Stationarity of Econometric Learning with Bounded Memory and a Predicted State Variable*

Tatiana Damjanovic[†] University of Exeter

Šarūnas Girdėnas[‡] University of Exeter

Keqing Liu[§] University of Exeter

February 2015

Abstract

In this paper, we consider a model where producers set their prices based on their prediction of the aggregated price level and an exogenous variable, which can be a demand or a cost-push shock. To form their expectations, they use OLS-type econometric learning with bounded memory. We show that the aggregated price follows the random coefficient autoregressive process and we prove that this process is covariance stationary.

Keywords: econometric learning, bounded memory, random coefficient autoregressive process, stationarity.

JEL classification: C22, C53, C62, D83, E31

1 Introduction

Econometric Learning was designed to model the forecast of the future economic variables in forward looking models. In contrast to the Rational Expectations Theory, which imposes a

*We would like to thank Adriana Cornea, Martin Ellison, Michele Berardi, Jack R. Rogers and Dooruj Rambaccussing for comments and discussions. Keqing Liu and Šarūnas Girdėnas are grateful for financial support from ESRC. Special acknowledge goes to the anonymous referee who corrected some typos in our proof.

[†]Email: T.Damjanovic@exeter.ac.uk.

[‡]Email: sg325@exeter.ac.uk.

[§]Email: kl307@exeter.ac.uk.

very strong assumption that the agents know the structure of the model, Econometric Learning only assumes that agents behave as professional econometricians. They collect the available data and use OLS regression to produce the forecast. As more data becomes available, this econometric forecast often converges to the rational expectations equilibria (Sargent, 1993). Although econometric learning relaxes many assumptions of the rational expectations mechanism, we think that one of them could still be too strong. In particular, it assumes that agents have access to the entire history of the variables, and they use all of them to form the forecast. Not only does that assumption require infinite memory, it also neglects the cost of data collection and processing.

Several papers facilitate the assumption of infinite memory and consider the case when the memory is bounded (for a survey, see Chevillon and Mavroeidis, 2014). However, the majority of the results are proven for non-stochastic models (Evans and Honkapohja, 2000). The only exception known to us is Honkapohja and Mitra (2003) who investigate learning with bounded memory in a stochastic environment. However, they consider a very special case of learning the intercept parameter, and their model does not account for the possibility of using some exogenous independent variables when the expectation is formed.

This paper picks up the research from Honkapohja and Mitra (2003) and explores the dynamic properties of econometric learning with bounded memory in a stochastic environment. We expand that paper by adding a stochastic exogenous variable which can be used for econometric forecasts.

The introduction of stochastic independent variable makes the mathematical framework more complex as compared to Honkapohja and Mitra (2003) where the model evolves according to **a** simple autoregressive process (AR). In this paper**,** the model is more complex since the transition matrix has random coefficients (the random coefficient autoregressive model, $RCAR$, as in Nicholls and Quinn**,** 1982). It is also more complex than Conlisk (1974), since our transition matrices are autocorrelated. Nevertheless, we proved the stationarity of the model. In addition, we formulate a sufficient condition for stationarity which can be more generally applied in the RCAR literature.

This paper is structured as follows. In Section 2 we present the model and introduce OLS-type learning with finite memory. In Section 3 we prove that the $RCAR$ process of price movement is covariance-stationary. Section 4 concludes the paper.

2 The model

We consider a model where producer j sets the current price $p_t(j)$ depending on the expected aggregated level of price p_t^e and the exogenous but not completely observable state variable $\widetilde{w}_t$:

$$p_t(j) = \alpha + \beta p_t^e + \delta \widetilde{w}_t \tag{1}$$

where α, δ are known constant parameters and $\widetilde{w}_t$ is the estimated value of the exogenous cost push shock which can negatively affect the profit. The cost push shock w_t is not observed in period t; however, every producer has access to the historical data of its past realisation of $\{w_s\}$.

This model is very similar to the cobweb model as presented in Kaldor (1934), Ezekiel (1938) and more recently in Evans and Honkapohja (2003). It is known to be stable when $|\beta| < 1$. We will restrict our analysis to this particular case. In equilibrium, each producer sets the same price, that is $p_t = p_t(j)$.

