Chapter 15

Distributed Lag Models

15.1 Introduction

- In this chapter we focus on the dynamic nature of the economy, and the corresponding dynamic characteristics of economic data.
- We recognize that a change in the level of an explanatory variable may have behavioral implications beyond the time period in which it occurred. The consequences of economic decisions that result in changes in economic variables can last a long time.
- When the income tax is increased, consumers have less disposable income, reducing their expenditures on goods and services, which reduces profits of suppliers, which

reduces the demand for productive inputs, which reduces the profits of the input suppliers, and so on.

- These effects do not occur instantaneously but are spread, or *distributed*, over future time periods. As shown in Figure 15.1, economic actions or decisions taken at one point in time, t, affect the economy at time t, but also at times $t + 1$, $t + 2$, and so on.
- Monetary and fiscal policy changes, for example, may take six to eight months to have a noticeable effect; then it may take twelve to eighteen months for the policy effects to work through the economy.
- Algebraically, we can represent this lag effect by saying that a change in a policy variable x_t has an effect upon economic outcomes y_t, y_{t+1}, y_{t+2}, … . If we turn this around slightly, then we can say that y_t is affected by the values of x_t, x_{t-1}, x_{t-2}, … , or

$$y_t = f(x_t, x_{t-1}, x_{t-2}, \ldots) \tag{15.1.1}$$

- To make policy changes policymakers must be concerned with the *timing* of the changes and the length of time it takes for the major effects to take place. To make policy, they must know *how much* of the policy change will take place at the time of the change, *how much* will take place one month after the change, *how much* will take place two months after the changes, and so on.
- Models like (15.1.1) are said to be *dynamic* since they describe the evolving economy and its reactions over time.
- One immediate question with models like (15.1.1) is how far back in time we must go, or the *length* of the distributed lag. *Infinite distributed lag* models portray the effects as lasting, essentially, forever. In *finite distributed lag* models we assume that the effect of a change in a (policy) variable x_t affects economic outcomes y_t only for a certain, fixed, period of time.

15.2 Finite Distributed Lag Models

15.2.1 An Economic Model

- Quarterly capital expenditures by manufacturing firms arise from appropriations decisions in prior periods. Once an investment project is decided on, funds for it are *appropriated*, or approved for expenditure. The actual expenditures arising from any appropriation decision are observed over subsequent quarters as plans are finalized, materials and labor are engaged in the project, and construction is carried out.
- If x_t is the amount of capital appropriations observed at a particular time, we can be sure that the effects of that decision, in the form of capital expenditures y_t, will be "distributed" over periods t, $t + 1$, $t + 2$, and so on until the projects are completed.
- Furthermore, since a certain amount of "start-up" time is required for any investment project, we would not be surprised to see the major effects of the appropriation decision delayed for several quarters.

- As the work on the investment projects draws to a close, we expect to observe the expenditures related to the appropriation x_t declining.
- Since capital appropriations at time t, x_t, affect capital expenditures in the current and future periods (y_t, y_{t+1}, y_{t+2}, …), until the appropriated projects are completed, we may say equivalently that current expenditures y_t are a function of current and past appropriations x_t, x_{t-1}, … .
- Furthermore, let us assert that after n quarters, where n is the lag length, the effect of any appropriation decision on capital expenditure is exhausted. We can represent this economic model as

$$y_t = f(x_t, x_{t-1}, x_{t-2}, \dots, x_{t-n}) \qquad (15.2.1)$$

- Current capital expenditures y_t depend on current capital appropriations, x_t, as well as the appropriations in the previous n periods, x_t, x_{t-1}, x_{t-2}, … , x_{t-n}. This distributed lag model is *finite* as the duration of the effects is a finite period of time, namely n periods. We now must convert this economic model into a statistical one so that we can give it empirical content.

15.2.2 The Econometric Model

- In order to convert model (15.2.1) into an econometric model we must choose a functional form, add an error term and make assumptions about the properties of the error term.
- As a first approximation let us assume that the functional form is linear, so that the finite lag model, with an additive error term, is

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_n x_{t-n} + e_t, \quad t = n + 1, \ldots, T \qquad (15.2.2)$$

where we assume that $E(e_t) = 0$, $\text{var}(e_t) = \sigma^2$, and $\text{cov}(e_t, e_s) = 0$.

