

# Learning in a Misspecified Multivariate Self-referential Linear Stochastic Model

*Eran A. Guse*

October 2005

CWPE 0548

**Not to be quoted without permission**

# Learning in a Misspecified Multivariate Self-Referential Linear Stochastic Model

Eran A. Guse\*  
University of Cambridge

October 19, 2005

## Abstract

This paper introduces a general method to study stability (under learning) of equilibria resulting from agents with misspecified perceptions of the law of motion of the economy. This is done by transforming the actual and perceived laws of motion into the form of seemingly unrelated regressions and then linearly projecting the actual law of motion into the same class as the perceived law of motion. I study the New Keynesian IS-LM model with inertia under all possible classes of restricted perceptions. It turns out that the results found in Bullard and Mitra (2002, 2003) are robust under misspecified expectations.

**Key Words:** Adaptive Learning; Expectational Stability; Monetary Policy Rules; Restricted Perceptions Equilibria; Seemingly Unrelated Regression

**JEL Classification:** E4; E5

---

\*Address correspondence to: Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, UK. Phone: +44 1223 335473. E-mail: eg295@econ.cam.ac.uk. I would like to thank George Evans, Seppo Honkapohja, Ben McCallum, and participants of the Bank of Finland Research Department seminars for useful comments and suggestions. All errors are mine.

# 1 Introduction

Rational expectations (RE) was introduced by Muth (1961) and later popularized by Lucas (1972,1976) and Sargent (1973) in the field of dynamic macroeconomics. Although it is the most widely used equilibrium concept in macroeconomics, it does have several limitations. One drawback to using RE is that it assumes that agents are aware of the true distribution of the economy and are able to form conditional expectations with this information. Therefore, as pointed out by Sargent (1993), an economist who uses RE is thus assuming that the agents in his model are more intelligent than he. The adaptive learning literature, discussed in Sargent (1993) and Evans and Honkapohja (2001), provides an alternative to RE in which one assumes that agents act as econometricians and estimate unknown parameters in a correctly specified model. Under certain conditions, it has been found that expectations in these models converge to the rational expectations equilibrium (REE).

Although adaptive learning may be a more realistic approach to modelling expectations, the common assumption of learning a correctly specified model may continue to be quite an abstraction from reality. In practice, econometricians cannot include all lags and exogenous variables when forecasting due to degrees of freedom restrictions, computational costs, or informational uncertainties. In fact, it is argued by White (1994) that all econometric models are necessarily misspecified. This paper will provide a general technique for a multivariate model in which to study stability, under adaptive learning, of the resulting equilibria when agents are modelled as econometricians who use a misspecified forecasting model.

There are many reasons why an agent may use a misspecified forecasting model to form expectations. Evans and Honkapohja (2001) suggest that agents may lack the full information set to use a fully specified forecasting model. In a laboratory, Adam (2002) discovers that agents may prefer to use a univariate forecasting model due to its simplicity. Such a forecasting model may be misspecified in a multivariate economic model. Perhaps agents take into account the performance of a forecasting model versus its costs as suggested in Adam (2005). If a fully specified forecasting model is too costly, then agents may choose to use a misspecified model that performs relatively well. Adam (2005) also suggests that another reason for misspecification is that agents may search for the best forecasting model in the short run. It may take many observations to discover the best forecasting model, so misspecification may be a short term phenomenon. Finally, less educated agents may use forecasting models with even fewer lags and exogenous variables as they may have a poorer understanding of the economy than a typical econometrician. Therefore, when modelling learning, it may be useful to further abstract from rationality by assuming that agents must choose their forecasting model subject to constraints.

In a misspecified model, Evans and Honkapohja (2001) call the resulting equilibria "restricted perceptions" equilibria (RPE) since agents are not fully aware of the true stochastic process of the aggregate variables they are forecasting. The RPE are fully rational in the class of restricted predictors as agents are ignorant of their restricted

perceptions. This ignorance is due to the orthogonality condition in least squares where the forecast errors are uncorrelated with the information used in the forecasting model. In a RPE, agents and their activities exist in economic models, but the agents' beliefs reside in their forecasting models. Therefore, the expectations are boundedly rational as if an agent were to step outside of the forecasting model, he would recognize the misspecification and prefer to use the correctly specified forecasting model. One limitation of the RPE is that excluded variables that should be included in a forecast will be correlated with the forecast errors. However, several reasons for misspecification, discussed above, also lead to agents being unaware of such a correlation.

Representation of the mapping from the perceived law of motion to the "projected" actual law of motion has so far been done case by case as there is currently no general tool for such representations. Most papers have considered mainly univariate models as calculation of more complex models proves to be quite difficult. I believe that since there is no general technique available for representing a T-mapping for a restricted perceptions equilibrium, economists have ignored many acceptable alternative equilibria in popular economic models such as the real business cycle model and the New Keynesian IS-LM model.

The purpose of this study is twofold. First, I will introduce a general technique to calculate projected T-maps for RPE's in order to evaluate the conditions for these equilibria to be learnable under simple adaptive learning rules. This technique will use the method of Seemingly Unrelated Regressions (SUR) originally analyzed by Zellner (1962) and presented e.g. by Hamilton (1994). By representing every regression as a single regression, the estimation restrictions can be quite general. Second, I will present an example of a restricted perceptions equilibrium in the New Keynesian IS-LM model as presented in Woodford (2003) and others. In this paper, I will follow Bullard and Mitra (2002) and use the T-map in order to evaluate how restricted perceptions will affect the stability properties of the model under learning.

Under a New Keynesian Monetary model with inertia, I show that the stability properties, under learning, do not change under a simple restricted forecasting rule. I show a relationship between determinacy, learnability, and inertia where determinacy only changes with inertia if the REE and RPE are not stable under learning. As in Bullard and Mitra (2002), the preferred policy rule is adjusting interest rates more than one for one (yet not too aggressively) with a change in inflation. The best policy to accomplish this, under the policy rules studied, is by following a Taylor rule with expected values of contemporaneous variables.

## **2 The Projected ALM**

In this section, a technique is presented which may be useful when evaluating stability of RPE's under learning. To do this, the ALM must be projected into the same class of models as the restricted perceived law of motion.

## 2.1 The Model, Determinacy, and Rational Expectations

Assume the following model:

$$z_t = A + BE_t^* z_t + DE_t^* z_{t+1} + Fz_{t-1} + Gv_t \quad (1)$$

where  $z_t$  is a  $(n \times 1)$  vector of endogenous variables,  $v_t$  is a vector of possibly autocorrelated shocks,<sup>1</sup> and  $E_t^*$  represents the not necessarily rational expectation operator at time  $t$ . The model is determinate when there is a unique nonexplosive REE and indeterminate if there are multiple nonexplosive solutions. Indeterminacy can be undesirable as agents may coordinate on an equilibrium that exhibits unwanted properties.

A model can be analyzed for determinacy by writing it up as a discrete difference equation with the associated extraneous noise terms capturing the errors in the agents' forecasts of the free variables. For the above model (assuming that  $D^{-1}$  exists)<sup>2</sup>, it turns out that determinacy can be determined by the matrix  $U$ :

$$U = \begin{pmatrix} D^{-1}(I - B) & -D^{-1}F \\ I & 0 \end{pmatrix}$$

If the number of eigenvalues outside the unit circle is equal to the number of explanatory variables,  $n$ , then the model is said to be determinate. If the number of eigenvalues outside the unit circle is greater than  $n$ , then the model is explosive. Finally, if the number of eigenvalues outside the unit circle is equal to  $n - m$ , then, following the terminology of Evans and McGough (2004a), the model is indeterminate of order  $m$ .<sup>3</sup>

A rational expectations equilibrium (REE) of this model can be presented as a minimal state variable (MSV) solution in a vector autoregression (VAR) form. Agents with rational expectations who know the correct MSV form will thus form expectations with the knowledge that  $z_t$  takes the following form:

$$z_t = a' + m'z_{t-1} + h(c'v_{t-1}) + \varepsilon_t \quad (2)$$

where  $h(x) = x$  if  $v_t$  is autocorrelated and  $h(x)$  is not included in the equation (or matrix) if  $v_t$  is i.i.d. Under a REE,  $a'$ ,  $m'$ , and  $h(c')$  are such that the stochastic process represented by equation (2) is consistent with the actual stochastic process of  $z_t$ .

