ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B, 00120 Helsinki 12, Finland, tel. 601322

Keskusteluaiheita
Discussion papers

Juha Ahtola*

CONSEQUENCES FROM IMPROPER USE OF ORDINARY LEAST SQUARES ESTIMATION WITH TIME SERIES DATA

No 209 12.05.1986

*) Financial support from the Yrjö Jahnsson Foundation is gratefully acknowledged.

ISSN 0781-6847

Abstract

This note warns against the routine use of ordinary least squares (OLS) estimation technique in stochastic difference equation models. Unless the error process is given proper attention the estimation will result in models whose interpretations may be totally wrong. If the error process is nonstationary OLS will yield models with a strong autoregressive component regardless of the true value of the autoregressive parameter. Two important situations, where problems of the above nature arise, are in certain tests for rational expectations hypotheses and in the estimation of the adjustment speed parameter in partial adjustment models. Rational expectations hypotheses may be accepted too easily and the partial adjustment models will indicate extremely slow adjustment towards the target, even though the truth might be a very fast adjustment.

We give theoretical reasoning behind these estimation problems and also point to the difficulty in recognizing any misspecification through residual analysis.

1. INTRODUCTION

In econometric modelling of time series and other data it is more than customary to derive the models to be estimated from the premises of relevant economic theory and then just for the estimation add a random error term to account for the noise around the deterministic relationship. After the estimation the implications of the estimated model are argued on the basis of the deterministic part of the model. For some of the empirical work this is a legitimate strategy e.g. when testing the adequacy of economic theories. Much of the legality however, hinges on the appropriateness of assuming the errors of the model to be random noise. A sizeable amount of prior belief on the correctness of the deterministic specification will be introduced through this assumption. In time series analysis this would particularly mean that all the dynamics of the variables have been captured into the deterministic part, and that the error part has no substantial role in the generation of the variable of interest regardless of the sequential nature of the data generating process.

It is, of course, natural why a model-plus-noise structure is desirable. Interpretations are straightforward and the estimation of the unknown parameters is best done by simple methods, usually OLS. If, in addition, a careful investigator does not afterwards see anything alarming in the residual autocorrelations, then the specification of the model needs no excuses and good faith can be put on the implications of the estimated model.

This note focusses on the properties of OLS estimates of stochastic difference equations, when there is nonstationarity in the dependent variable due to an omitted variable or a proper error term. Two peculiar features result from the OLS estimations. Firstly, the estimate of the autoregressive parameter in the stochastic difference equation specification is shown to converge to unity regardless of the true value of the parameter. The second peculiar feature is the difficulty in finding anything suspicious in the residuals of the estimated model, which is the most common source for detecting specification problems.

Since the arguments we present are asymptotic it does not matter which of the two sources is the cause for nonstationarity as long as the possible omitted variable is uncorrelated with the included explanatory variable. Therefore the problems in the estimation are not due to the usual omitted variables problem but to the improper account for nonstationarity in the errors.

2. A SIMPLE REGRESSION MODEL

In this section we consider a simple regression model

$$y_t = \beta x_{t-1} + \varepsilon_t, \qquad t = 1,2,\ldots,n, \qquad (1)$$

where x_{t-1} is assumed stationary, with an absolutely summable autocovariance function, and for the ease of exposition with mean zero. A key assumption regulating model (1) here is that ε_t is nonstationary, such that either

(i) $\varepsilon_t = \varepsilon_{t-1} + a_t$

or

(ii) $\varepsilon_t = \gamma z_{t-1} + a_t$,

where a_t is Gaussian white noise, and in (ii) $z_{t-1} = z_{t-2} + \eta_{t-1}$, with η_{t-1} Gaussian white noise and independent of a_t and x_{t-1}.

A common source for nonstationarity in ε_t may be the case (ii) as Treadway (1983) has stressed. However, for our purposes (i) and (ii) are equivalent in a sense of yielding exactly the same problems for the estimators we consider. We will first consider one type of a potential danger in using data arising from a model like (1).

Consider a situation, where some rational expectations hypothesis would imply $E_{t-1}(\Delta y_t | I_{t-1}) = 0$, where $\Delta = I-L$ and I_{t-1} is some relevant information set of the past history up to and including period t-1, see e.g. Hall (1978).

