Lecture 6: Dynamic Models

R.G. Pierse

1 Introduction

Up until now we have maintained the assumption that

X values are *fixed in repeated sampling* (A4)

In this lecture we look at dynamic models, where the regressors may include lags of the dependent variable as well as lags of other regressors. In this case assumption A4 can no longer be maintained.

1.1 The reasons for lags

There are several reasons why there may be some *inertia* in economic relationships, leading to the need to consider lagged variables in econometric equations.

1.1.1 Psychological reasons

As a result of the force of habit, agents may be slow to adjust to changes in economic variables. This can be perfectly consistent with the agents' behaviour being 'rational'. Friedman's theory of permanent income, Friedman (1957), distinguishes between transitory changes and permanent changes to income, only the latter of which should influence consumption. *Ex ante*, however, the agent does not know whether a change is transitory or permanent. This suggests that the rational agent should not immediately respond to a change in income but should wait until it is clear whether the change is persistent or merely a random blip.

More generally, in a world of incomplete information, agents have to learn about the economy. This learning takes time so that responses will not be immediate.

1.1.2 Technological reasons

There may be high costs attached to making changes. In this case frequent small changes will be inefficient and agents will wait until a series of changes has cumulated enough to justify the cost of making a response. Such arguments may explain the inertia of a firm in adjusting its prices. Some adjustments have a long gestation period before they can be effected. Classic examples are changes to the capital stock (both investment and disinvestment).

1.1.3 Institutional reasons

Sometimes institutional structures and laws contribute to inertia. Firms enter into contractual arrangements which means that they cannot always respond instantly to changes in the prices of factors of production. Investors can get locked into holding an asset even when changes in interest rates mean that it no longer offers the highest yield.

2 The Pure *AR* model

Consider the simplest possible case where there is a single regressor which is the lag of the dependent variable:

$$Y_t = \beta Y_{t-1} + u_t \quad , \quad t = 1, \cdots, T \tag{2.1}$$

where

$$-1 < \beta < 1$$

and u_t retains all the usual assumptions of the classical model. By successive substitution in (2.1) we can write

$$Y_t = u_t + \beta u_{t-1} + \beta^2 u_{t-2} + \cdots + \beta^{t-1} u_1 + \beta^t Y_0$$

or, lagging (2.1) and then substituting out,

$$Y_{t-1} = u_{t-1} + \beta u_{t-2} + \beta^2 u_{t-3} + \cdots + \beta^{t-2} u_1 + \beta^{t-1} Y_0$$

where the initial value Y_0 is taken as *fixed* with $Y_0 = 0$. Note that the regressors Y_{t-1} are now a linear function of the error term and so can *no longer be treated as fixed.*

For *OLS* to have its optimal properties we require that, for each of the k regressors, X_j, $j = 1, \cdots, k$,

$$\mathrm{E}(X_{jt} u_t) = 0\,. \tag{2.2}$$

In the *fixed regressor* case, this condition holds because $\mathrm{E}(X_{jt} u_t) = X_{jt}\,\mathrm{E}(u_t)$. When the X_j's include lagged dependent variables, however, then the regressors are not independent of the error term. In particular

$$Y_{t-1} u_{t-1} = u_{t-1}^2 + \beta u_{t-2} u_{t-1} + \beta^2 u_{t-3} u_{t-1} + \cdots + \beta^{t-2} u_1 u_{t-1} + \beta^{t-1} Y_0 u_{t-1}$$

and so

$$\mathrm{E}(Y_{t-1}u_{t-1}) = \mathrm{E}(u_{t-1}^2) = \sigma^2 \,.$$

More generally

$$\mathrm{E}(Y_{t-1}u_{t-i}) = \beta^{i-1}\,\mathrm{E}(u_{t-i}^2) = \beta^{i-1}\sigma^2 \quad , \quad i > 0$$

so that the regressors are correlated with *lagged values* of the error term. However, for the current value

$$Y_{t-1}u_t = u_{t-1}u_t + \beta u_{t-2}u_t + \beta^2 u_{t-3}u_t + \cdots + \beta^{t-2}u_1u_t + \beta^{t-1}Y_0u_t$$

and

$$\mathrm{E}(Y_{t-1}u_t) = 0 \,.$$

Thus, despite the correlation of the regressors with lags of the error term, the condition (2.2) still holds in this model and OLS retains its optimal properties.