Economic agents with RE have a great deal of knowledge of  $z_t$ . It is common to ask whether a REE is robust when agents form expectations using less sophisticated schemes than RE. Assume that agents know the true form of  $z_t$  represented by equation (2), but they must learn the parameters  $a'$ ,  $m'$ , and  $h(c')$ . Equation (2) is known as the perceived law of motion (PLM). Expectations from the PLM feedback into the

<sup>1</sup>If  $v_t$  is autocorrelated then assume that  $v_t = \rho v_{t-1} + \varepsilon_t$  where  $\varepsilon_t \sim N(0, \Sigma)$  and all the eigenvalues of  $\rho$  are inside the unit circle.

<sup>2</sup>If  $D$  is singular, then one must break the general form down to predetermined and non-predetermined variables and follow Blanchard and Kahn (1980).

<sup>3</sup>In Evans and McGough (2004a), all of the eigenvalues must be real for the model to be indeterminate of order  $m$ . As I am not considering sunspots, I just report the number of eigenvalues less than  $n$ .

structural model, equation (1), giving the actual law of motion (ALM):

$$z_t = A + Ba' + D(I + m')a' + (Bm' + Dm'^2 + F)z_{t-1} + h((Bc' + D(m'c' + c'\rho) + G\rho)v_{t-1}) + G\varepsilon_t \quad (3)$$

or

$$z_t = T'_a + T'_m z_{t-1} + h(T'_c v_{t-1}) + G\varepsilon_t$$

One can write the above form as a mapping, or T-map, from the PLM to the ALM:

$$T \begin{pmatrix} a' \\ m' \\ h(c') \end{pmatrix} = \begin{pmatrix} A + Ba' + D(I + m')a' \\ Bm' + Dm'^2 + F \\ h(Bc' + D(m'c' + c'\rho) + G\rho) \end{pmatrix} \quad (4)$$

The condition for an equilibrium to be (locally) stable under a simple learning rule, such as ordinary least squares, is known as Expectational Stability, or E-stability. Consider the following ordinary differential equation (ODE):

$$\frac{d\phi}{d\tau} = T(\phi) - \phi \quad (5)$$

where  $T$  is the mapping from the perceived law of motion,  $\phi$ , to the implied actual law of motion,  $T(\phi)$  and  $\tau$  denotes "notional" or "artificial" time. In this case,  $T(\phi)$  is represented by equation (4) and

$$\phi = \begin{pmatrix} a' \\ m' \\ h(c') \end{pmatrix}$$

Evans and Honkapohja (2001) define an equilibrium (fixed point of the ODE) to be E-stable if the ODE is stable when evaluated at the equilibrium values.<sup>4</sup> For stability, the eigenvalues of the Jacobian matrix of the ODE must have negative real parts. It is commonly known that (Marcet and Sargent (1989) and Evans and Honkapohja (2001)), under least squares learning, an E-stable equilibrium is (locally) learnable. Learnability of an equilibrium may be regarded as a necessary condition for the relevance of that equilibrium.

## 2.2 Constrained Forecasting and the RPE

Agents may be unable, or unwilling, to use the above correctly specified forecasting model (equation (2)) due to computational costs, degrees of freedom restrictions, poor information, or just a poor understanding of the economy. In this case, agents may

---

<sup>4</sup>To compute the stability conditions, the ODE must be vectorized (See chapter 10 of Evans and Honkapohja (2001)). For any matrix  $X$ , let  $vec(X)$  denote the vector obtained by stacking the columns of  $X$ .

learn using a restricted forecasting model of the following form:

$$z_t = \hat{a}' + \hat{m}'z_{t-1} + h(\hat{c}'v_{t-1}) + \varepsilon_t \quad (6)$$

where some, or all of the  $\hat{m}$ 's, are equal to zero and the corresponding components of  $\hat{c}$  are also equal to zero. The expectations from the restricted PLM, equation (6), feedback into the structural equation (1) to give the ALM:

$$z_t = T'_a + T'_m z_{t-1} + h(T'_c v_{t-1}) + G\varepsilon_t \quad (7)$$

where

$$\begin{aligned} T'_a &= A + B\hat{a}' + D(I + \hat{m}')\hat{a}' \\ T'_m &= B\hat{m}' + D\hat{m}'^2 + F \\ T'_c &= h(B\hat{c}' + D(\hat{m}'\hat{c}' + \hat{c}'\rho) + G\rho) \end{aligned}$$

represent the T-maps for  $a'$ ,  $m'$ , and  $c'$  under a correctly specified PLM to the ALM. Note that the ALM here is in the same form as it is in equation (3), however, some of the  $\hat{m}$ 's and  $\hat{c}$ 's are constrained to be zero. Since this ALM has a complete VAR form, one cannot present a T-mapping from the restricted PLM to this ALM. This ALM must be projected into the same form as the restricted PLM. However, at its current form, a single linear projection of the ALM is impossible, so I will state the PLM in a different form.

Recall the PLM:

$$z_t = \hat{a}' + \hat{m}'z_{t-1} + h(\hat{c}'v_{t-1}) + \varepsilon_t$$

where some, or all, of the coefficients in  $\hat{m}$  are zero. The zero's in  $\hat{m}$  can be eliminated by assuming that the agents estimate each variable separately:

$$z_{it} = \hat{a}_i + w'_{it}b_i + h(v'_{i,t-1}c_i) + \varepsilon_{it} \quad (8)$$

where  $w_{it}$  and  $v_{i,t-1}$  are  $(k_i \times 1)$  vectors of predetermined explanatory variables agents use to estimate variable  $i$ . The vectors  $b_i$  and  $c_i$  are correspondingly represented by the  $i^{th}$  column of the matrices  $\hat{m}$  and  $\hat{c}$  where all of the zero's are removed. The  $n$  regressions can be written as a system of seemingly unrelated regressions (SUR) in the following manner:

$$z_t = \mathcal{A} + \mathcal{W}'_t \mathcal{B} + \varepsilon_t \quad (9)$$

where<sup>5</sup>

$$\begin{aligned} \mathcal{A} &= (\hat{a}_1 \quad \hat{a}_2 \quad \cdots \quad \hat{a}_n)' \\ \mathcal{W}'_t &= [\mathcal{W}'_{1t}, h(\mathcal{W}'_{2t})] \end{aligned}$$

---

<sup>5</sup>If  $b_i$  is an empty set, then the  $i^{th}$  row of both  $\mathcal{W}'_{1t}$  and  $\mathcal{W}'_{2t}$  is a zero vector.  $w'_{i+1,t}$  (and  $v'_{i+1,t-1}$ ) will appear one row below starting in the first column that has only zeros in the rows above.

$$\mathcal{W}'_{1t} = \begin{bmatrix} w'_{1t} & 0' & \cdots & 0' \\ 0' & w'_{2t} & \cdots & 0' \\ \vdots & \vdots & \ddots & \vdots \\ 0' & 0' & \cdots & w'_{nt} \end{bmatrix}$$

$$\mathcal{W}'_{2t} = \begin{bmatrix} v'_{1,t-1} & 0' & \cdots & 0' \\ 0' & v'_{2,t-1} & \cdots & 0' \\ \vdots & \vdots & \ddots & \vdots \\ 0' & 0' & \cdots & v'_{n,t-1} \end{bmatrix}$$

$$\mathcal{B} = (\mathcal{B}'_1, h(\mathcal{B}'_2))'$$

$$\mathcal{B}_1 = (b'_1 \quad b'_2 \quad \cdots \quad b'_n)'$$

and

$$\mathcal{B}_2 = (c'_1 \quad c'_2 \quad \cdots \quad c'_n)'$$

$\mathcal{W}'_t$  is  $(n \times l)$ ,  $\mathcal{B}$  is  $(l \times 1)$ ,  $k = \sum_{i=1}^n k_i$ , and  $l = k + h(k)$ . The vector  $\mathcal{B}$  is equivalently represented by  $\begin{pmatrix} \text{vec}(\hat{m}) \\ \text{vec}(\hat{c}) \end{pmatrix}$  with all the zeros removed. This form is equivalent to the original form and will help give a T-map from the PLM to the projected ALM.

The next step is to put the ALM into a similar form. Consider the current form of the ALM:

$$z_t = T'_a + T'_m z_{t-1} + h(T'_c v_{t-1}) + G\varepsilon_t \quad (10)$$

This can first be put into an alternative form by taking the transpose:

$$z'_t = T_a + z'_{t-1} T_m + h(v'_{t-1} T_c) + \varepsilon'_t G' \quad (11)$$

One can now vectorize this system of equations to make it similar to the SUR form of the PLM:

$$z_t = T'_a + \mathcal{X}'_t T_x + G\varepsilon_t \quad (12)$$

where

$$\mathcal{X}'_t = ((I \otimes z_{t-1})', h((I \otimes v_{t-1})'))$$

$$T_x = \begin{pmatrix} \text{vec}(T_m) \\ h(\text{vec}(T_c)) \end{pmatrix}$$

Note that  $\mathcal{X}'_t$  is similar to the matrix  $\mathcal{W}'_t$  where here every variable is included in every separate regression. Now that the PLM and ALM are in SUR form, one can linearly project this system of seeming unrelated regressions into another class of seemingly unrelated regressions of the form of the PLM.

