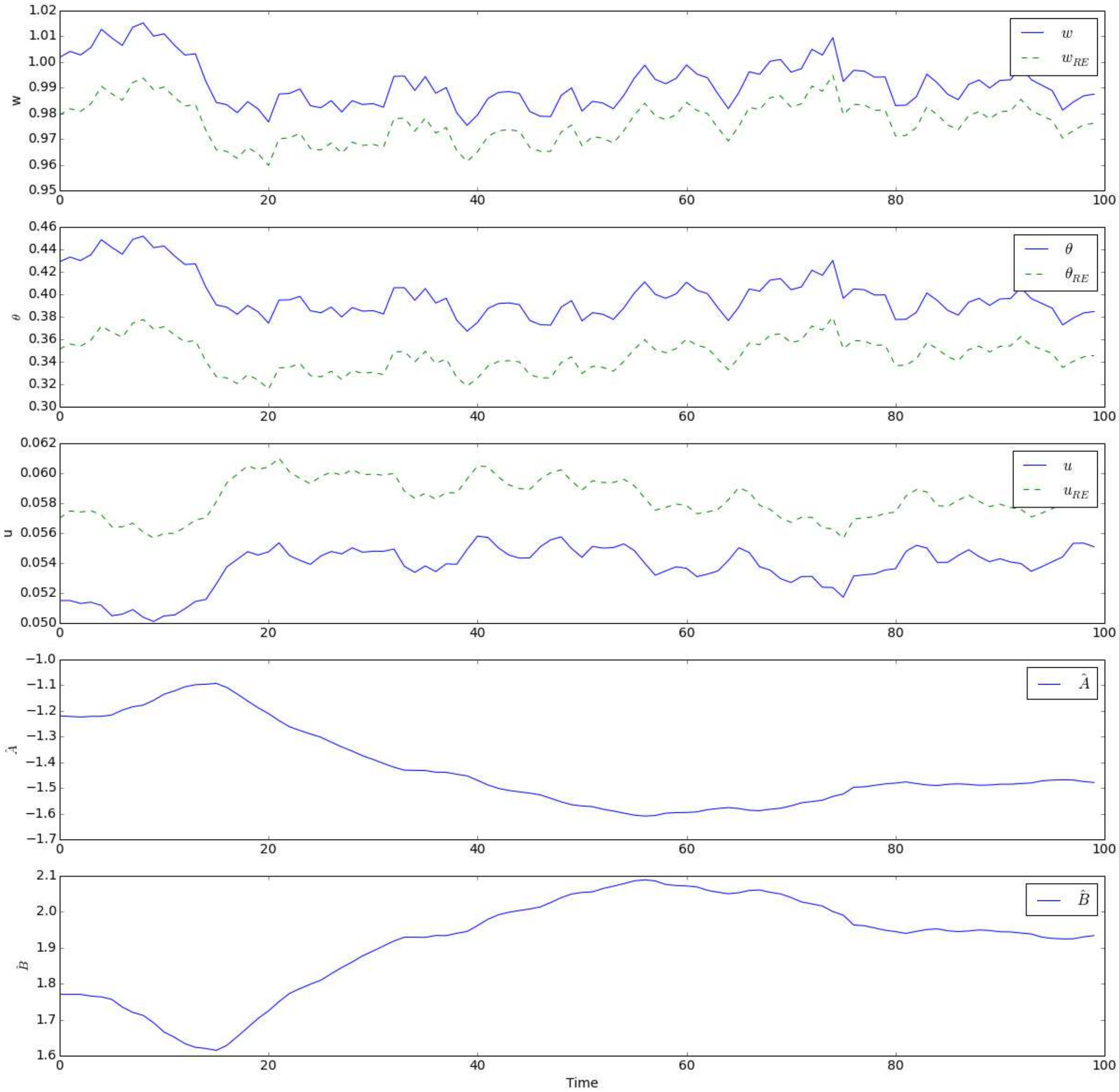


Figure 3: Learning (I) - Comparison of recursive least squares learning with rational expectations eq.