2.1 Inference in dynamic models

Once we drop the assumption of fixed regressors, then inference is no longer exact. The distribution of the estimator $\widehat{\beta}$ will no longer be *exactly* normal because it is no longer a linear function of the error term u_t. Similarly, t-ratios and F-statistics will no longer *exactly* follow the t- or F- distributions. However, these distributions will continue to be *approximately* valid and we will continue to use them. Asymptotically, when the number of observations $T \to \infty$, then the (unknown) exact distributions of the estimators and test statistics collapse to the standard normal, t- or F- distributions.

However, this is not true for the Durbin-Watson statistic for testing for first order serial correlation. This test is *invalid* if the regressors include lags of the dependent variable and will be biased towards 2, i.e. non-rejection of the null. Consequently, it *should not be used* in dynamic models. The Breusch-Godfrey LM statistic however continues to be valid in dynamic models.

2.1.1 The Durbin h statistic

An alternative statistic for testing for first order autocorrelation in dynamic models was proposed by Durbin (1970). It is based on a transformation of the Durbin-watson statistic d:

$$h = (1 - \frac{1}{2}d)\sqrt{\frac{T}{1 - T\widehat{Var}(\widehat{\beta})}} \sim_a N(0,1)$$

This is a large sample test that is *asymptotically* distributed as a normal variate with zero mean and unit variance. Note that if $T\widehat{\text{var}}(\widehat{\beta}) > 1$ then the test cannot be computed since it involves the square root of a negative number. In general the Breusch-Godfrey test is to be preferred.

3 Autocorrelation in dynamic models

Consider the simple dynamic model again where now the error process is assumed to follow a first order autoregressive process.

$$Y_t = \beta Y_{t-1} + u_t \quad , \quad t = 1, \cdots, T \tag{3.1}$$

with

$$u_t = \rho u_{t-1} + \varepsilon_t \quad , \quad t = 1, \cdots, T \tag{3.2}$$

where

$$-1 < \beta < 1 \quad \text{and} \quad -1 < \rho < 1$$

and ε_t has all the standard properties:

$$\mathrm{E}(\varepsilon_t) = 0 \quad , \quad \mathrm{E}(\varepsilon_t^2) = \sigma_\varepsilon^2 \quad , \quad t = 1, \cdots, T$$

and

$$\mathrm{E}(\varepsilon_t \varepsilon_s) = 0 \quad , \quad t, s = 1, \cdots, T \quad s \neq t\ .$$

Recall that the error u_t has autocorrelations given by

$$\mathrm{E}(u_t u_{t-s}) = \frac{\sigma_\varepsilon^2 \rho^s}{1 - \rho^2} \quad , \quad s \geq 0\ .$$

In the presence of autocorrelation we can now show that *OLS is no longer unbiased*. From the definition of the *OLS* estimator we have

$$\mathrm{E}(\widehat{\beta}) = \mathrm{E}\left(\frac{\sum_{t=1}^T Y_{t-1} Y_t}{\sum_{t=1}^T Y_{t-1}^2}\right) = \beta + \mathrm{E}\left(\frac{\sum_{t=1}^T Y_{t-1} u_t}{\sum_{t=1}^T Y_{t-1}^2}\right)$$

but

$$\begin{aligned} \mathrm{E}(Y_{t-1} u_t) &= \mathrm{E}(u_{t-1} u_t) + \beta\ E(u_{t-2} u_t) + \cdots + \beta^{t-2}\,\mathrm{E}(u_1 u_t) + \beta^{t-1} Y_0\,\mathrm{E}(u_t) \\ &= \frac{\sigma_\varepsilon^2}{1 - \rho^2}(\rho + \beta\rho^2 + \beta^2\rho^3 + \cdots + \beta^{t-2}\rho^{t-1}) \neq 0\,. \end{aligned}$$

Thus, in this model,

$$\mathrm{E}(\widehat{\beta}) \neq \beta\,.$$

Hence autocorrelation in dynamic models is much more serious than in fixed regressor models since it causes *OLS* estimators to be biased.

4 Distributed Lag Models

A distributed lag model is one where the effect of the regressors is distributed over time. Consider the general *infinite order* distributed lag model with a single explanatory variable X:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + u_t \,. \qquad (4.1)$$

This equation cannot be estimated as it stands because it involves an infinite number of parameters. In practice, we need to consider the *finite l-th order* distributed lag model

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_l X_{t-l} + u_t \qquad (4.2)$$

where $l < T$.