A convenient route to set this implication under scrutiny is first to look at y_t alone and see whether its autoregressive representation has a unit root. Now OLS estimation of

$$y_t = \phi y_{t-1} + \text{error}_t$$

will give an estimate $\hat{\phi}$ which converges to unity when (1) is the true model. For a proof in case (i) see Ahtola (1986). The case (ii) can be handled similarly. Therefore it is easy to start to look at the behavior of Δy_t and regress it on the variables in the information set. Consider we regress Δy_t on x_{t-1}.

$$\Delta y_t = \beta x_{t-1} + u_t \qquad (2)$$

where $u_t = a_t - \beta x_{t-2}$, in case (i)

or $u_t = a_t - a_{t-1} + \gamma\eta_{t-1} - \beta x_{t-2}$, in case (ii).

OLS in (2) gives

$$\hat{\beta} = \beta + \frac{\Sigma x_{t-1} u_t}{\Sigma x_{t-1}^2} .$$

It is immediately seen that plim $\hat{\beta} = \beta(1 - \rho_x(1))$, where $\rho_x(1)$ is the autocorrelation coefficient of the x-series at lag 1.

As often is the case $\rho_x(1)$ may be quite large and consequently the bias towards zero in $\hat{\beta}$ can be considerable. Since the purpose of the regression (2) is to test H_0: $\beta = 0$ vs. H_A: $\beta \neq 0$, the bias in $\hat{\beta}$ is likely to favor H_0 in practice. Thus the rational expectations hypothesis would not easily be rejected, even though the true model were $y_t = \beta x_{t-1} + \gamma z_{t-1} + a_t$.

There is another problem in the above analysis, too. This problem is rather serious, since it makes the analysis look very innocent even for a careful investigator. Namely, the residuals, $\hat{u}_t$, from (2) may show no sign of autocorrelation. For instance in case (i)

$$\hat{u}_t = \Delta y_t - \hat{\beta} x_{t-1} \doteq \Delta y_t - \beta x_{t-1} + \beta \rho_x(1) x_{t-1}$$

$$= u_t + \beta \rho_x(1) x_{t-1} = a_t - \beta(x_{t-2} - \rho_x(1) x_{t-1})$$

Now, $x_{t-2} - \rho_x(1) x_{t-1}$ is white noise if x_t obeys an AR(1) model. Therefore the residual series can often in practice look like white noise.

3. A SIMPLE STOCHASTIC DIFFERENCE EQUATION

We now turn to a simple stochastic difference equation

$$y_t = \phi_1 y_{t-1} + \beta x_{t-1} + \varepsilon_t \quad . \tag{3}$$

The assumptions for (3) are the same as those with (1), and typically $0 \leq \phi_1 < 1$. Specification (3.1) is the familiar form to estimate e.g. the partial adjustment model

$$y_t^* = \beta^* x_t \tag{3a}$$

$$\Delta y_t = (1 - \phi_1)(y_{t-1}^* - y_{t-1}) \tag{3b}$$

by adding an error term to the derived difference equation form. We believe, it is usually the case that both (3a) and (3b), that is, the target level and the adjustment process, should be thought of as behavioral relationships with their own error terms. Restricting these error terms to be white noise precludes the possibility for the target to adjust, possibly slowly, due to changing environment, (cf. Feldstein and Auerbach (1976)) and the possibility for a fast actual adjustment (small ϕ_1) towards the target. In fact, we will see that OLS estimation in (3) gives plim $\hat{\phi}_1 = 1$, that is, very slow adjustment speeds will be estimated from (3).

Furthermore, the OLS estimate of β can be considerably biased towards zero, and finally the estimated residuals may show no apparent autocorrelation.

<u>Theorem 1</u> Let the true model be (3), with $|\phi_1| < 1$ and the assumptions for x_{t-1} and ε_t be the same as those with model (1). If we apply OLS to estimate ϕ_1 and β in (3) then

$$\text{plim } \hat{\phi}_1 = 1 \quad \text{and plim } \hat{\beta} = \beta\left(\sum_{j=1}^{\infty} \phi_1^{j-1}(\rho_x(j-1) - \rho_x(j))\right).$$

Proof: See the Appendix.

Note that if x_{t-1} is AR(1) with $\rho_x(j) = \rho^j$ then plim $\hat{\beta} = \beta \frac{1-\rho}{1-\rho\phi_1}$. Therefore, $\hat{\beta}$ is strongly biased towards zero when ρ is large and ϕ_1 small.