It is important to note that the ALM of equation (12) is equivalent to equation (4). Evans and Honkapohja (2001, pp 140-141) discuss that E-stability is invariant to 1 to



1 transformations. Therefore, E-stability conditions will not change when the ALM is put into SUR form. The next step of linearly projecting the ALM into the same class as the PLM (in SUR form) will also not affect the E-stability conditions.

Suppose that a forecast of some  $Y_t$  is restricted to be a linear forecast of some matrix,  $X_t$ , which is a partitioned matrix of block diagonal matrices representing the variables used for each separate regression. Then, the forecast of  $Y_t$  is the following:

$$Y_t = p(\mathcal{A}) + X_t'p(\mathcal{B}) \quad (13)$$

Using this forecast, it follows that

$$p(\mathcal{A}) = EY_t - EX_t'p(\mathcal{B}) \quad (14)$$

which will prove to be helpful when solving for the T-map of the constant terms.  $p(\mathcal{A})$  and  $p(\mathcal{B})$  are chosen such that the forecast errors are uncorrelated with the SUR matrix of regressors:

$$E[X_t(Y_t - (p(\mathcal{A}) + X_t'p(\mathcal{B})))] = 0 \quad (15)$$

This linear projection will produce the least squares orthogonality condition such that agents using the matrix  $X_t$ , will be unaware they may not be using all relevant information when forecasting  $Y_t$ .

For the PLM and ALM in SUR form, we can write the forecast as follows:

$$z_t = p(\mathcal{A}) + \mathcal{W}_t'p(\mathcal{B}) \quad (16)$$

or

$$T_a' + \mathcal{X}_t'T_x + G\varepsilon_t = p(\mathcal{A}) + \mathcal{W}_t'p(\mathcal{B}) \quad (17)$$

The constants can be solved by using the expectations operator:

$$p(\mathcal{A}) = T_a' + E(\mathcal{X}_t')T_x - E(\mathcal{W}_t')p(\mathcal{B}) \quad (18)$$

Next, the forecast error must be uncorrelated with the SUR matrix  $\mathcal{W}_t$ :

$$E[\mathcal{W}_t(T_a' + \mathcal{X}_t'T_x + G\varepsilon_t - p(\mathcal{A}) - \mathcal{W}_t'p(\mathcal{B}))] = 0 \quad (19)$$

Using the solution for  $p(\mathcal{A})$  from (18) this turns out to be:

$$E[\mathcal{W}_t((\mathcal{X}_t' - E(\mathcal{X}_t'))T_x + G\varepsilon_t - (\mathcal{W}_t' - E(\mathcal{W}_t'))p(\mathcal{B}))] = 0$$

Doing some algebra gives

$$cov[\mathcal{W}_t, \mathcal{X}_t']T_x = var[\mathcal{W}_t]p(\mathcal{B}) \quad (20)$$

and this can be further reduced to<sup>6</sup>

$$p(\mathcal{B}) = \Omega^{-1} cov[\mathcal{W}_t, \mathcal{X}_t'] T_x \quad (21)$$

where

$$\Omega = var[\mathcal{W}_t]$$

The technique used to solve for the moment matrices following Adam, Evans, and Honkapohja (2005) is presented in appendix A. Equation (21) demonstrates that the solution to  $p(\mathcal{B})$  has the same form as any other linear regression: it is equal to the inverse of the variance matrix times the covariance of the matrix of regressors and the endogenous variables which are being estimated. Note that if  $\mathcal{W}_t = \mathcal{X}_t$ , then we have the following:

$$\begin{aligned} p(\mathcal{B}) &= T_x \\ p(\mathcal{A}) &= T'_a \end{aligned}$$

$p(\mathcal{B})$  and  $p(\mathcal{A})$  have the same form as the SUR vectors  $\mathcal{B}$  and  $\mathcal{A}$  respectively and represent the projected T-map (or the "p-map") which maps the restricted PLM to the projected ALM. The Projected ALM is thus the following:

$$z_t = p(\mathcal{A}) + \mathcal{W}'_t p(\mathcal{B}) + \varepsilon_t \quad (22)$$

Corollary 1 shows how one can find the E-stability conditions of the RPE using the p-map.

Corollary 1. *A RPE is E-stable if the standard E-stability ODE*

$$\frac{d\phi}{d\tau} = \begin{pmatrix} p(\mathcal{A}) \\ p(\mathcal{B}) \end{pmatrix} - \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix}$$

*is stable at the RPE.*

The proof of this corollary is trivial and thus not given. The proof follows the fact that the E-stability conditions for an equilibrium are the same under the standard form or the SUR form. The restricted ALM in the standard form is transformed to SUR form and then projected to the same form as the PLM in SUR form. The transformation to the SUR and the linear projection do not change the E-stability conditions as they are 1 to 1 transformations. As the agents are learning using least squares, the standard E-stability ODE can be used to determine whether a RPE is locally learnable.

---

<sup>6</sup>It is shown in appendix B, after the moment matrices are defined in appendix A, that if the model is stationary, then  $\Omega$  is invertible.

### 2.3 Cobweb Example

The SUR transformation is most convenient when agents must forecast  $n > 1$  variables. Using the SUR form, one needs to put the model in the form above and then do one linear projection to get the projected ALM. One can obtain the projected ALM without using the SUR technique. However, this can be quite cumbersome with  $n > 1$  variables. This requires linearly projecting each component of the ALM ( $n$  total) separately which is, in essence, treating each variable as a unrelated regression. Therefore, for many models, it will be easier for one to put the PLM and ALM into the SUR forms and then do a single linear projection.

I present a misspecification of a cobweb model as presented in pages 318-320 of Evans and Honkapohja (2001) as an example of how to use the SUR technique.<sup>7</sup> Consider a simple cobweb model in the reduced form:

$$p_t = \mu + \alpha E_{t-1}^* p_t + \gamma' w_{t-1} + \eta_t \quad (23)$$

where  $w_{t-1}$  is a vector of observable exogenous variables and  $\eta_t$  is an unobservable white noise shock independent of the  $w_t$  process. Furthermore,  $w_t$  follows a stationary VAR form. Suppose that agents misspecify this model assuming that  $p_t$  can be accurately forecasted using  $w_{1,t-1}$  where  $w'_t = (w'_{1t}, w'_{2t})$  and  $\gamma' = (\gamma'_1, \gamma'_2)$ . The restricted PLM is the following:

$$p_t = a + c' w_{1,t-1} + \varepsilon_t \quad (24)$$

where  $\varepsilon_t$  is believed to be white noise. The ALM is obtained by taking expectations of equation (24), and plugging this into equation (23) to get:

$$p_t = (\mu + \alpha a) + (\alpha c' + \gamma'_1) w_{1,t-1} + \gamma'_2 w_{2,t-1} + \eta_t$$

Since  $w_t$  follows a stationary VAR, the ALM can be put into a VAR form:

$$\begin{pmatrix} p_t \\ w_t \end{pmatrix} = \begin{pmatrix} \mu + \alpha a \\ \phi \end{pmatrix} + \begin{pmatrix} (\alpha c' + \gamma'_1, \gamma'_2) \\ \rho' \end{pmatrix} w_{t-1} + \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix}$$

Notice that I do not include  $p_{t-1}$  on the right hand side and impose zeros as this will complicate the math and give the same results below. Next, put the ALM into SUR form:

$$\begin{pmatrix} p_t \\ w_t \end{pmatrix} = \begin{pmatrix} \mu + \alpha a \\ \phi \end{pmatrix} + (I \otimes w_{t-1})' \begin{pmatrix} \alpha c + \gamma_1 \\ \gamma_2 \\ \text{vec}(\rho) \end{pmatrix} + \begin{pmatrix} \eta_t \\ \zeta_t \end{pmatrix}$$

---

<sup>7</sup>This is a univariate model, however, it can also be expressed as a multivariate model as shown below.