What is the effect of a *sustained* change in X on Y? Clearly there is an immediate *impact* or *short run* effect given by

$$\frac{\partial Y_t}{\partial X_t} = \beta_0 \,.$$

However, in the next time period we have the effect $\partial Y_{t+1}/\partial X_{t+1} = \beta_0$ from X_{t+1} but also a lagged effect $\partial Y_{t+1}/\partial X_t = \beta_1$ coming from X_t. In the time period after that, there is a further effect from X_{t-2}, and so on. The change in X_t continues to have an effect on Y, up to l periods into the future. The *long run* or *equilibrium* effect of the change to X_t is given by the sum of all the effects which is

$$\sum_{i=0}^{l} \beta_i \,.$$

Note that the long run effect may have a different sign from the impact effect, or may even be zero. One example is the famous *J-curve* effect of a devaluation which initially worsens the balance of payments but in the long run will be expected to improve it. Similarly, in a monetarist model, increasing the supply of money will have a short run positive effect on real GDP. However, after the economy has completely adjusted, the long run effect will be zero.

Suppose that there are also lagged dependent variables in the model as in

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_l X_{t-l} + u_t \,. \qquad (4.3)$$

Then the *impact* or *short run* effect of a change to X_t on Y is given by

$$\frac{\partial Y_t}{\partial X_t} = \beta_0$$

as before. However, in future periods there is now an additional impact of the change in X_t through the lagged Y_t terms and this will continue into the infinite future. The long run or *equilibrium* effect of the change is now given by

$$\frac{\sum_{i=0}^{l} \beta_i}{(1 - \sum_{i=1}^{p} \alpha_i)} \, .$$

4.1 The Koyck lag scheme

The model (4.2) can be estimated by *OLS* although it is quite likely that the lags of X will be highly correlated so that the equation will be subject to multicollinearity. It is useful to be able to impose some restrictions on the pattern of lag weights to reduce the number of parameters to be estimated.

Suppose that the lag coefficients in (4.1)follow a pattern of exponentially declining weights given by

$$\beta_k = \beta_0 \lambda^k \tag{4.4}$$

where $0 < \lambda < 1$. These weights are known as Koyck weights after Koyck (1954). Substituting (4.4) into (4.1) we have

$$Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \cdots + u_t \, .$$

Lagging and multiplying by λ gives

$$\lambda Y_{t-1} = \lambda \alpha + \beta_0 \lambda X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \beta_0 \lambda^3 X_{t-3} + \cdots + \lambda u_{t-1}$$

and, subtracting from the original equation and rearranging,

$$Y_t = \lambda Y_{t-1} + (1 - \lambda)\alpha + \beta_0 X_t + v_t$$

where $v_t = u_t - \lambda u_{t-1}$ is a first order *moving average* process. Note that this equation is a *dynamic equation with serial correlation. OLS* estimation in this model will lead to biased estimates. A consistent estimator will be considered in the next section.

4.2 The Adaptive Expectations model

Consider the model

$$Y_t = \beta_0 + \beta_1 X_{t+1}^e + u_t \tag{4.5}$$

where X_{t+1}^e is the expected or equilibrium value of X_{t+1} which *is not directly observable*. Suppose that expectations are formed by the equation

$$X_{t+1}^e - X_t^e = \gamma (X_t - X_t^e)$$

or

$$X_{t+1}^e = \gamma X_t + (1-\gamma)X_t^e$$

where $0 < \gamma \leq 1$. This hypothesis is known as the *Adaptive Expectations Hypothesis* or *AEH* and has been suggested by Friedman (1957) *inter alia*. It implies that expectations are adjusted in the light of past experience. $X_t - X_t^e$ represents the mistake made in forming expectations in the previous period. The expectation formed in the current period is corrected by a factor γ times this past mistake where the correction factor is less than unity.

Substituting the expectations equation into the model gives

$$Y_t = \beta_0 + \beta_1 \gamma X_t + \beta_1 (1-\gamma) X_t^e + u_t \,. \tag{4.6}$$

Lagging (4.5), multiplying by $1-\gamma$, subtracting from (4.6), and then rearranging, gives

$$Y_t = (1-\gamma)Y_{t-1} + \gamma\beta_0 + \gamma\beta_1 X_t + +v_t$$

where $v_t = u_t - (1-\gamma)u_{t-1}$ is a first order moving average process. This is thus a dynamic equation with first order serial correlation. *OLS* estimation will lead to biased estimates.

4.3 Estimating the Adaptive Expectations Model

The problem with *OLS* estimation of the *AEH* model is that one of the regressors, Y_{t-1} is correlated with the error term v_t. A solution for this problem is to use the estimation method of *Instrumental Variables* or *IV*.