The most important implication of the above theorem is that the short run dynamics of the model will be badly misinterpreted. The short run effect of x_{t-1} on y_t will be underestimated and the adjustment speed of y_t towards its target will be (possibly grossly) underestimated. Thus the estimated model could give an indication of a very sluggish dynamics from x to y although the true model could be $\Delta y_t = \beta \Delta x_{t-1} + a_t$, where changes in x transmit quickly to changes in y.

Again, the residuals $\hat{u}_t = y_t - \hat{\phi}_1 y_{t-1} - \hat{\beta} x_{t-1}$ may indicate that the OLS estimated model is in no doubt, since autocorrelation in $\hat{u}_t$ may be hard to find. We will use the AR(1) model for the x-series to demonstrate this.

Now,

$$\begin{aligned} \hat{u}_t &= y_t - \hat{\phi}_1 y_{t-1} - \hat{\beta} x_{t-1} \\ &\doteq \Delta y_t - \frac{1-\rho}{1-\rho\phi_1} \beta x_{t-1} . \end{aligned}$$

Thus, in case (i)

$$(1 - \phi_1 L)\hat{u}_t \doteq (1 - \phi_1 L)\Delta y_t - \frac{1-\rho}{1-\rho\phi_1} \beta(1 - \phi_1 L)x_{t-1}$$

$$= \beta(1 - L)x_{t-1} - \frac{1 - \rho}{1 - \rho\phi_1}\beta(1 - \phi_1 L)x_{t-1} + a_t$$

$$= \beta\frac{1 - \rho}{1 - \rho\phi_1}(\rho x_{t-1} - x_{t-2}) + a_t ,$$

which is white noise.

Therefore $\hat{u}_t$ behaves like an AR(1) model with autoregressive parameter ϕ_1. If the true ϕ_1 is small, then autocorrelation in $\hat{u}_t$ is difficult to detect and we would probably be happy with the estimated model.

As a special case $\phi_1 = 0$, and therefore we see that all the problems encountered earlier in Section 2 are still present if we use the OLS regression of (3) as a basis for a test of the rational expectations hypothesis.

4. CONCLUSIONS

To conclude this note, we can summarize some practical implications from the neglect to consider y_t as arising from a model with autocorrelated errors:

- strong autoregressive relationships for y_t are often estimated.
- failure to recognize the influence of regressor variables and especially too easy acceptance of rational expectations hypotheses.
- the short run dynamics of the model will be misinterpreted and, in particular, adjustment speeds towards the target will be easily underestimated
- failure to see any misspecification from the residuals.

References

Ahtola, J. (1986) "Asymptotic inference in regression models with auto-regressive errors having roots on the unit circle" Economics Letters, to appear.

Feldstein, M.S. and Auerbach, A. (1976) "Inventory behavior in durable-goods manufacturing: The target adjustment model. Brookings Papers on Economic Activity 1976:2.

Hall, R.E. (1978) "Stochastic Implications of the life cycle-permanent income hypothesis: Theory and Evidence", Journal of Political Economy, 86, 971-987.

Treadway, A. (1983) "On the properties typical of economic time series", invited paper, First Catalan Symposium on Statistics, Barcelona.

APPENDIX

Proof of Theorem

The OLS estimates of ϕ_1 and β are

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \sum y_{t-1}^2 & \sum y_{t-1}x_{t-1} \\ \sum y_{t-1}x_{t-1} & \sum x_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_{t-1}y_t \\ \sum x_{t-1}y_t \end{bmatrix}$$

First, let us look at $\hat{\phi}_1$:

$$\hat{\phi}_1 = \frac{(\sum x_{t-1}^2)(\sum y_{t-1}y_t) - (\sum y_{t-1}x_{t-1})(\sum x_{t-1}y_t)}{(\sum y_{t-1}^2)(\sum x_{t-1}^2) - (\sum y_{t-1}x_{t-1})^2}$$

It is straightforward to show that

$\sum x_{t-1}\varepsilon_t = O_p(n)$, which implies that

$\sum y_{t-1}x_{t-1} = O_p(n)$ and $\sum x_{t-1}y_t = O_p(n)$.

Therefore dividing the numerator and denominator by $n\sum y_{t-1}^2$, which is $O_p(n^3)$, we get

$$\hat{\phi}_1 = \frac{\frac{1}{n}\sum x_{t-1}^2 \frac{\sum y_{t-1}y_t}{\sum y_{t-1}^2} - O_p(n^{-1})}{\frac{1}{n}\sum x_{t-1}^2 - O_p(n^{-1})}$$

Since $(1 - \phi_1 L)(1 - L)y_t$ is stationary, we immediately see that $\text{plim} \frac{\sum y_{t-1}y_t}{\sum y_{t-1}^2} = 1$. Also $\text{plim} \frac{1}{n} \sum x_{t-1}^2 = \sigma_x^2 > 0$. Therefore $\text{plim}\hat{\phi}_1 = 1$.