Assume that agents know the true process of  $w_t$  to get the PLM in SUR form:<sup>8</sup>

$$\begin{pmatrix} p_t \\ w_t \end{pmatrix} = \begin{pmatrix} a \\ \phi \end{pmatrix} + \begin{pmatrix} w_{1,t-1} & 0 \\ 0 & (I \otimes w_{t-1})' \end{pmatrix} \begin{pmatrix} c \\ vec(\rho) \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \zeta_t \end{pmatrix}.$$

The variance of the SUR matrix of regressors is:

$$\Omega = \begin{pmatrix} \Phi_{11} & 0 \\ 0 & (I \otimes \Phi) \end{pmatrix}$$

and the covariance matrix of the SUR matrix of regressors and the full information matrix is:

$$cov[\mathcal{W}_t, \mathcal{X}'] = \begin{pmatrix} \Phi_{11} & \Phi_{12} & 0 \\ 0 & 0 & (I \otimes \Phi) \end{pmatrix}$$

where

$$Var \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} = \Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12} & \Phi_{22} \end{pmatrix}$$

Using equation (21), the p-mapping from the PLM to the projected ALM is

$$\begin{aligned} p(a) &= \mu + \alpha a \\ p(c) &= \alpha c + \gamma_1 + \Phi_{11}^{-1} \Phi_{12} \gamma_2 \end{aligned}$$

and the RPE is

$$\begin{pmatrix} \bar{a} \\ \bar{c} \end{pmatrix} = \begin{pmatrix} (1 - \alpha)^{-1} \mu \\ (1 - \alpha)^{-1} (\gamma_1 + \Phi_{11}^{-1} \Phi_{12} \gamma_2) \end{pmatrix}$$

as shown in Evans and Honkapohja (2001).

This example shows how the SUR technique may be quite cumbersome and unnecessary in a univariate model. However, in a multivariate model, the techniques used in Evans and Honkapohja (2001) and Adam, Evans, and Honkapohja (2005) become increasingly more computational versus the SUR technique. For example, if the above cobweb model is multivariate where  $p_t$  is a vector of prices, then each price could be misspecified in the following PLM:

$$p_t = a + \tilde{c}' w_{t-1} + \varepsilon_t$$

where some of the  $\tilde{c}$ 's are constrained to be zero. Other projection techniques would require a linear projection of each separate equation. The SUR technique only requires transforming the PLM and ALM into a SUR form and then doing a single linear projection.

---

<sup>8</sup>A misspecification on  $w_t$  would have no effect on the RPE of  $p_t$  as expectations of  $w_t$  are not in the linear model. However, if the structural model also included  $E_{t-1} p_{t+1}$ , expectations of  $w_t$  would enter the ALM and thus affect the RPE of  $p_t$ .

### 3 Theory

The second purpose of this study is to study restricted perceptions in the New Keynesian Monetary Model. In this section, I will develop the theory necessary to analyze RPE's under several interest rate rules. First, I will present the model of interest.

#### 3.1 Monetary Model

The model used in this paper is the New Keynesian Monetary Model with inertia in both the IS and AS curves:

$$\begin{aligned} IS & : x_t = -\phi(i_t - E_t\pi_{t+1}) + \delta E_t x_{t+1} + (1 - \delta)x_{t-1} + g_t \\ AS & : \pi_t = \beta(\gamma E_t\pi_{t+1} + (1 - \gamma)\pi_{t-1}) + \lambda x_t + u_t \end{aligned}$$

where  $x_t$  is the proportional output gap,  $\pi_t$  is the inflation rate, and  $g_t$  and  $u_t$  are independent, exogenous, stationary, zero mean AR(1) shocks with damping parameters  $0 \leq \rho_g < 1$  and  $0 \leq \rho_u < 1$  respectively.

The IS curve is obtained from a linearized model of consumer optimization amended to include inertia. Inertia is included as it has been argued<sup>9</sup> that consumption smoothing is not perfect and does not follow a random walk as suggested by Hall (1978). Inertia can be captured via habit formation as in Fuhrer (2000) and Smets (2003) or from the existence of rule of thumb consumers as in Cambell and Mankiw (1989), Mankiw (2000), and Amato and Laubach (2003). Here,  $\phi$  relates to the intertemporal elasticity of substitution of the representative household and  $0 < \delta < 1$  is the inertial term due to backward looking consumers.

The AS curve is the forward looking Phillips curve amended to include inertia as it has been found by Fuhrer and Moore (1995) and Gali and Gertler (1999) to be important when estimating the Phillips curve.<sup>10</sup> This can be captured from habit formation as in McCallum and Nelson (1999a) and Fuhrer (2000), rule of thumb price setters as in Gali and Gertler (1999) and Amato and Laubach (2003), or price indexation as in Christiano et al. (2001). Here,  $0 < \beta < 1$  is the discount factor,  $\lambda$  is the degree of price stickiness, and  $0 < \gamma < 1$  is the inertial term due to backward looking price setters.

The above structural model is closed by specifying a policy rule describing how interest rates are set. It has been shown that the region and nature of a model's indeterminacy depends on the nature of this policy rule. I will analyze three of the four policy rules as presented in Bullard and Mitra (2002) in this hybrid New Keynesian

---

<sup>9</sup>A long list of papers include Mankiw et al. (1985), Deaton (1992), Shea (1995), and Parker (1999).

<sup>10</sup>The existence of backward looking agents in the Phillips curve is a controversial topic as several authors including Wolman (1999) and Dotsey (2002) show the estimated model used by Gali and Gerler (1999) was biased for a positive result of backward looking behavior.

Model:<sup>11</sup>

$$\begin{aligned}
 PR_1 & : i_t = a_\pi \pi_{t-1} + \alpha_x x_{t-1} \\
 PR_2 & : i_t = \alpha_\pi E_t \pi_{t+1} + \alpha_x E_t x_{t+1} \\
 PR_3 & : i_t = \alpha_\pi E_t \pi_t + \alpha_x E_t x_t
 \end{aligned}$$

In each policy rule,  $\pi_t$  can be interpreted as the deviation from its target. As usual the intercepts of the above equations have been omitted for convenience. I assume throughout that  $\alpha_\pi, \alpha_x \geq 0$  so that the  $a_\pi \pi_{t-1}$  term in  $PR_1$  is the degree to which the monetary policy authority raises nominal interest rates in response to a upward deviation of  $\pi_{t-1}$  from its target.

## 4 Results

The model under the three policy rules will be analyzed using three different calibrations of the parameters in the IS-AS curves, however, the discussion will be concerned with "areas" which each calibration has in common (discussed below). Consider the following calibrations:

Table 1. Calibrations

Author(s)	$\phi$	$\lambda$
W	1/.157	.024
CGG	4	.075
MN	.164	.3

These calibrations are due to Woodford (1999), Clarida, Gali, and Gertler (2000), and McCallum and Nelson (1999b). Several papers have discussed determinacy and learnability of this model under Rational Expectations (Bullard and Mitra 2002, 2003) and Stationary Sunspot Equilibria (Honkapohja and Mitra 2004, Evans and McGough 2004a, 2004b), however, the learnability of RPE's in such a model has not yet been discussed.

Evans and McGough (2004b) assume that the parameters for backward looking behavior,  $\delta$  and  $\gamma$ , are set to 0.5. However, it turns out that the characteristics of the below described areas changes with the level of backward looking behavior.<sup>12</sup> Therefore, I consider several combinations of these two parameters. Following Evans and McGough, I also set  $\beta = 1$ .

---

<sup>11</sup>Taylor's original policy rule where the central bank adjusts the nominal interest rate due to contemporaneous data is not included as McCallum (1999) argues that this rule is non-operational. Stability results for this Taylor rule are equivalent to the results under  $PR_3$ .

<sup>12</sup>Bullard and Mitra (2003) discuss varying these parameters, however, they only consider the determinate region D (discussed below). They find that these parameters do not affect the E-stability results in this region.

Previously, papers have looked at the calibrations separately, however, it turns out that this is not necessary. Consider the following figures that show a change in determinacy as the policy parameters,  $\alpha_x$  and  $\alpha_\pi$ , vary for each calibration.

Figures 1, 2, 3 about here

It turns out that, for the 3 calibrations above, the areas labeled A-D usually share the same determinacy and learnability properties for any given policy rule and level of backward looking agents.<sup>13</sup> Therefore, to simplify matters, I will discuss the determinacy and learnability properties that these calibrations share for each area.