2.1 OLS Learning

As w_t is the only state variable, the producer expects the aggregated price to depend on the variable

$$p_t = \alpha_2 + \beta_2 w_t, \tag{2}$$

where α_2 and β_2 are unknown parameters with producer estimates based on available historical data $\{p_s, w_s\}$. The price expectation is then

$$p_t^e = \widehat{\alpha}_{2,t-1} + \widehat{\beta}_{2,t-1} \widetilde{w}_t \tag{3}$$

where $\widehat{\alpha}_{2,t}$ and $\widehat{\beta}_{2,t}$ are estimated coefficients and $\widetilde{w}_t$ is a proxy for w_t. The classical OLS-type learning model assumes that agents forecast future prices by running the OLS regression using equation (2) and that at time t, the available information set consists of the entire history of prices and the exogenous state variable $\{p_s, w_s\}_{s=0}^{t-1}$. Coefficients $\widehat{\alpha}_{2,t}$ and $\widehat{\beta}_{2,t}$ are OLS estimators on the information set $\{p_s, w_s\}_{s=0}^{t}$.

2.2 Learning with Bounded Memory

Learning with bounded memory in our paper simply means that the agent is only using a limited number of observations T to form expectations.[1] The forecast will be made using the same OLS algorithm as in the classical case (3); however, we assume that only a finite set of historical data, $\{p_s, w_s\}_{s=t-T}^{t-1}$, is used to estimate the coefficients. Consequently, the estimators $\widehat{\alpha}_{2,t}$ and $\widehat{\beta}_{2,t}$ are defined as follows:

$$\widehat{\beta}_{2,t-1} = \frac{\sum_{i=1}^{T}\left[(w_{t-i}-\overline{w}_{t-1})(p_{t-i}-\overline{p}_{t-1})\right]}{\sum_{i=1}^{T}\left[(w_{t-i}-\overline{w}_{t-1})^2\right]}, \tag{4}$$

$$\widehat{\alpha}_{2,t-1} = \overline{p}_{t-1} - \widehat{\beta}_{2,t-1}\overline{w}_{t-1}, \tag{5}$$

$$\overline{w}_{t-1} = \frac{1}{T}\sum_{i=1}^{T} w_{t-i}, \tag{6}$$

$$\overline{p}_{t-1} = \frac{1}{T}\sum_{i=1}^{T} p_{t-i}. \tag{7}$$

Finally, as the agents cannot observe the realization of w_t at the time when they set their prices, the forecast $\widetilde{w}_t$ is used. The forecast is based on available historical data $\{w_s\}_{s=t-T}^{t-1}$, and consists of the weighted sum as in Mitra and Honkapohja (2003). Formally, $\widetilde{w}_t$ can be written as

$$\widetilde{w}_t = \sum_{i=1}^{t-1} \gamma_{i,t} w_{t-i}, \tag{8}$$

where $\gamma_{i,t}$ is the expected probability that $w_t = w_{t-i}$ and therefore,

$$\sum_{i=1}^{t-1} \gamma_{i,t} = 1. \tag{9}$$

Our set up covers an extensive range of models. For example, if w_t follows a Markov process with high persistency, the best prediction for w_t is w_{t-1}. In this case, $\gamma_{1t} = 1$, and $\gamma_{it} = 0$ for $i > 1$. In particular, for $T = 2$, $\gamma_1 = 1$, $\gamma_2 = 0$, the price p_t follows a simple autoregressive process with $p_t^e = p_{t-1}$. If w_t is $i.i.d.$ distributed, the best proxy for w_t might be $\overline{w}_{t-1}$. In this case, $\gamma_{i,t} = \frac{1}{T}$, and the price p_t follows the $AR(T)$ process with $p_t^e = \overline{p}_{t-1}$. Our model will also work if $\gamma_{i,t}$ corresponds to precautionary predictors with larger weights attached to the worse realisations as in the Robust Control or The Ambiguity Aversion theories.

[1] This is similar to Honkapohja and Mitra (2003) **where** a simplified version of the model without state variable is considered.

The complete model consists of (1), (3), (8), (4), (5), (6) and (7). Our aim is to show that p_t is stationary for all $T > 1$.

First, we show that the aggregated price p_t follows a Random Coefficient Autoregressive ($RCAR$) process.