- Note that if we have T observations on the pairs (y_t, x_t) then only $T - n$ *complete* observations are available for estimation since n observations are "lost" in creating x_{t-1}, x_{t-2}, … , x_{t-n}.
- In this finite distributed lag the parameter α is the intercept and the parameter β_i is called a **distributed lag weight** to reflect the fact that it measures the effect of changes in past appropriations, Δx_{t-i}, on expected current expenditures, $\Delta E(y_t)$, all other things held constant. That is,

$$\frac{\partial E(y_t)}{\partial x_{t-i}} = \beta_i \tag{15.2.3}$$

- Equation (15.2.2) can be estimated by least squares if the error term e_t has the usual desirable properties. However, collinearity is often a serious problem in such models.

Recall from Chapter 8 that collinearity is often a serious problem caused by explanatory variables that are correlated with one another.

- In Equation (15.2.2) the variables x_t and x_{t-1}, and other pairs of lagged x's as well, are likely to be closely related when using time-series data. If x_t follows a pattern over time, then x_{t-1} will follow a similar pattern, thus causing x_t and x_{t-1} to be correlated. There may be serious consequences from applying least squares to these data.
- Some of these consequences are imprecise least squares estimation, leading to wide interval estimates, coefficients that are statistically insignificant, estimated coefficients that may have incorrect signs, and results that are very sensitive to changes in model specification or the sample period. These consequences mean that the least squares estimates may be unreliable.
- Since the pattern of lag weights will often be used for policy analysis, this imprecision may have adverse social consequences. Imposing a tax cut at the *wrong* time in the business cycle can do much harm.

15.2.3 An Empirical Illustration

- To give an empirical illustration of this type of model, consider data on quarterly capital expenditures and appropriations for U. S. manufacturing firms. Some of the observations are shown in Table 15.1.
- We assume that $n = 8$ periods are required to exhaust the expenditure effects of a capital appropriation in manufacturing. The basis for this choice is investigated in Section 15.2.5, since the lag length n is actually an unknown constant. The least squares parameter estimates for the finite lag model (15.2.2) are given in Table 15.2.

Table 15.2 Least Squares Estimates for the Unrestricted Finite Distributed Lag Model

Variable	Estimate	Std. Error	*t*-value	*p*-value
const.	3.414	53.709	0.622	0.5359

x_t	0.038	0.035	1.107	0.2721
x_{t-1}	0.067	0.069	0.981	0.3300
x_{t-2}	0.181	0.089	2.028	0.0463
x_{t-3}	0.194	0.093	2.101	0.0392
x_{t-4}	0.170	0.093	1.824	0.0723
x_{t-5}	0.052	0.092	0.571	0.5701
x_{t-6}	0.052	0.094	0.559	0.5780
x_{t-7}	0.056	0.094	0.597	0.5526
x_{t-8}	0.127	0.060	2.124	0.0372

- The R^2 for the estimated relation is 0.99 and the overall *F*-test value is 1174.8. The statistical model "fits" the data well and the *F*-test of the joint hypotheses that all distributed lag weights $\beta_i = 0$, $i = 0, \ldots, 8$, is rejected at the $\alpha = .01$ level of significance.

- Examining the individual parameter estimates, we notice several disquieting facts. First, only the lag weights b_2, b_3, b_4, and b_8 are statistically significantly different from zero based on individual *t*-tests, reflecting the fact that the estimates' standard errors are large relative to the estimated coefficients.
- Second, the estimated lag weights b_7 and b_8 are *larger* than the estimated lag weights for lags of 5 and 6 periods. This does not agree with our anticipation that the lag effects of appropriations should decrease with time and in the most distant periods should be small and approaching zero.
- These characteristics are symptomatic of collinearity in the data. The simple correlations among the current and lagged values of capital appropriations are large. Consequently, a high level of *linear* dependence is indicated among the explanatory variables. Thus, we conclude that the least squares estimates in Table 15.2 are subject to great sampling variability and are unreliable, owing to the limited independent information provided by each explanatory variable x_{t-i}.

- In Chapter 8 we noted that one way to combat the ill-effects of collinearity is to use restricted least squares. By replacing restrictions on the model parameters we reduce the variances of the estimator.
- In the context of distributed lag models we often have an idea of the pattern of the time effects, which we can translate into parameter restrictions. In the following section we restrict the lag weights to fall on a polynomial.