There are a few general results that occur under all three policy rules. First, the main result is that all 16 possible classes of equilibria (REE and 15 possible forms of RPE) generally share E-stability properties.<sup>14</sup> Therefore, under the above model, forecast misspecification does not cause instability under learning and the results of Bullard and Mitra (2002, 2003) are robust under model misspecification. Second, if an REE is stable under learning with no inertia, then determinacy properties of the REE do not change when there are backward looking agents. Third, if there is no learnable REE in a determinate or indeterminate area, then the determinacy properties of the REE will change as more agents become backward looking.

#### 4.1 Misspecified Forecasting Rule

Previous literature has assumed that agents learn the equilibrium using a forecasting model consistent with the REE. Why should we assume that agents take into account the relationship between inflation and the output gap when forecasting these variables? It is possible that agents may choose to make their forecast without using all the variables due to lack of information, poor specification testing abilities, or the fully specified forecasting rule may be too costly. For this section, I will follow the results in Adam (2002) and assume that agents form expectations of each variable using a univariate AR(1) forecasting rule of the form:

$$\begin{aligned}x_t &= \hat{a}_1 + \hat{m}_1 x_{t-1} + \hat{v}_{1t} \\ \pi_t &= \hat{a}_2 + \hat{m}_2 \pi_{t-1} + \hat{v}_{2t}\end{aligned}$$

As the results for every class of RPE only differs in the non-learnable regions (some solutions may be explosive while others may have a non-learnable stationary solution), I discuss this particular class as a general result. The above technique will be used to form the T-map and ODE in order to determine E-stability of the RPE.

---

<sup>13</sup>For some levels of backward looking behavior, there can be multiple determinacy results within a single area. This is discussed below.

<sup>14</sup>When  $\gamma$  and  $\delta$  are near zero, there is a small sliver near  $\phi_\pi = 1$  and  $\phi_x = 0$  in area D (under policy rules 1 and 3) where the REE and several other RPE are explosive and other RPE are E-stable.

## 4.2 Policy Rule 1

Recall that policy rule 1 is a backward looking Taylor rule of the following form:

$$PR_1 : i_t = a_\pi \pi_{t-1} + \alpha_x x_{t-1}$$

This policy rule has been analyzed by Bullard and Mitra (2002,2003) and Evans and McGough (2004a). The figure in Bullard and Mitra (2002) is similar to Figures 1-3 except the vertical line at  $\phi_\pi = 1$  is replaced by the line (where  $\beta = .99$ )

$$\phi_\pi + \frac{(1 - \beta)}{\lambda} \phi_x = 1$$

representing the Taylor Principle with  $\beta = .99$ .<sup>15</sup> Under a situation where no agents are backward looking, areas A and D are determinate, however, the REE is learnable only within area D. Therefore, Bullard and Mitra find that such a rule would not be advisable as agents may fail to coordinate on the equilibrium, even though it is unique. As area B is explosive, following the Taylor principle does not guarantee existence or learnability.

Tables 2 shows the determinacy and stability results under several combinations of  $\gamma$  and  $\delta$  for the REE while table 3 shows the E-stability results<sup>16</sup> for the RPE. For the REE, the text shows determinacy for each given  $(\gamma, \delta)$  and the superscript shows if there is a learnable equilibrium. The text in table 3 shows learnability results for the RPE. Det means that the REE is determinate and therefore, there is a unique stationary REE. EX means that there is exist no stationary REE or RPE. 1 or 2 means that the model is indeterminate of order 1 or 2 respectively. In some cases, there may be two determinacy results in one area for some  $(\gamma, \delta)$ . I denote "1, EX" for the model being indeterminate of order 1 for small  $\phi_x$  and explosive for larger levels of  $\phi_x$  within a particular area. "EX, Det" is the explosive "sliver" discussed in Evans and McGough (2004b) where no stationary solution exists in area D near  $\phi_\pi = 1$  and  $\phi_x = 0$  and a unique stationary solution exists when these two parameters are slightly increased. This "sliver" is quite small relative to the size of area D. U means that there is no learnable equilibrium while E is where there exists an E-stable REE or RPE.

Table 2. Results of REE under Policy Rule 1

<sup>15</sup>In this model with inertia, the Taylor Principle is represented by  $\phi_\pi > 1$  as  $\beta = 1$ . When  $\beta = .99$  the two figures are the same, however, there are no major changes in the results below.

<sup>16</sup>There is no general method to determine determinacy under misspecified expectations. This is a topic of current research.



$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	<i>EX</i>	<i>EX</i>	<i>EX</i>	<i>EX, Det<sup>E</sup></i>
0.1	0.5	<i>EX</i>	<i>EX</i>	$1^U, EX$	<i>Det<sup>E</sup></i>
0.1	0.9	<i>EX</i>	<i>EX</i>	$1^U, EX$	<i>Det<sup>E</sup></i>
0.5	0.5	<i>Det<sup>U</sup></i>	<i>EX</i>	$1^U$	<i>Det<sup>E</sup></i>
0.9	0.1	<i>Det<sup>U</sup></i>	<i>EX</i>	$1^U$	<i>Det<sup>E</sup></i>
0.9	0.5	<i>Det<sup>U</sup></i>	<i>EX</i>	$1^U$	<i>Det<sup>E</sup></i>
0.9	0.9	<i>Det<sup>U</sup></i>	<i>EX</i>	$1^U$	<i>Det<sup>E</sup></i>

Table 3. Results of RPE under Policy Rule 1

$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	<i>EX</i>	<i>EX</i>	<i>EX</i>	<i>E</i>
0.1	0.5	<i>EX</i>	<i>EX</i>	<i>EX</i>	<i>E</i>
0.1	0.9	<i>EX</i>	<i>EX</i>	<i>EX</i>	<i>E</i>
0.5	0.5	<i>EX</i>	<i>EX</i>	<i>EX</i>	<i>E</i>
0.9	0.1	<i>EX</i>	<i>EX</i>	<i>U</i>	<i>E</i>
0.9	0.5	<i>EX</i>	<i>EX</i>	<i>U</i>	<i>E</i>
0.9	0.9	<i>EX</i>	<i>EX</i>	<i>U</i>	<i>E</i>

The stability results in Bullard and Mitra (2002) continue to follow in the model with inertia as shown in Bullard and Mitra (2003).<sup>17</sup> It turns out that these results are also robust when we are considering RPE rather than REE. For areas A and C, by increasing inertia in the AS (increasing  $\gamma$ ), one can see that the REE becomes explosive. The determinate and learnable equilibrium in area D continues to be learnable for all  $(\gamma, \delta)$  except for a small area when  $\gamma = \delta = 0.1$ . The RPE within this explosive "sliver" discussed in Evans and McGough (2004b) not only exists, but is learnable. For the given restricted forecasting model, the RPE is learnable for any  $(\gamma, \delta)$  and  $(\phi_\pi, \phi_x)$  within area D.<sup>18</sup>

### 4.3 Policy Rule 2

Recall that policy rule 2 is forward looking Taylor rule of the form:

$$i_t = \phi_\pi E_t \pi_{t+1} + \phi_x E_t x_{t+1}$$

Here, the central bank adjusts the interest rate based on its expectations of next period's inflation and output gap. Under least squares learning, private agents are assumed to recursively estimate the parameters of their PLM and form expectations of inflation and output given these estimates. One could assume that the private agents and central

<sup>17</sup>The results are only reported under policy rule 2 as the general result is the same for all 3 policy rules. They did not, however, report that determinacy changes as inertia increases.

<sup>18</sup>This is not a general result for all RPE. This explosive sliver exists in 6 of the 16 different classes of equilibria.

bank have different expectations as the assumption of rational expectations has been relaxed.<sup>19</sup> To simplify matters, I will assume that the expectations of private agents and the central bank are homogeneous as it has been done in Bullard and Mitra (2002). This can be justified as the central bank setting interest rates as a reaction to private sector forecasts as in Bernanke and Woodford (1997).

When there are no backward looking agents, Bullard and Mitra (2002) showed that the REE is determinate in area D and indeterminate in areas A, B, and C. The REE are learnable in area D and area B. Honkapohja and Mitra (2004) and Evans and McGough (2004a) have also shown that there are learnable sunspot equilibria in area B. This suggests that if the central bank is too aggressive with the output gap, agents may coordinate to an undesirable equilibrium.

The determinacy and learnability results of the REE and RPE under policy rule 2 are shown in tables 4 and 5. Here, as indeterminacy does not imply non-learnability,  $1^E$  is entered when the model is indeterminate of order one and there is a learnable equilibrium. Once again, the stability results in Bullard and Mitra (2002) are robust to model misspecification under policy rule 2. Increasing inertia leads to area C becoming explosive and area A becoming determinate. Determinacy properties of the areas with learnable solutions (B and D) do not change with inertia changes. Here, the RPE and REE share E-stability properties as it was in policy rule 1.