Proposition 1 *The actual price follows an autoregressive process of order T with random coefficients as in (10)*

$$p_t = \alpha + \beta \left(\sum_{i=1}^{T} Z_{i,t} p_{t-i} \right) + \delta \widetilde{w}_t \tag{10}$$

where

$$Z_{i,t} = \frac{1}{T} + \frac{(w_{t-i} - \overline{w}_{t-1}) \left(\left(\sum_{i=1}^{T} \gamma_{i,t} (w_{t-i} - \overline{w}_{t-1}) \right) \right)}{\sum_{i=1}^{T} [(w_{t-i} - \overline{w}_{t-1})^2]}. \tag{11}$$

3 Stationarity of Bounded Memory Learning

Proposition 1 allows us to write our model in the $RCAR$ representation.

$$y_t = \varepsilon_t + M_t y_{t-1}, \tag{12}$$

where $M_t = \beta Z_t + S$ and S is **a** lower shift matrix,

$$y_t = \begin{pmatrix} p_t \\ p_{t-1} \\ .... \\ p_{t-T+1} \end{pmatrix}, \; Z_t = \begin{pmatrix} Z_{1,t} & Z_{2,t} & ... & Z_{T,t} \\ 0 & 0 & ... & 0 \\ .... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{pmatrix} \text{ and } \varepsilon_t = \begin{pmatrix} \alpha + \delta \widetilde{w}_t \\ 0 \\ .... \\ 0 \end{pmatrix}.$$

We begin our investigation of stationarity of the model (12) by setting up additional properties of coefficients $Z_{i,t}$.

Lemma 2 *For any realisation of w_t, i) $\sum_{i=1}^{T} Z_{i,t} = 1$ and ii) $\sum_{i=1}^{T} Z_{i,t}^2 \leq 1$.*

Proof. It is convenient to define $k_{i,t} = \frac{(w_{t-i} - \overline{w}_{t-1})}{\left(\sum_{i=1}^{T} (w_{t-i-1} - \overline{w}_{t-1})^2 \right)^{\frac{1}{2}}}$. Then**,** according to (11), $Z_{i,t} = \frac{1}{T} + k'_{i,t} \gamma'_t k_t$, where $k_{i,t}$ can be any number with the following restrictions

$$\sum_{i=1}^{T} k_{i,t} = 0, \tag{13}$$

$$\sum_{i=1}^{T} k_{i,t}^2 = 1. \tag{14}$$

Now we can compute:

$$\sum_{i=1}^{T} Z_{i,t}^2 = \sum_{i=1}^{T} \left(\frac{1}{T} + k_{i,t}\gamma_t' k_t \right)^2 = \frac{1}{T} + \left(\gamma_t' k_t\right)^2. \tag{15}$$

Maximisation of (15) subject to constraints (13) and (14) implies $(\gamma' k_t)^2 = \sum_{i=1}^{T} \gamma_{i,t}^2 - \frac{1}{T} \left(\sum_{i=1}^{T} \gamma_{i,t} \right)^2$ at the maximum. Evaluating (15) entails $\sum_{i=1}^{T} Z_{i,t}^2 = \frac{1}{T} + (\gamma' k_t)^2 \le \frac{1}{T} + \left(1 - \frac{1}{T}\right) = 1$. ■

Lemma 2 also implies that $|Z_{i,t}| < 1$. For further discussion, it is convenient to define a random matrix $G_{t,n} = \prod_{k=1,n} (M_{t-k})$. To show that it is finite, we will first establish the boundaries for every element of such a matrix.

Proposition 3 *Consider matrix M_{t-k} such that $M_{t-k} = \beta Z_{t-k} + S$, where S is a lower shift matrix, $|\beta| < 1$ and element $z_{i,j}$ of matrix Z_{t-k} satisfies*

$$\begin{aligned} |z_{1,j}| &\le 1, \\ z_{i,j} &= 0, \quad if \quad i > 1. \end{aligned}$$

Then, for any memory length T and $\widetilde{\beta} \in (\beta, 1)$, there exists a finite boundary c_T such that for any n, every element of the product of n matrixes, $G_{t,n}$, is bounded in absolute value by $c_T \widetilde{\beta}^n$ and therefore

$$|G_{t,n}| < c_T \widetilde{\beta}^n J,$$

where J is a $T \times T$ matrix of ones.

Proof. See Appendix 5. ■

Having established these results, we could investigate the stationarity of y_t by proving the existence of the unconditional expectations $E[y_t]$ and $E[y_t y_t']$.