15.2.4 Polynomial Distributed Lags

- Imposing a shape on the lag distribution will reduce the effects of collinearity. Let us assume that the lag weights follow a smooth pattern that can be represented by a low degree polynomial. Shirley Almon introduced this idea, and the resulting finite lag model is often called the **Almon distributed lag**, or a **polynomial distributed lag**.

- For example, suppose we select a second-order polynomial to represent the pattern of lag weights. Then the effect of a change in x_{t-i} on $E(y_t)$ is

$$\frac{\partial E(y_t)}{\partial x_{t-i}} = \beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2, \quad i = 0, \ \dots \ , \ n \tag{15.2.4}$$

- An example of this quadratic polynomial lag is depicted in Figure 15.2. The polynomial lag in Figure 15.2 depicts a situation that commonly arises when modeling the effects of monetary and fiscal policy. At time t the effect of a change in a policy variable is

$$\frac{\partial E(y_t)}{\partial x_t} = \beta_0 = \gamma_0$$

The immediate impact might well be less than the impact after several quarters, or months. After reaching its maximum, the policy effect diminishes for the remainder of the finite lag.

- For illustrative purposes again suppose that the lag length is $n = 4$ periods. Then the finite lag model is

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \beta_4 x_{t-4} + e_t, \quad t = 5, \ldots, T \qquad (15.2.5)$$

- Then the relations in Equation (15.2.4) become

$$\begin{aligned} \beta_0 &= \gamma_0 & i &= 0 \\ \beta_1 &= \gamma_0 + \gamma_1 + \gamma_2 & i &= 1 \\ \beta_2 &= \gamma_0 + 2\gamma_1 + 4\gamma_2 & i &= 2 \qquad (16.2.6) \\ \beta_3 &= \gamma_0 + 3\gamma_1 + 9\gamma_2 & i &= 3 \end{aligned}$$

$$\beta_4 = \gamma_0 + 4\gamma_1 + 16\gamma_2 \qquad i = 4$$

- In order to estimate the parameters describing the polynomial lag, γ_0, γ_1, and γ_2, we substitute Equation (15.2.6) into the finite lag model Equation (15.2.5) to obtain

$$\begin{aligned} y_t &= \alpha + \gamma_0 x_t + (\gamma_0 + \gamma_1 + \gamma_2)x_{t-1} + (\gamma_0 + 2\gamma_1 + 4\gamma_2)x_{t-2} \\ &\quad + (\gamma_0 + 3\gamma_1 + 9\gamma_2)x_{t-3} + (\gamma_0 + 4\gamma_1 + 16\gamma_2)x_{t-4} + e_t \\ &= \alpha + \gamma_0 z_{t0} + \gamma_1 z_{t1} + \gamma_2 z_{t2} + e_t \end{aligned} \tag{15.2.7}$$

In Equation (15.2.7) we have defined the constructed variables z_{tk} as

$$z_{t0} = x_t + x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4}$$
$$z_{t1} = x_{t-1} + 2x_{t-2} + 3x_{t-3} + 4x_{t-4}$$
$$z_{t2} = x_{t-1} + 4x_{t-2} + 9x_{t-3} + 16x_{t-4}$$

- Once these variables are created the polynomial coefficients are estimated by applying least squares to Equation (15.2.7).
- If we denote the estimated values of γ_k by $\hat{\gamma}_k$, then we can obtain the estimated lag weights as

$$\hat{\beta}_i = \hat{\gamma}_0 + \gamma_1 i + \gamma_2 i^2, \quad i = 0, \; \ldots \;, \; n \tag{15.2.8}$$

Whatever the degree of the polynomial, the general procedure is an extension of what we have described for the quadratic polynomial.

- Equation (15.2.7) is a restricted model. We have replaced (n + 1) = 5 distributed lag weights with 3 polynomial coefficients. This implies that in constraining the distributed lag weights to a polynomial of degree 2 we have imposed J = (n + 1) – 3 = 2 parameter restrictions.

- We may wish to check the compatibility of the quadratic polynomial lag model with the data by performing an *F*-test, comparing the sum of squared errors from the restricted model in Equation (15.2.7) to the sum of squared errors from the unrestricted model (15.2.5).
- As an illustration, we will fit a second-order polynomial lag to the capital expenditure data in Table 15.1, with a lag length of $n = 8$ periods. In Table 15.3 are the estimated polynomial coefficients from Equation (15.2.7).