With initial parameter estimates assumed to be 'realistically' far away from true REE values

Figure 4: Learning (I) - Convergence of parameter estimates

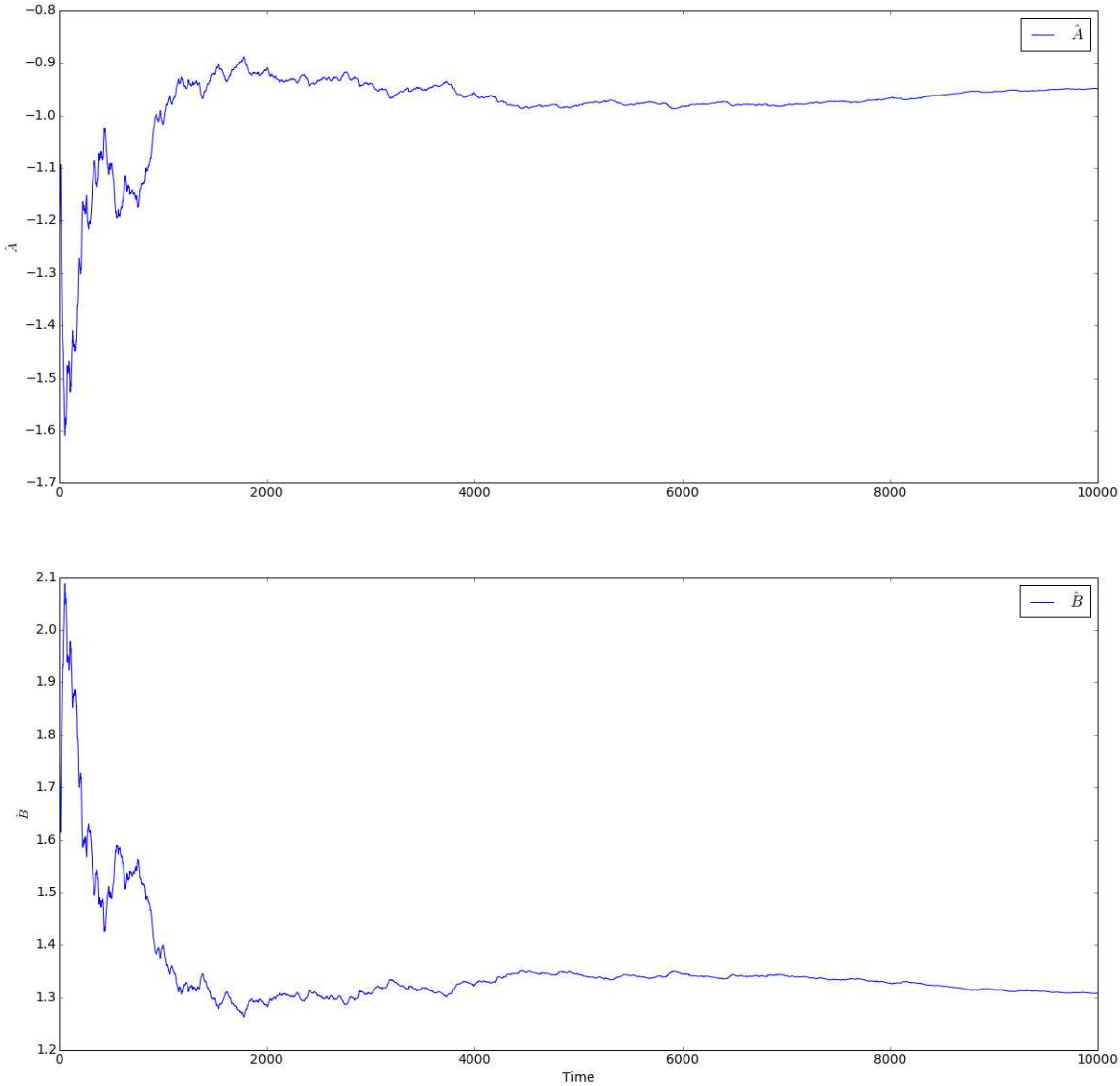