Consider the model

$$Y_t = \alpha + \beta X_t + u_t$$

where the regressor X_t is *correlated* with the error term u_t. Suppose that it is possible to find a variable Z_t which is highly correlated with X_t but is *not* correlated with the error term u_t. Such a variable is called an *instrument*. The *IV* estimator for β is then defined by

$$\widetilde{\beta}_{IV} = \frac{\sum_{t=1}^T z_t y_t}{\sum_{t=1}^T z_t x_t} \,.$$

Note that this is *not* the same as replacing X by Z and then regressing Y on Z.

In the case of the *AEH* model, a suitable instrument might be X_{t-1}. More generally, the *IV* method is a very useful estimation method for coping with problems of correlation between regressors and error term. The problem with the method is that of finding suitable instruments.

4.4 The Rational Expectations Hypothesis

Suppose that X_t is increasing. If expectations are formed by the *AEH*, then they will consistently underestimate X_t. Conversely, when X_t is decreasing, adaptive expectations will consistently overestimate X_t. It is hard to believe that rational agents will not revise their expectations when they observe this happening. This argument led Muth (1961) to the formulation of an alternative hypothesis for expectations formation: the *Rational Expectations hypothesis* or *REH*. This hypothesis is that expectations will always be unbiased so that

$$X_t = X_t^e + \omega_t$$

where ω_t is an error with zero mean, independent of X_t^e. This hypothesis means that there can be no systematic component in the forecast error which agents could correct. Let I_{t-1} represent all the information available at time $t-1$ when the expectation is being formed. Then the *REH* says that

$$X_t^e = \mathrm{E}(X_t|I_{t-1})\,.$$

Informally, this means that agents must have full knowledge about the 'true' mechanism or model that is generating the data, which is a very strong assumption.

4.5 The Partial Adjustment model

Consider the model

$$Y_t^* = \beta_0 + \beta_1 X_t + u_t$$

where Y_t^* is the desired or equilibrium value of Y_t which *is not directly observable.* Suppose that adjustment to the equilibrium level takes the form:

$$Y_t - Y_{t-1} = \delta(Y_t^* - Y_{t-1})$$

where $0 < \delta \leq 1$. This hypothesis is known as the *Partial Adjustment Hypothesis* or *PAH.* δ represents the speed of adjustment, and if $\delta < 1$ then adjustment is not complete. There are several reasons why adjustment may be incomplete such as adjustment costs, institutional or psychological inertias etc.

Substituting the equilibrium equation into the adjustment equation, and rearranging, gives

$$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta u_t\,.$$

This is a dynamic equation but the error process is serially uncorrelated so that *OLS* is optimal.

4.6 Almon lags

The Koyck distributed lag model is based on a very restrictive assumption of the pattern of lag weights. Consider again the general distributed lag model but, this time, assume the distributed lag is of *finite* order with l terms:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_l X_{t-l} + u_t$$

or

$$Y_t = \alpha + \sum_{j=0}^{l} \beta_j X_{t-j} + u_t \,.$$

Almon (1965) considers the parameters β_j as functions of the lag j. Then, by Weierstrass's theorem, this function can be approximated to any desired degree of accuracy by an mth order polynomial

$$\beta(j) = \gamma_0 + \gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + \cdots + \gamma_m j^m$$

where $m < l$. The Almon assumption allows for flexible patterns of weights while involving estimating less parameters than in the unrestricted weights case. Typically, only a low order of m will be required to approximate the β_j well. Suppose $m = 2$. Then

$$\begin{aligned} Y_t &= \alpha + \sum_{j=0}^{l} (\gamma_0 + \gamma_1 j + \gamma_2 j^2) X_{t-j} + u_t \\ &= \alpha + \gamma_0 \sum_{j=0}^{l} X_{t-j} + \gamma_1 \sum_{j=0}^{l} j\, X_{t-j} + \gamma_2 \sum_{j=0}^{l} j^2\, X_{t-j} + u_t \\ &= \alpha + \gamma_0 Z_{0t} + \gamma_1 Z_{1t} + \gamma_2 Z_{2t} + u_t \,. \end{aligned}$$

Implementing the Almon technique requires that we know the values of both l and m *a priori*.

References

[1] Almon, S. (1965), 'The distributed lag between capital appropriations and expenditures', *Econometrica*, 33, 178–196.

[2] Durbin, J. (1970), 'Testing for serial correlation in Least-Squares regression when some of the regressors are lagged dependent variables', *Econometrica*, 38, 410–421.

[3] Friedman, M. (1957), *A Theory of the Consumption Function*, Princeton University Press, Princeton, NJ.

[4] Koyck, L.M. (1954), *Distributed Lags and Investment Analysis*, North-Holland, Amsterdam.

[5] Muth, J.F. (1961), 'Rational expectations and the theory of price movements', *Econometrica*, 29, 315–335.