Table 15.3 Estimated (Almon) Polynomial Coefficients

Parameter	Estimates	*t*-value	*p*-value
α	51.573	0.970	0.3351
γ_0	0.067	4.411	0.0001
γ_1	0.038	2.984	0.0038

γ_2	–0.005	–3.156	0.0023

- In Table 15.4 we present the distributed lag weights calculated using Equation (15.2.8). The reported standard errors are based on the fact that the estimated distributed lag weights are combinations of the estimates in Table 15.3.

Table 15.4 Estimated Almon Distributed Lag Weights from Polynomial of Degree Two

Parameter	Estimate	Std. Error	*t*-value	*p*-value
β_0	0.067	0.015	4.41	0.0001
β_1	0.100	0.005	19.60	0.0001
β_2	0.123	0.005	22.74	0.0001
β_3	0.136	0.009	14.40	0.0001
β_4	0.138	0.011	12.86	0.0001

β_5	0.130	0.009	14.31	0.0001
β_6	0.112	0.005	20.92	0.0001
β_7	0.083	0.007	11.32	0.0001
β_8	0.044	0.018	2.47	0.0156

- Since the estimated weights in Table 15.4 are linear combinations of the estimated polynomial coefficients in Table 15.3, as shown in Equation (15.2.8), their estimated variances are calculated using Equation (2.5.8), from Chapter 2.
- Constraining the distributed lag weights to fall on a polynomial of degree two has drastically affected their values as compared to the unconstrained values in Table 15.2.
- Also, note that the standard errors of the estimated coefficients are much smaller than those in the unconstrained model indicating more precise parameter estimation.

Remark: Recall that imposing restrictions on parameters leads to bias unless the restrictions are true. In this case we do not really believe that the distributed lag weights fall exactly on a polynomial of degree two. However, if this assumption approximates reality, then the constrained estimator will exhibit a small amount of bias. Our objective is to trade a large reduction in sampling variance for the introduction of some bias, increasing the probability of obtaining estimates *close* to the true values.

- In Figure 15.3 we plot the unrestricted estimates of lag weights and the restricted estimates. Note that the restricted estimates display the increasing-then-decreasing "humped" shape that economic reasoning led us to expect. The effect of a change in capital appropriations x_t at time t leads to an increase in capital expenditures in the current period, y_t, by a relatively small amount. However, the expenditures arising from the appropriation decision increase during the next four quarters, before the effect begins to taper off.

15.2.5 Selection of the Length of the Finite Lag

- Numerous procedures have been suggested for selecting the length *n* of a finite distributed lag. None of the proposed methods is entirely satisfactory. The issue is an important one, however, because fitting a polynomial lag model in which the lag length is either over- or understated may lead to biases in the estimation of the lag weights, even if an appropriate polynomial degree has been selected.
- We offer two suggestions that are based on "goodness-of-fit" criteria. Begin by selecting a lag length *N* that is the *maximum* that you are willing to consider. The unrestricted finite lag model is then

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \ldots + \beta_N x_{t-N} + e_t \qquad (15.2.9)$$

- We wish to assess the goodness of fit for lag lengths $n \le N$. The usual measures of goodness-of-fit, R^2 and $\bar{R}^2$, have been found not to be useful for this task.
- Two goodness-of-fit measures that are more appropriate are Akaike's *AIC* criterion

$$AIC = \ln \frac{SSE_n}{T-N} + \frac{2(n+2)}{T-N} \qquad (15.2.10)$$

and Schwarz's *SC* criterion

$$SC = \ln \frac{SSE_n}{T-N} + \frac{(N+2)\ln(T-N)}{T-N} \qquad (15.2.11)$$

- For each of these measures we seek that lag length n^* that minimizes the criterion used. Since adding more lagged variables reduces *SSE*, the second part of each of the criteria is a *penalty function* for adding additional lags.
- These measures weigh reductions in sum of squared errors obtained by adding additional lags against the penalty imposed by each. They are useful for comparing lag lengths of alternative models estimated using the same number of observations.