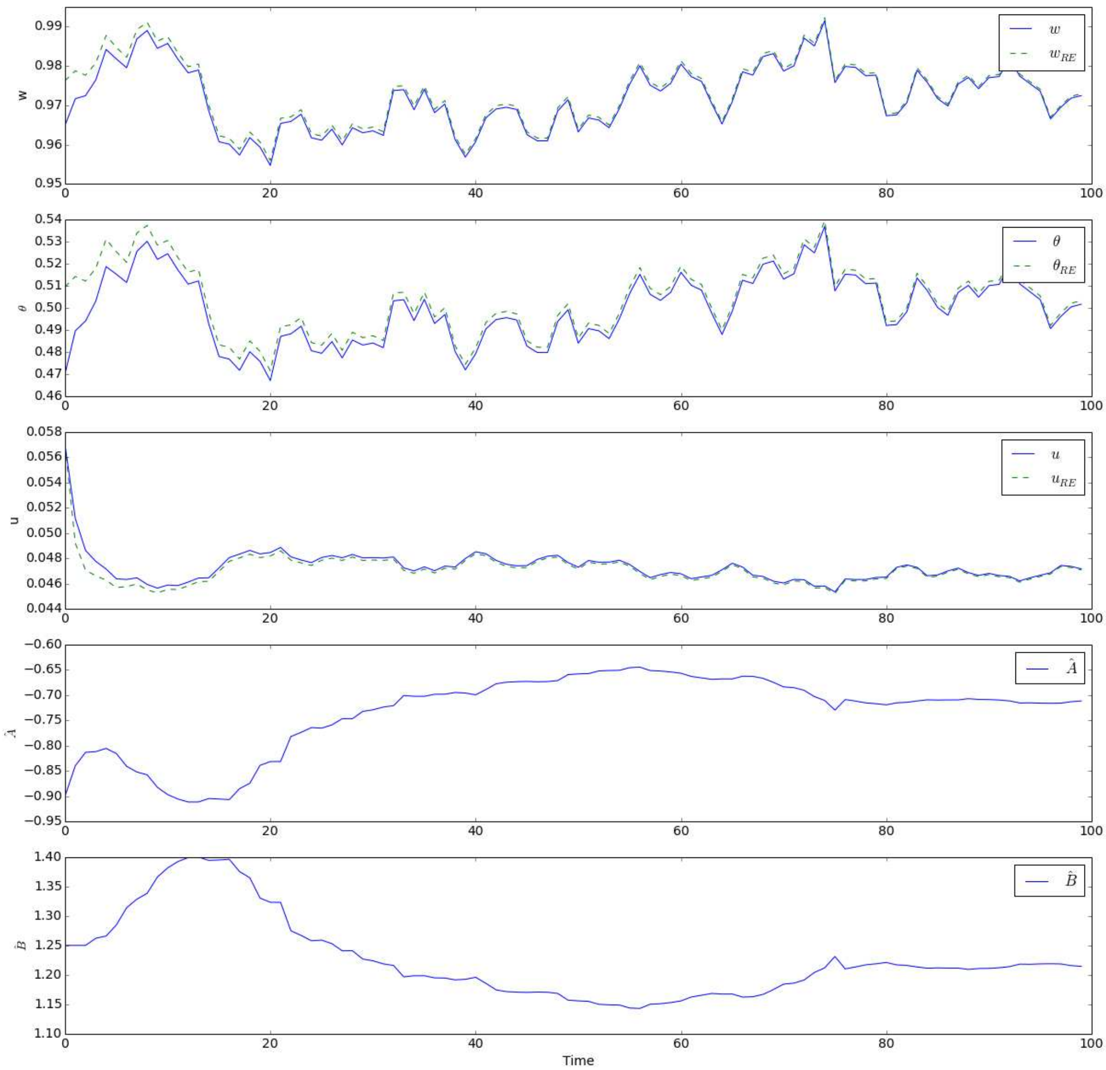
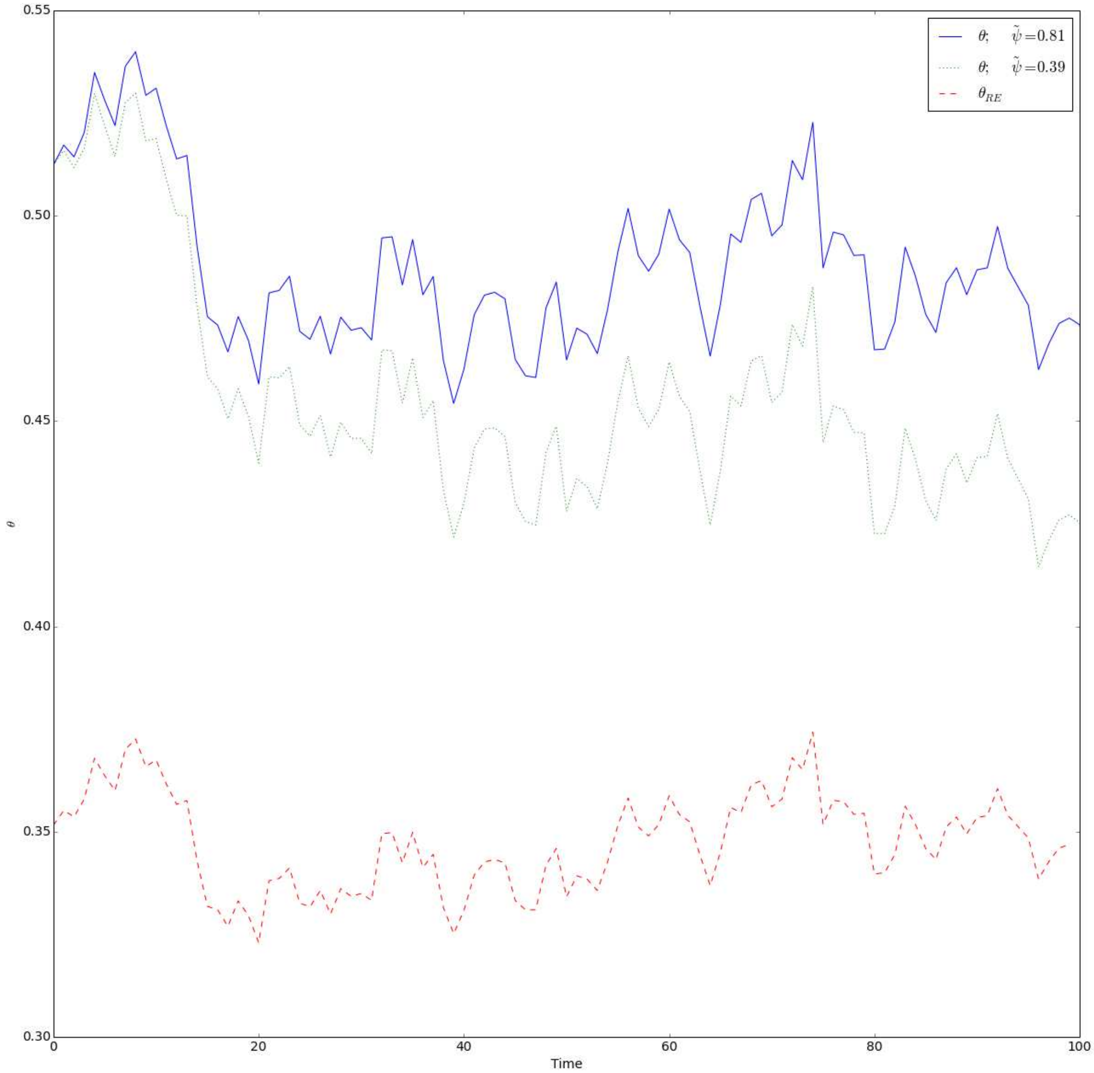