Proposition 4 *Process (12) is covariance stationary if there exist unconditional expectations of $E[|\varepsilon_t|]$ and $E[|\varepsilon_t \varepsilon_t'|]$.*

Proof. To prove stationarity, we will iterate the backward expression (12):

$$y_t = \varepsilon_t + M_t y_{t-1} = \varepsilon_t + M_t \varepsilon_{t-1} + M_t M_{t-1} y_{t-1} = \sum_{k=0}^{\infty} G_{t,k} \varepsilon_{t-k}. \tag{16}$$

First, we will prove that the expectation of y_t is finite by applying Proposition 3:

$$E\left[y_t\right] < E\left[|y_t|\right] < E\left[\sum_{k=0}^{\infty}|Q^{t,k}||\varepsilon_{t-k}|\right] < Jc_T E\left[|\varepsilon_t|\sum_{k=0}^{\infty}\widetilde{\beta}^n\right] = \frac{c_T}{1-\widetilde{\beta}}J\left[E|\varepsilon_t|\right].$$

Thus, we have proved that $E\left[|y_t|\right]$ is finite if $E\left[|\varepsilon_t|\right]$ exists. To complete the proof, we need to show that $E\left[y_t y_t'\right]$ is also finite:

$$y_t y_t' = \varepsilon_t\varepsilon_t' + M_t y_{t-1}\varepsilon_t' + \varepsilon_t y_{t-1}' M_t' + M_t y_{t-1} y_{t-1}' M_t'. \tag{17}$$

We iterate it backwards to obtain:

$$y_t y_t' = \sum_{k=0}^{\infty} Q_{t,k}\left[\varepsilon_{t-k}\varepsilon_{t-k}' + M_{t-k}y_{t-k-1}\varepsilon_{t-k}' + \varepsilon_{t-k}y_{t-k-1}'M_{t-k}'\right]Q_{t,k}.$$

Finally, we will show that the expectations of the absolute value of the product are bounded[2]:

$$\begin{aligned}
E\left[|y_t y_t'|\right] &= E\sum_{k=0}^{\infty}\left[|Q_{t,k}|\left[|\varepsilon_{t-k}\varepsilon_{t-k}'| + |M_{t-k}||y_{t-k-1}||\varepsilon_{t-k}'| + |\varepsilon_{t-k}||y_{t-k-1}'||M_{t-k}'|\right]|Q_{t,k}|\right] \\
&< c_T^2 J\left(E\left[|\varepsilon_t\varepsilon_t'|\right] + JE\left[|y_t|\right]E\left[|\varepsilon_t'|\right] + E\left[|\varepsilon_t|\right]E\left[|y_t'|\right]J\right)J\sum_{k=0}^{\infty}\widetilde{\beta}^{2K} \\
&= c_T^2 J\left(E\left[|\varepsilon_t\varepsilon_t'|\right] + JE\left[|y_t|\right]E\left[|\varepsilon_t'|\right] + E\left[|\varepsilon_t|\right]E\left[|y_t'|\right]J\right)J\frac{1}{1-\widetilde{\beta}^2}.
\end{aligned}$$

■

Another interesting implication of Proposition 3 is that the spectral radius of M_t is smaller than one.

Lemma 5 *For any realization of the stochastic matrix M_t, its eigenvalues are less than one in absolute value.*

Proof. Consider $G_n = \left(M_t\right)^n$. Applying proposition 3 we can claim that $|G_n| < c_T\widetilde{\beta}^n J$:

$$\lim_{n\to\infty}\left(M_t\right)^n = 0$$

[2] We use that $|M_t| < J$.

which is necessary and sufficient for eigenvalues to be less than one in absolute value. ■

4 Conclusion

In this paper, we have investigated properties of econometric (OLS-type) learning with a bounded memory. We have shown that the eigenvalues of the transition matrix lie in the unit circle for any length of memory T. Furthermore, we have found that the price p_t follows a covariance stationary process. Our results could be tested in a DSGE framework, similarly to Berardi and Galimberti (2014).

References

[1] Berardi, Michele and Galimberti, Jaqueson K. ,2014, A note on the representative adaptive learning algorithm. *Economics Letters* 124, 104-107.

[2] Chevillon, G., Mavroeidis, S., 2014, 'Learning Can Generate Long Memory', mimeo.

[3] Conlisk, J., 1974, 'Stability in a Random Coefficient Model', *International Economic Review*, Vol. 15, pp. 529-533.