Table 4. Results of REE under Policy Rule 2

$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	$Det^U$	$1^E$	$EX$	$Det^E$
0.1	0.5	$Det^U$	$1^E$	$1^U, EX$	$Det^E$
0.1	0.9	$Det^U$	$1^E$	$1^U, EX$	$Det^E$
0.5	0.5	$2^U$	$1^E$	$1^U$	$Det^E$
0.9	0.1	$2^U$	$1^E$	$1^U$	$Det^E$
0.9	0.5	$2^U$	$1^E$	$1^U$	$Det^E$
0.9	0.9	$2^U$	$1^E$	$1^U$	$Det^E$

Table 5. Results of RPE under Policy Rule 2

$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	$EX$	$E$	$EX$	$E$
0.1	0.5	$EX$	$E$	$EX$	$E$
0.1	0.9	$EX$	$E$	$EX$	$E$
0.5	0.5	$EX$	$E$	$EX$	$E$
0.9	0.1	$U$	$E$	$U$	$E$
0.9	0.5	$U$	$E$	$U$	$E$
0.9	0.9	$U$	$E$	$U$	$E$

<sup>19</sup>Honkapohja and Mitra (2005) studies this model without inertia when the central bank and private agents have different expectations.

#### 4.4 Policy Rule 3

Recall that policy rule 3 is a Taylor rule where the central bank looks at expectations of contemporaneous variables:

$$i_t = \phi_\pi E_t \pi_t + \phi_x E_t x_t$$

I will continue to assume that the central bank uses the private sectors forecasts in this rule. Bullard and Mitra (2002) show that this policy rule, with no backward looking agents, is the only one in which following the Taylor principle will guarantee determinacy and learnability. Therefore, this may be the best rule for a central banker to follow as the other rules may lead to no equilibria (policy rule 1) or learnable sunspot equilibria (policy rule 2). Given the results below, I believe that this policy rule continues to be the best rule even under inertia and restricted perceptions. The results for the policy rule are shown in Tables 6 and 7.

Table 6. Results of REE under Policy Rule 3

$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	<i>EX</i>	<i>Det<sup>E</sup></i>	<i>EX</i>	<i>EX, Det<sup>E</sup></i>
0.1	0.5	<i>EX</i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup>, EX</i>	<i>Det<sup>E</sup></i>
0.1	0.9	<i>EX</i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup>, EX</i>	<i>Det<sup>E</sup></i>
0.5	0.5	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>
0.9	0.1	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>
0.9	0.5	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>
0.9	0.9	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>	<i>1<sup>U</sup></i>	<i>Det<sup>E</sup></i>

Table 7. Results of RPE under Policy Rule 3

$\gamma$	$\delta$	Area A	Area B	Area C	Area D
0.1	0.1	<i>EX</i>	<i>E</i>	<i>EX</i>	<i>E</i>
0.1	0.5	<i>EX</i>	<i>E</i>	<i>EX</i>	<i>E</i>
0.1	0.9	<i>EX</i>	<i>E</i>	<i>EX</i>	<i>E</i>
0.5	0.5	<i>EX</i>	<i>E</i>	<i>EX</i>	<i>E</i>
0.9	0.1	<i>U</i>	<i>E</i>	<i>U</i>	<i>E</i>
0.9	0.5	<i>U</i>	<i>E</i>	<i>U</i>	<i>E</i>
0.9	0.9	<i>U</i>	<i>E</i>	<i>U</i>	<i>E</i>

Increasing inertia leads to the areas of non-learnable equilibria changing determinacy results. In this case, areas A and C become explosive as  $\gamma$  is decreased. Determinacy properties of the areas with learnable REE, with the exception of the explosive sliver at  $(\gamma, \delta) = (0.1, 0.1)$ , do not change with inertia change. E-stability of the REE and RPE are quite similar here like it was in the previous cases: the REE and RPE share stability results except in the explosive sliver where the RPE is learnable and the REE is explosive.<sup>20</sup>

<sup>20</sup>This explosive sliver only exists in 6 of the 16 different classes of models.

## 5 Conclusion

Model misspecification is another way to model bounded rationality. Here agents are boundedly rational as their beliefs reside in their forecasting model even though their activities exist in the economic model. This bounded rationality is due to the fact that forecast errors are uncorrelated with information used in a forecasting model. Therefore, an agent that stepped outside of their incorrectly specified forecasting model could see that they are not using enough information. As no econometric model is ever fully specified, it may be useful to study a further abstraction from rationality where agents learn about economic variables by using incorrectly specified forecasting models.

E-stability, a condition of local learnability of an equilibrium, is found by forming an ordinary differential equation using a mapping from the perceived law of motion (PLM) to the actual law of motion (ALM). However, in a standard form, when the PLM is incorrectly specified, the ALM is not in the same form of the PLM. As there is no mapping from the PLM to the ALM of the same form, the ALM must be projected into the same class of the PLM. I showed that under a fully specified model, the standard form of the ALM can be transformed into a Seemingly Unrelated Regression form by taking the transpose of the ALM and then vectorizing it. It turns out that the E-stability conditions under this form do not change which is important in the transformation of the incorrectly specified PLM. To project the ALM of the incorrectly specified PLM, the PLM is put in the form of a SUR. The ALM is then transformed to this form. Next, the ALM is projected into the same class as the incorrectly specified PLM using a simple linear projection. As agents are still using least squares to learn, the standard E-stability ODE may be used to determine whether equilibria are locally learnable.

Next, I presented an example of Restricted Perceptions in a standard New Keynesian IS-LM model with inertia under three different Taylor rules. The E-stability results given by Bullard and Mitra (2002) are robust under restrictive perceptions. Furthermore, in the policy areas of E-stability, determinacy (for a correctly specified model) does not change when backward-looking behavior increases, and determinacy changes in the non-explosive, E-unstable policy areas when backward looking behavior increases. This suggests that there may be a relationship between determinacy, E-stability, and inertia in the New Keynesian IS-LM model.

In the above model, it turns out that restrictive perceptions are not important for learnability when considering policy making. However, the actual laws of motion of an RPE and REE are different. This will be important when one considers optimal policy as this will be different under the two expectational regimes. In future work, I aim to study optimal monetary policy when the central banker is unaware of which expectational regime is being used by the private agents.

### Appendix A. The Moment Matrices

Next, one must solve for the moment matrices in order for calculation of the T-map. Consider the implied ALM from equation (10):

$$z_t = T'_a + T'_m z_{t-1} + h (T'_c v_{t-1}) + G \varepsilon_t$$

To solve for the variance and covariance matrices above, the first step is to put this equation in a VAR(1) form as follows:

$$\begin{pmatrix} z_t \\ h(v_t) \end{pmatrix} = \begin{pmatrix} T'_a \\ h(0) \end{pmatrix} + \begin{pmatrix} T'_m & h(T'_c) \\ h(0) & h(\rho) \end{pmatrix} \begin{pmatrix} z_{t-1} \\ h(v_{t-1}) \end{pmatrix} + \begin{pmatrix} G \\ h(I) \end{pmatrix} \varepsilon_t$$

or

$$Y_t = T'_A + T'_Y Y_{t-1} + T'_G \varepsilon_t$$

Under the assumption that  $z_t$  follows a stationary process, the variance is:

$$var(Y_t) = T'_Y var(Y_t) T_Y + T'_G var(\varepsilon_t) T_G. \quad (25)$$

The variance can then be solved using one of two methods. First, following the method used in Adam, Evans, and Honkapohja (2005), the variance can be vectorized in the following manner:

$$vec(var(Y_t)) = (I - T'_Y \otimes T'_Y)^{-1} vec(T'_G var(\varepsilon_t) T_G) \quad (26)$$

Knowing the dimension of  $var(Y_t)$ , one can unvectorize equation (26) in order to get  $var(Y_t)$  which will be used in order to calculate  $var(\mathcal{W}_t)$  and  $cov[\mathcal{W}_t, \mathcal{X}']$ . For many models, if the number of variables in  $Y_t$  is greater than 2, the above method becomes computationally impossible. For more complex models, as equation (25) is Lyapunov, the variance and RPE can be estimated numerically. This is a two step process where the RPE is updated using equation (21), and then the variance matrix is updated using equation (25). It will be helpful to write up the variance matrix in a partitioned form:

$$var(Y_t) = \begin{bmatrix} var(z_t) & h(cov(z_t, v_t)) \\ h(cov(z_t, v_t)) & h(var(v_t)) \end{bmatrix}$$