Next, look at $\hat{\beta}$:

$$\hat{\beta} = \frac{(\sum y_{t-1}^2)(\sum x_{t-1}y_t) - (\sum y_{t-1}x_{t-1})(\sum y_{t-1}y_t)}{(\sum y_{t-1}^2)(\sum x_{t-1}^2) - (\sum y_{t-1}x_{t-1})^2}$$

Again dividing the numerator and denominator by $n\sum y_{t-1}^2$, we get

$$\hat{\beta} = \frac{\frac{1}{n}\sum x_{t-1}y_t - \frac{1}{n}\sum y_{t-1}x_{t-1}(1 + o_p(1))}{\frac{1}{n}\sum x_{t-1}^2 - o_p(n^{-1})}$$

Inserting $y_t = \phi_1 y_{t-1} + \beta x_{t-1} + \varepsilon_t$, the numerator becomes

$$\frac{1}{n} \sum x_{t-1}\{(\phi_1 - 1)y_{t-1} + \beta x_{t-1} + \varepsilon_t\} + o_p(1)$$

$$= \frac{1}{n} \sum x_{t-1}\Delta y_t + o_p(1) .$$

Since $\Delta y_t = \beta x_{t-1} + u_t$, where $u_t = \phi_1 \Delta y_{t-1} - \beta x_{t-2} + a_t$, then after repeatedly inserting

$$\Delta y_{t-j-1} = \beta x_{t-j-2} + u_{t-j-1}, \quad j = 0,1,\ldots$$

into u_t, we quickly see that

$$\text{plim } \hat{\beta} = \beta\left(\sum_{j=1}^{\infty} \phi_1^{j-1}(\rho_x(j-1) - \rho_x(j))\right). \qquad \text{Q.E.D.}$$

ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)
The Research Institute of the Finnish Economy
Lönnrotinkatu 4 B, SF-00120 HELSINKI Puh./Tel. (90) 601 322

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 193 KARI ALHO, An Analysis of Financial Markets and Central Bank Policy: A Flow-of-Funds Approach to the Case of Finland. 15.01.1986. 44 p.

No 194 PAAVO OKKO, Julkisen rahoitustuen tehokkuus ja sen kohdentuminen eteläsuomalaisiin teollisuusyrityksiin. 15.01.1986. 46 s.

No 195 JUSSI KARKO, The Measurement of Productivity and Technical Change at Industry Level: An Application of Micro-Data to Industry Analysis. 16.01.1986. 40 p.

No 196 MARKKU RAHIALA, Teollisuusyritysten tuotantosuunnitelmien toteutumiseen vaikuttavat tekijät suhdannebarometriaineiston valossa tarkasteltuina. 20.01.1986. 53 s.

No 197 ERKKI KOSKELA, Taxation and Timber Supply under Uncertainty and Liquidity Constraints. 31.01.1986. 24 p.

No 198 PEKKA YLÄ-ANTTILA, Metsäteollisuuden kannattavuusvaihteluiden kokonaistaloudellisista vaikutuksista. 13.02.1986. 14 s.

No 199 JUHA KETTUNEN, Kansaneläke- ja sairausvakuutuksen rahoituksesta. 10.03.1986. 28 s.

No 200 JUKKA LESKELÄ, Välitysvaluutat ja ulkomaankaupan laskutus. 10.03.1986. 22 s.

No 201 VESA KANNIAINEN - HANNU HERNESNIEMI, Asset Structure, Indebtedness, and the Rate of Return on Capital in a Sample of Finnish Manufacturing Firms in 1961 - 1983. 11.03.1986. 31 s.

No 202 ANTTI RIPATTI, Teollisuus- ja ulkomaankauppatilaston yhdisteen hyödyntäminen. 20.03.1986. 31 s.

No 203 SYNNÖVE VUORI, Returns to R & D in Finnish and Swedish Manufacturing Industries. 20.03.1986. 23 p.

No 204 VESA KANNIAINEN, On the Effects of Inflation: The Debtor-Creditor Hypothesis Reconsidered. 20.03.1986. 15 p.

No 205 PEKKA ILMAKUNNAS, Aggregation of Micro Forecasts. 01.04.1986. 17 p.