[4] Evans, G., W., Honkapohja, S., 2000, 'Convergence for Difference Equations with Vanishing Time-Dependence, with Applications to Adaptive Learning', *Economic Theory*, vol. 15, pp. 717-725.

[5] Ezekiel, M., 1938, 'The Cobweb Theorem', *Quarterly Journal of Economics,* vol. 52, pp. 255-280.

[6] Honkapohja, S., Mitra, K., 2003, 'Learning with Bounded Memory in Stochastic Models, *Journal of Economic Dynamics and Control,* vol. 27, pp. 1437-1457.

[7] Kaldor, N., 1934, 'A Classificatory Note on the Determinateness of Equilibrium', *The Review of Economic Studies,* vol. 1, pp. 122-136.

[8] Nicholls, D. F. and Quinn, B. G. (1982) Random Coefficient Autoregressive Models: An Introduction. New York: Springer-Verlag

[9] Sargent, Thomas J., 1993, Bounded Rationality in Macroeconomics, Clarendon Press, Oxford.

5 Appendix: Proof of Proposition 3

For any memory length T, and constant $\widetilde{\beta} > \beta$, there exists a boundary c_T such that every element of the product of n matrices M_t is bounded in absolute value by $c_T \widetilde{\beta}^n$:

$$|G_{t,n}|_{ij} = \left| \prod_{i=1,n} (M_{t-i}) \right|_{ij} < c_T \widetilde{\beta}^n, \tag{18}$$

where the matrix M_t can be represented as follows

$$M_t = \beta Z_t + S, \tag{19}$$

where Z_t has the form of

$$Z_t = \begin{pmatrix} Z_{1,t} & Z_{2,t} & ... & Z_{T-1,t} \\ 0 & 0 & ... & 0 \\ .... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{pmatrix}, \tag{20}$$

where each element $Z_{i,t}$ is smaller than 1 in absolute value, $|Z_{i,t}| < 1$; and S is the lower shift matrix

$$S = \begin{pmatrix} 0 & 0 & ... & 0 & 0 \\ 1 & 0 & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ .... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & 0 \end{pmatrix}. \tag{21}$$

Proof. First we will compute the product

$$G = \prod_{i=1,n} (M_{t-i}) = (\beta Z_{t-1} + S)(\beta Z_{t-2} + S) .... (\beta Z_{t-n} + S) \tag{22}$$

using the property of matrix S. For any matrix A, the first row of SA is zero. Moreover, if the first k rows of A are zeros, then the first of $k+1$ rows of SA are also zeros.

To compute the product (22) we need to sum up the products of n matrixes, each of them is either Z or S. However, if S appears more than $T-1$ times, the product is zero. Therefore, we can restrict our attention to only those cases when S appears less than T times.
The number of products with S being exactly on k places is $n!/k!/(n-k)!$ and therefore, the total number of non-zero products is less than $n!/((T-1)!(n-T+1)!) * (T-1)$.
Moreover, we can claim that every component is a matrix with elements less than $(\beta z)^{n-T}$, where $z = \max_i |Z_{i,t}| \leq 1$. Therefore, every element of

$$[(\beta Z_{t-1} + S)(\beta Z_{t-2} + S) .... (\beta Z_{t-n} + S)]_{ij} < \frac{n!}{(T-2)!(n-T+1)!} \beta^{n-T} < n^T \beta^{n-T}.$$

Consider the sequence $\{a_n\}$, defined as $a_n = n^T \beta^{n-T}$,

$$\frac{a_{n+1}}{a_n} = \left(\frac{n+1}{n}\right)^T \beta.$$

Let $\widetilde{\beta} \in (\beta, 1)$, then we can find $n^*(\widetilde{\beta}, T, \beta)$, such that for any $n > n^*$,

$$\frac{a_{n+1}}{a_n} = \left(\frac{n+1}{n}\right)^T \beta < \widetilde{\beta}, \tag{23}$$

in particular

$$n^* = ceil\left(\left[\left(\frac{\widetilde{\beta}}{\beta}\right)^{1/T} - 1\right]^{-1}\right).$$

It follows from (23) that for any positive k

$$a_{n^*+k} < \widetilde{\beta}^k a_{n^*} = \widetilde{\beta}^k n^{*T} \beta^{n^*-T}. \tag{24}$$

To complete the proof we define c_T

$$c_T = \max_{n \leq n^*} \left(a_n \widetilde{\beta}^{-n}\right).$$

■