The variance matrix of the SUR information matrix,  $\mathcal{W}_t$ , can be represented in the following manner. It is a partitioned matrix consisting of components of the matrix  $var(Y_t)$  and can be represented in the following manner:

$$\Omega = \begin{bmatrix} var(\mathcal{W}_{1t}) & h(cov(\mathcal{W}_{1t}, \mathcal{W}_{2t})) \\ h(cov(\mathcal{W}_{1t}, \mathcal{W}_{2t})) & h(var(\mathcal{W}_{2t})) \end{bmatrix}$$

or

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{bmatrix}$$

All the components of this symmetric matrix can be expressed using components of the matrix  $var(Y_t)$ . Each partitioned matrix is a block diagonal matrix consisting of

components of  $\text{var}(Y_t)$ . Each matrix takes the following form:<sup>21</sup>

$$\Omega_{jk} = \begin{bmatrix} \Omega_{jk,1} & 0 & \cdots & 0 \\ 0 & \Omega_{jk,2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \Omega_{jk,n} \end{bmatrix}$$

Now let  $(N)_{-i}$  represent the matrix  $N$  where all of the rows and columns of the variables not considered in the regression of variable  $i$  are removed from  $N$ . The blocks of each matrix in the partitioned matrix above thus can be represented in the following manner:

$$\Omega_{jk,i} = \left( \text{var}(Y_t)_{jk} \right)_{-i}$$

The covariance matrix,  $\text{cov}[\mathcal{W}_t, \mathcal{X}']$ , can be represented in a similar manner. Let  $(N)_{-i}^r$  represent the matrix  $N$  where all of the rows of the variables not considered in the regression of variable  $i$  are removed from  $N$ . This covariance matrix can also be represented by a symmetric partitioned matrix consisting of components of the matrix  $\text{var}(Y_t)$  in the following form:

$$\text{cov}[\mathcal{W}_t, \mathcal{X}'] = \begin{bmatrix} \Pi_{11} & h(\Pi_{12}) \\ h(\Pi_{12}) & h(\Pi_{22}) \end{bmatrix}$$

Each matrix in the partitioned matrix can be represented like above:<sup>22</sup>

$$\Pi_{jk} = \begin{bmatrix} \Pi_{jk,1} & 0 & \cdots & 0 \\ 0 & \Pi_{jk,2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \Pi_{jk,n} \end{bmatrix}$$

If all variables are included in the regression for variable  $i$ , then  $\Pi_{jk,i} = \text{var}(Y_t)_{jk}$ . However, if not all of the variables are included, then

$$\Pi_{jk,i} = \left( \text{var}(Y_t)_{jk} \right)_{-i}^r$$

By calculating all of the  $\Pi_{jk,i}$ 's, we can get the representation for  $\text{cov}[\mathcal{W}_t, \mathcal{X}']$ .

## Appendix B. Proof of invertibility of $\Omega$

For the T-map to be well defined, it must be that the matrix  $\Omega$  is invertible. Consider the matrix  $\text{var}(Y_t)$  as defined above. Assume that this matrix is invertible so that the

<sup>21</sup>If  $b_i$  is the empty set, then  $\Omega_{jk,i}$  is the empty set as well. Therefore, the number of blocks in this block diagonal matrix is equal to the number of forecast equations that include lagged variables.

<sup>22</sup>The number of blocks in this matrix is also equal to the number of forecast equations that include lagged variables.

system is stationary. If  $(\text{var}(Y_t))^{-1}$  exists, then we know that

$$\det(\text{var}(Y_t)) \neq 0.$$

Call  $\Phi = \text{var}(Y_t)$ . This positive definite, symmetric matrix has a unique representation of the following:

$$\Phi = ADA'$$

where  $A$  is a lower diagonal matrix with 1's on the diagonal and  $D$  is a diagonal matrix. Therefore, the determinant of  $\Phi$  can be represented as follows:

$$\begin{aligned} \det \Phi &= \det A * \det D * \det A' \\ &= 1 * \det D * 1 \\ &= \det D \end{aligned}$$

Since  $D$ , is a diagonal matrix,

$$\det \Phi = \prod_{j=1}^n d_{jj}$$

Now, consider,  $\Omega$  which recall is represented as follows:

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$$

Recall that a matrix is singular if and only if a subset of its rows (or columns) is linearly dependent. Therefore, one way to prove invertability of a matrix is to perform elementary operations on the rows and columns of the matrix and show that the resulting matrix is not singular (See Horn and Johnson (1985)). As all the blocks of this matrix are block diagonal matrices of the same dimensions, the matrix can be arranged in the following manner:<sup>23</sup>

$$\hat{\Omega} = \begin{bmatrix} \hat{\Omega}_1 & 0 & \cdots & 0 \\ 0 & \hat{\Omega}_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \hat{\Omega}_n \end{bmatrix}$$

where

$$\hat{\Omega}_i = \begin{bmatrix} \Omega_{11,i} & \Omega_{12,i} \\ \Omega_{21,i} & \Omega_{22,i} \end{bmatrix}$$

which is equivalent to

$$\hat{\Omega}_i = (\Phi)_{-i}$$

Here,  $(N)_{-i}$  represents the original matrix where the rows and columns of the non-represented variables (in the regression for variable  $i$ ) are excluded. Since  $\Phi$  is a symmetric matrix,

---

<sup>23</sup>The number of blocks in this matrix is actually equal to the number of forecast equations that include lagged variables. This does not affect this proof.

the triangular factorization is similar. We can represent  $\hat{\Omega}_i$  in the following manner:

$$\hat{\Omega}_i = (A)_{-i} (D)_{-i} (A')_{-i}$$

Therefore, it clearly follows that

$$\det \hat{\Omega}_i = \det (D)_{-i}$$

Since  $\det D \neq 0$ , it follows that  $\det (D)_{-i} \neq 0$  and

$$\det \hat{\Omega}_i \neq 0$$

for all  $i$ . Therefore,

$$\begin{aligned} \det \hat{\Omega} &= \prod_{i=1}^n \det \hat{\Omega}_i \\ &\neq 0 \end{aligned}$$

Since  $\hat{\Omega}$  is not singular, it follows that  $\Omega^{-1}$  exists.

## References

- [1] Adam, K., Evans, G., Honkapohja, S. (2005) 'Are Stationary Hyperinflation Paths Learnable?', *Journal of Economic Dynamics and Control*, Forthcoming.
- [2] Amato, J., Laubach, T. (2003) 'Rule of Thumb Behaviour and Monetary Policy', *European Economic Review* 47, 791-831.
- [3] Bernanke, B., Woodford, M. (1997) 'Inflation Forecasts and Monetary Policy', *Journal of Money, Credit, and Banking* 24, 653-684.
- [4] Blanchard, O., Kahn, C., (1980) 'The Solution of Linear Difference Models Under Rational Expectations', *Econometrica* 48, 1305-1311.
- [5] Bullard, J., Mitra, K. (2002) 'Learning About Monetary Policy Rules', *Journal of Monetary Economics* 49, 1105-1129.
- [6] Bullard, J., Mitra, K. (2003) 'Determinacy, Learnability, and Monetary Policy Inertia', *Federal Reserve of St. Louis Working Paper*, 2000-030B.
- [7] Cambell, J., Mankiw, N.G. (1989) 'Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence', In: Blanchard, O., Fischer, S. (Eds.), *NBER Macroeconomics Annual*. MIT Press, Cambridge, MA, pp. 185-214.
- [8] Christiano, L., Eichenbaum, M., Evans, C. (2001) 'Nominal Rigidities and the Dynamics Effects of a Shock to Monetary Policy', Working Paper no. 01-07, *Federal Reserve Bank of Cleveland*.