Figure 5: Learning (II) - Comparison of recursive least squares learning with rational expectations eq.



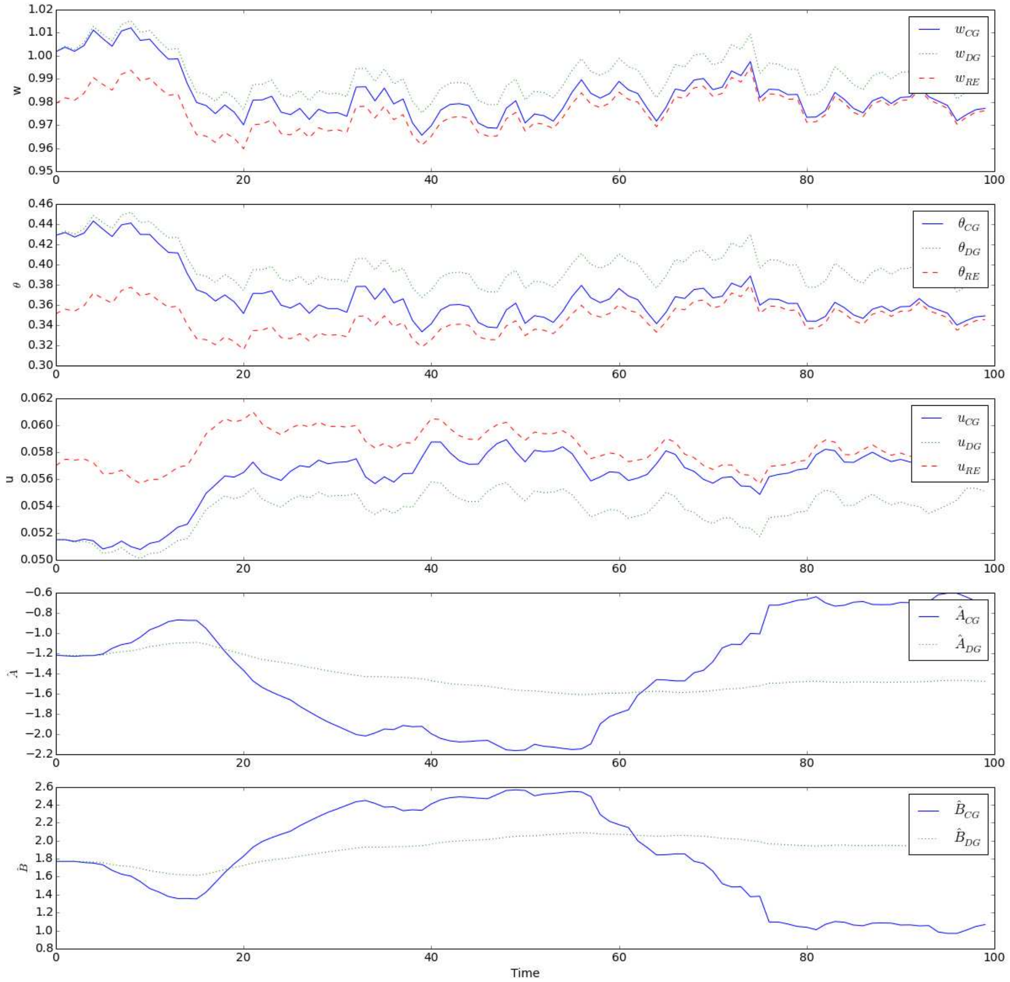
Following a -10% shock to the flow value of unemployment, with initial parameter estimates assumed to be at the true REE values

Figure 6: Learning (III) - Speed of convergence comparison, changing worker bargaining power β



Given $\beta = 0.1$, then $\tilde{\psi} = 0.81$. For $\beta = \alpha$, $\tilde{\psi} = 0.39$.

Figure 7: Learning (IV) - Comparison of decreasing and constant gain learning



With initial parameter estimates assumed to be 'realistically' far away from true REE values