15.3 The Geometric Lag

An *infinite distributed lag model* in its most general form is:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \cdots + e_t$$

$$= \alpha + \sum_{i=0}^{\infty} \beta_i x_{t-i} + e_t \tag{15.3.1}$$

- In this model y_t is taken to be a function of x_t and *all* its previous values. There may also be other explanatory variables on the right-hand side of the equation. The model in Equation (15.3.1) is impossible to estimate since there are an infinite number of parameters.
- Models have been developed that are *parsimonious*, and which reduce the number of parameters to estimate. The cost of reducing the number of parameters is that these

models must assume particular patterns for the parameters β_i, which are called *distributed lag weights*.

- One popular model is the **geometric lag**, in which the lag weights are positive and decline geometrically. That is

$$\beta_i = \beta\phi^i, \quad |\phi| < 1 \qquad (15.3.2)$$

The parameter β is a scaling factor and the parameter ϕ is less than 1 in absolute value. The pattern of lag weights β_i is shown in Figure 15.4.

- The lag weights $\beta_i = \beta\phi^i$ decline toward zero as i gets larger. The most recent past is more heavily weighted than the more distant past, and, although the weights never reach zero, beyond a point they become negligible.
- Substituting Equation (15.3.2) into Equation (15.3.1) we obtain,

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \ldots + e_t$$

$$= \alpha + \beta(x_t + \phi x_{t-1} + \phi^2 x_{t-2} + \phi^3 x_{t-3} + \ldots) + e_t \qquad (15.3.3)$$

which is the **infinite geometric distributed lag model**. In this model there are 3 parameters, α-an intercept parameter, β-a scale factor, and ϕ-which controls the rate at which the weights decline.

- In Equation (15.3.3) the effect of a one unit change in x_{t-i} on $E(y_t)$ is

$$\frac{\partial E(y_t)}{\partial x_{t-i}} = \beta_i = \beta\phi^i \qquad (15.3.4)$$

This equation says that the change in the average value of y in period t given a change in x in period $t-i$, all other factors held constant, is $\beta_i = \beta\phi^i$.

- The change in $E(y_t)$ given a unit change in x_t is β; it is called an **impact multiplier** since it measures the change in the current period.
- If the change in period t is sustained for another period then the combined effect $\beta + \beta\phi$ is felt in period $t + 1$. If the change is sustained for three periods the effect on $E(y_{t+2})$ is $\beta + \beta\phi + \beta\phi^2$. These sums are called **interim multipliers** and are the effect of sustained changes in x_t.
- If the change is sustained permanently then the total effect, or the **long run multiplier**, is

$$\beta(1+\phi+\phi^2+\phi^3+\cdots)=\frac{\beta}{1-\phi}$$

15.4 The Koyck Transformation

How shall we estimate the geometric lag model represented by Equation (15.3.3)? As it stands, it is an impossible task, since the model has an infinite number of terms, and the parameters β and ϕ are multiplied together, making the model *nonlinear in the parameters*. One way around these difficulties is to use the **Koyck transformation**, in deference to L. M. Koyck, who developed it.

- To apply the Koyck transformation, lag equation (15.3.3) one period, multiply by ϕ, and subtract that result from lag equation (15.3.3). We obtain

$$\begin{aligned} y_t - \phi y_{t-1} &= [\alpha + \beta(x_t + \phi x_{t-1} + \phi^2 x_{t-2} + \phi^3 x_{t-3} + \cdots) + e_t] \\ &\quad - \phi[\alpha + \beta(x_{t-1} + \phi x_{t-2} + \phi^2 x_{t-3} + \phi^3 x_{t-4} + \cdots) + e_{t-1}] \\ &= \alpha(1-\phi) + \beta x_t + (e_t - \phi e_{t-1}) \end{aligned} \tag{15.4.1}$$

- Solving for y_t we obtain the Koyck form of the geometric lag,

$$\begin{aligned} y_t &= \alpha(1 - \phi) + \phi y_{t-1} + \beta x_t + (e_t - \phi e_{t-1}) \\ &= \beta_1 + \beta_2 y_{t-1} + \beta_3 x_t + v_t \end{aligned} \qquad (15.4.2)$$

where $\beta_1 = \alpha(1 - \phi)$, $\beta_2 = \phi$, $\beta_3 = \beta$ and the random error $v_t = (e_t - \phi e_{t-1})$.