- [9] Clarida, R., Gali, J., Gertler, M. (2000) 'Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory', *Quarterly Journal of Economics*, 115, 147-180.
- [10] Deaton, A. (1992) *Understanding Consumption*. Clarendon Press, Oxford.
- [11] Dotsey, M. (2002) 'Pitfalls in Interpreting Tests of Backward-Looking Pricing in New Keynesian Models', *Federal Reserve Bank of Richmond Economic Quarterly* 88, 37-50.
- [12] Evans, G.W., Honkapohja, S., (2001) *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, NJ.
- [13] Evans, G.W., McGough, B., (2004) 'Monetary Policy, Indeterminacy and Learning', *Journal of Economic Dynamics and Control*, Forthcoming.
- [14] Evans, G.W., McGough, B., (2004) 'Monetary Policy and Stable Indeterminacy with Inertia', *Economic Letters*, Forthcoming.
- [15] Fuhrer, J. (2000) 'Habit Formation in Consumption and its Implications for Monetary-policy Models', *American Economic Review* 90, 367-390.
- [16] Fuhrer, J., Moore, G. (1995) 'Inflation Persistence', *Quarterly Journal of Economics* 110, 127-159.
- [17] Gali, J. and Gertler, M.(1999) 'Inflation Dynamics: A Structural Econometric Analysis', *Journal of Monetary Economics* 44, 195-222.
- [18] Hall, R. (1978) 'Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence', *Journal of Political Economy* 86, 971-987.
- [19] Hamilton, J. (1994) *Time Series Analysis*, Princeton University Press, Princeton, NJ.
- [20] Honkapohja, S., Mitra, K. (2004) 'Are Non-fundamental Equilibria Learnable in Models of Monetary Policy?', *Journal of Monetary Economics* 51, 1743-1770.
- [21] Honkapohja, S., Mitra, K. (2005) 'Learning Stability in Economies with Heterogenous Agents', *Mimeo*.
- [22] Horn, R., Johnson, C., (1985) *Matrix Analysis*. Cambridge University Press, Cambridge, UK.
- [23] Lucas, R.E., Jr., (1972) 'Expectations and the Neutrality of Money', *Journal of Economic Theory* 4, 103-124.
- [24] Lucas, R.E. (1976) 'Econometric Policy Evaluation: A Critique', *Carnegie-Rochester Conference Series on Public Policy* 1, 19-46.

- [25] Mankiw, N.G. (2000) 'The Savers-Spenders Theory of Fiscal Policy', *American Economics Review* 90, 120-125.
- [26] Mankiw, N.G., Rotemberg, J., Summers, L. (1985) 'Intertemporal Substitution in Macroeconomics', *Quarterly Journal of Economics* 100, 225-251.
- [27] McCallum, B., Nelson, E. (1999) 'Nominal Income Targeting in an Open-Economy Optimising Model', *Journal of Monetary Economics* 43, 553-578.
- [28] McCallum, B., Nelson, E. (1999) 'Performance of Operational Policy Rules in an Estimated Semi-Classical Model', In: Taylor, J., Woodford, M. (Eds.), *Handbook of Macroeconomics*, Volume 1. Elsevier, Amsterdam.
- [29] Muth, J. F., (1961) 'Rational Expectations and the Theory of Price Movements', *Econometrica* 29, 315-335.
- [30] Parker, J. (1999) 'The Response of Household Consumption to Predictable Changes in Social Security Taxes', *American Economic Review* 89, 959-973.
- [31] Sargent, T.J., (1973) 'Rational Expectations, the Real Rate of Interest and the Natural Rate of Unemployment', *Brookings Papers on Economic Activity* 2, 429-472.
- [32] Sargent, T.J., (1993) *Bounded Rationality in Macroeconomics*. Oxford University Press, Oxford.
- [33] Shea, J. (1995) 'Union Contracts and the Life-Cycle/Permanent-Income Hypothesis', *American Economic Review* 85, 186-200.
- [34] Smets, F. (2003) 'Maintaining Price Stability: How Long is the Medium Term?', *Journal of Monetary Economics* 50, 1293-1309.
- [35] White, H. (1994) *Estimation, Inference, and Specification Analysis*. Cambridge University Press, NY.
- [36] Wolman, A. (1999) 'Sticky Prices, Marginal Cost, and the Behavior of Inflation', *Federal Reserve Bank of Richmond Economic Quarterly* 85, 29-48.
- [37] Woodford, M., (1999) 'Optimal Monetary Policy Inertia', *The Manchester School*, Supplement, 67, 1-35.
- [38] Zellner, A. (1962) 'An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias', *Journal of the American Statistical Association* 57, 348-368.

Figure 1. Woodford Calibration

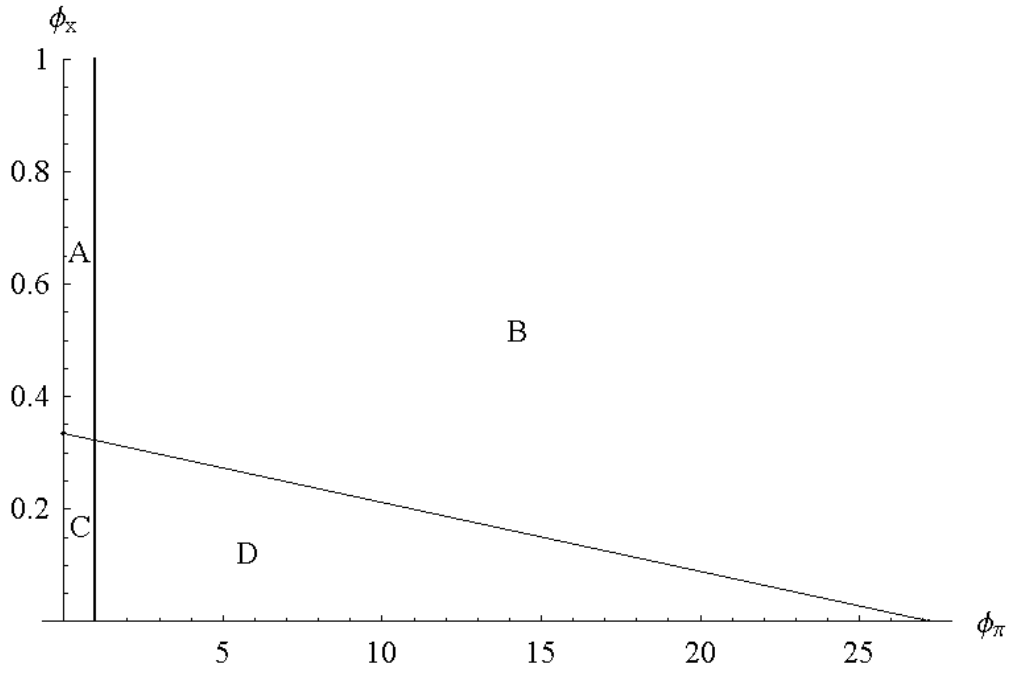


Figure 2: CGG Calibration

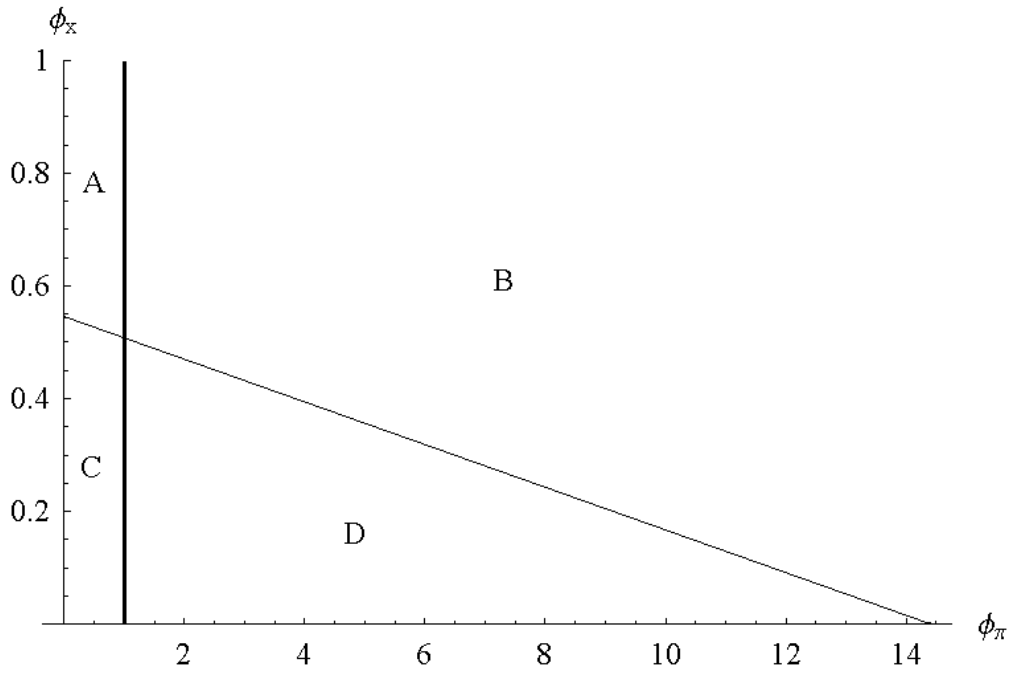


Figure 3: MN Calibration