15.4.1 Instrumental Variables Estimation of the Koyck Model

- The last line of equation (15.4.2) looks like a multiple regression model, with two special characteristics. The first is that one of the explanatory variables is the *lagged dependent variable*, y_{t-1}. The second is that the error term v_t depends on e_t and on e_{t-1}. Consequently, y_{t-1} and the error term v_t must be *correlated*, since equation (15.3.3) shows that y_{t-1} depends directly on e_{t-1}.

- In Chapter 13.2.4, we showed that such a correlation between an explanatory variable and the error term causes the least squares estimator of the parameters is biased and *inconsistent*. Consequently, in Equation (15.4.2) we should not use the least squares estimator to obtain estimates of β_1, β_2, and β_3.
- We can estimate the parameters in Equation (15.4.2) consistently using the instrumental variables estimator.
- The "problem" variable in Equation (15.4.2) is y_{t-1}, since it is the one correlated with the error term v_t.
- An appropriate instrument for y_{t-1} is x_{t-1}, which is correlated with y_{t-1} (from Equation (15.4.2)) and which is uncorrelated with the error term v_t (since it is exogenous).
- Instrumental variables estimation can be carried out using two-stage least squares. Replace y_{t-1} in Equation (15.4.2) by $\hat{y}_{t-1} = a_0 + a_1 x_{t-1}$, where the coefficients a_0 and a_1 are obtained by a simple least squares regression of y_{t-1} on x_{t-1}, to obtain

$$y_t = \delta_1 + \delta_2 \hat{y}_{t-1} + \delta_3 x_t + error \qquad (15.4.3)$$

- Least squares estimation of Equation (15.4.3) is equivalent to instrumental variables estimation, as we have shown in Chapter 13.3.5. The variable x_{t-1} is an instrument for y_{t-1}, and x_t is an instrument for itself.
- Using the instrumental variables estimates of δ_1, δ_2, and δ_3, we can derive estimates of α, β, and ϕ in the geometric lag model.

15.4.2 Testing for Autocorrelation in Models with Lagged Dependent Variables

- In the context of the lagged dependent variable model (15.4.2), obtained by the Koyck transformation, we know that the error term is correlated with the lagged dependent variable on the right-hand side of the model, and thus that the usual least squares estimator fails.

- Suppose, however, that we have obtained a model like (15.4.2) through other reasoning, and that we do not know whether the error term is correlated with the lagged dependent variable or not. If it is not, then we can use least squares estimation. If it is, we should not use least squares estimation.
- The key question is whether the error term, v_t in Equation (15.4.2), is serially correlated or not, since if it is, then it is also correlated with y_{t-1}. The Durbin-Watson test is not applicable in this case, because in a model with an explanatory variable that is a lagged dependent variable it is biased towards finding no autocorrelation.
- A test that is valid in large samples is the LM test for autocorrelation introduced in Chapter 12.6.2. Estimate Equation (15.4.2) by least squares and compute the least squares residuals, $\hat{e}_t$. Then estimate the artificial regression

$$y_t = a_1 + a_2 y_{t-1} + a_3 x_t + a_4 \hat{e}_{t-1} + error \qquad (15.4.4)$$

- Test the significance of the coefficient on $\hat{e}_{t-1}$ using the usual *t*-test. If the coefficient is significant, then reject the null hypothesis of no autocorrelation. This alternative test is also useful in other general circumstances and can be extended to include least squares residuals with more than one lag.

15.5 Autoregressive Distributed Lags

- There are some obvious problems with the two distributed lags models we have discussed. The finite lag model requires us to choose the lag length and then deal with collinearity in the resulting model. The polynomial distributed lag (PDL) addresses the collinearity by requiring the lag weights to fall on a smooth curve. While the PDL is flexible, it is still a very strong assumption to make about the structure of lag weights.
- The infinite lag model removes the problem of specifying the lag length, but requires us to impose structure on the lag weights to get around the problem that there are an infinite number of parameters.
- The geometric lag is one such structure, but it imposes the condition that successive lag weights decline geometrically. This model would not do in a situation in which

the peak effect does not occur for several periods, such as when modeling monetary or fiscal policy.

- In this section we present an alternative model that may be useful when neither a polynomial distributed lag nor a geometric lag is suitable.

15.5.1 The Autoregressive Distributed Lag Model

- The autoregressive-distributed lag (ARDL) is an infinite lag model that is both flexible and parsimonious.
- An example of an ARDL is as follows:

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \gamma_1 y_{t-1} + e_t \qquad (15.5.1)$$

In this model we include the explanatory variable x_t, and one or more of its lags, with one or more lagged values of the dependent variable.

- The model in (15.5.1) is denoted as ARDL(1, 1) as it contains one lagged value of x and one lagged value of y. A model containing p lags of x and q lags of y is denoted ARDL(p, q).
- If the usual error assumptions on the error term e hold, then the parameters of Equation (15.5.1) can be estimated by least squares.
- Despite its simple appearance the ARDL(1,1) model represents an infinite lag. To see this we repeatedly substitute for the lagged dependent variable on the right-hand side of Equation (15.5.1). The lagged value y_{t-1} is given by

$$y_{t-1} = \mu + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \gamma_1 y_{t-2} + e_{t-1} \qquad (15.5.2)$$

Substitute Equation (15.5.2) into Equation (15.5.1) and rearrange,

$$
\begin{aligned}
y_t &= \mu + \beta_0 x_t + \beta_1 x_{t-1} + \gamma_1[\mu + \beta_0 x_{t-1} + \beta_1 x_{t-2} + \gamma_1 y_{t-2} + e_{t-1}] + e_t \\
&= \mu(1+\gamma_1) + \beta_0 x_t + (\beta_1 + \gamma_1\beta_0)x_{t-1} + \gamma_1\beta_1 x_{t-2} + \gamma_1^2 y_{t-2} + (\gamma_1 e_{t-1} + e_t)
\end{aligned}
\qquad (15.5.3)
$$

Substitute the lagged value $y_{t\text{-}2} = \mu + \beta_0 x_{t\text{-}2} + \beta_1 x_{t\text{-}3} + \gamma_1 y_{t\text{-}3} + e_{t\text{-}2}$ into Equation (15.5.3) to obtain

$$
\begin{aligned}
y_t &= \mu(1+\gamma_1+\gamma_1^2) + \beta_0 x_t + (\beta_1 + \gamma_1\beta_0)x_{t-1} + \gamma_1(\beta_1 + \gamma_1\beta_0)x_{t-2} + \gamma_1^2\beta_1 x_{t-2} \\
&\quad + \gamma_1^3 y_{t-3} + (\gamma_1^2 e_{t-2} + \gamma_1 e_{t-1} + e_t)
\end{aligned}
\qquad (15.5.4)
$$

Continue this process, and assuming that $|\gamma_1| < 1$, we obtain in the limit

$$
y_t = \alpha + \beta_0 x_t + \sum_{i=1}^{\infty} \gamma_1^{(i-1)}(\beta_1 + \gamma_1\beta_0)x_{t-1} + u_t \qquad (15.5.5)
$$

where $\alpha = \mu(1+\gamma_1+\gamma_1^2+\gamma_1^3+\ldots) = \mu/(1-\gamma_1)$ and $u_t = e_t + \gamma_1 e_{t-1} + \gamma_1^2 e_{t-2} + \gamma_1^3 e_{t-3} + \ldots$.

Equation (15.5.5) is an infinite distributed lag model,

$$y_t = \alpha + \sum_{i=0}^{\infty} \alpha_i x_{t-1} + u_t \tag{15.5.6}$$

with lag weights

$$\begin{aligned} \alpha_0 &= \beta_0 \\ \alpha_1 &= (\beta_1 + \gamma_1 \beta_0) \\ \alpha_2 &= \gamma_1(\beta_1 + \gamma_1 \beta_0) = \gamma_1 \alpha_1 \\ \alpha_3 &= \gamma_1^2 \alpha_1 \\ &\vdots \\ \alpha_s &= \gamma_1^{(s-1)} \alpha_1 \end{aligned} \tag{15.5.7}$$

- Estimating the ARDL(1, 1) model yields an infinite lag model with weights given by Equation (15.5.7).
- Similarly, the ARDL(2, 2) model, given by

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + e_t \qquad (15.5.8)$$

yields the infinite lag Equation (15.5.6) with weights

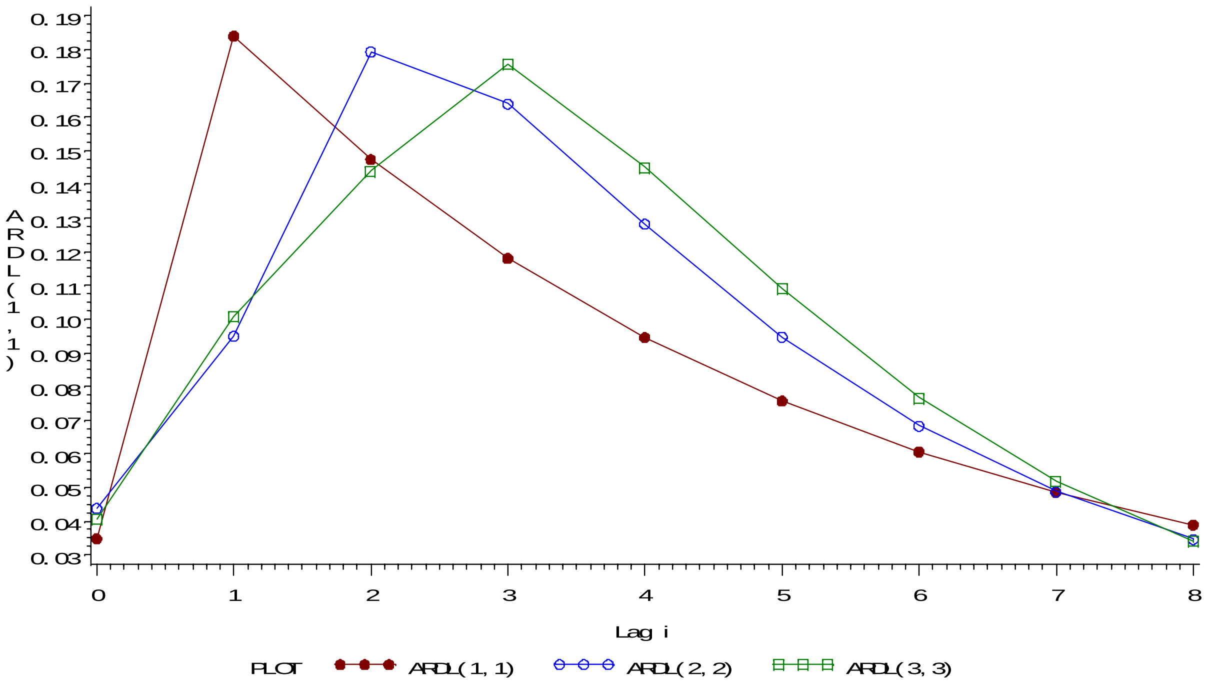
$$\begin{aligned}
\alpha_0 &= \beta_0 \\
\alpha_1 &= (\beta_1 + \gamma_1 \beta_0) \\
\alpha_2 &= \alpha_0 \gamma_2 + \alpha_1 \gamma_1 + \beta_2 \\
\alpha_3 &= \alpha_2 \gamma_1 + \alpha_1 \gamma_2 \\
\alpha_4 &= \alpha_3 \gamma_1 + \alpha_2 \gamma_2 \\
&\vdots \\
\alpha_s &= \alpha_{s-1} \gamma_1 + \alpha_{s-2} \gamma_2
\end{aligned} \qquad (15.5.9)$$

- It can be shown that the infinite lag arising from the ARDL(p, q) model is flexible enough to approximate any shape infinite lag distribution with sufficiently large values of p and q.

15.5.2 An Illustration of the ARDL Model

- To illustrate the estimation of an infinite ARDL, let us use the capital expenditure data in Table 15.1. Figure 15.5 shows the first eight lag weights from three alternative ARDL models.

Figure 15.5 ARDL models



- We see that unlike a geometric lag, the lag weights implied by the ARDL models capture the delay in the peak lag effect.

- As the order of the ARDL(p, q) increases, the lag weights exhibit a more flexible shape, and the peak effect is further delayed.
- The ARDL(3, 3) model yields lag weights not unlike the polynomial distributed lag of order two, shown in Figure 15.3; one difference is that the maximum weight is now at lag 3 instead of lag 4, which is more in line with the unrestricted lag weights.

Exercise

15.2	15.3	15.4		