$$\hat{S} = 579.57 + 6.66\ T \qquad \text{in quarter III}$$
$$\hat{S} = 622.71 + 6.66\ T \qquad \text{in quarter IV}$$

These results remain unchanged when four dummies are used, one for each of the four seasons, *but* the constant from the regression equation is dropped. Using the four seasonal dummies and the constant together would make it impossible to estimate the OLS regression (see Sec. 9.2).

DISTRIBUTED LAG MODELS

8.10 (*a*) What is meant by a *distributed lag model*? (*b*) Write the equation for a general distributed lag model with an infinite number of lags and for one with k lags. (*c*) What practical difficulties arise in estimating a distributed lag model with k lags?

(*a*) Often the effect of a policy variable may be distributed over a series of time periods (i.e., the dependent variable may be "sluggish" to respond to a policy change), requiring a series of lagged explanatory variables to account for the full adjustment process through time. A *distributed lag model* is one in which the current value of the dependent variable Y_t depends on the weighted sum of present and past values of the independent variables (X_t, X_{t-1}, X_{t-2}, etc.) and the error term, with generally different weights assigned to various time periods (usually declining successively for earlier time periods).

(*b*)
$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + \cdots + u_t \tag{8.9}$$
$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + \cdots + b_k X_{t-k} + u_t \tag{8.9a}$$

Note that in Eqs. (*8.9*) and (*8.9a*), a is constant, while b_0 is the coefficient of X_t. This has been done in order to simplify the algebraic manipulation in Prob. 8.11(*a*).

(*c*) In the estimation of a distributed lag model, the inclusion of each lagged term uses up one degree of freedom. When the number of independent lagged terms k is small, the model can be estimated with OLS, as done in Chap. 7. However, with k large (in relation to the length of the time series), an inadequate number of degrees of freedom may be left to estimate the model or to be confident in the estimated parameters. Moreover, the lagged explanatory variables in a distributed lag model are likely to be strongly correlated, so it may be difficult to adequately separate their independent effects on the dependent variable [see Prob. 7.3(*b*)].

8.11 (*a*) Derive the Koyck distributed lag model. (*b*) What problems arise in the estimation of this model? (*Hint for part a*: Start with the general distributed lag model and assume that the weights decline geometrically, with λ referring to a constant larger than 0 and smaller than 1; then lag the relationship by one period, multiply by λ, and subtract it from the original relationship.)

(*a*) Starting with Eq. (*8.9*), it is assumed that all the usual assumptions of OLS are satisfied (see Prob. 7.1):

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + \cdots + u_t \tag{8.9}$$

Geometrically declining weights and $0 < \lambda < 1$ gives

$$b_i = \lambda^i b_0 \qquad i = 1, 2, \ldots \tag{8.16}$$

Substituting Eq. (*8.16*) into Eq. (*8.9*), we obtain

$$Y_t = a + b_0 X_t + \lambda b_0 X_{t-1} + \lambda^2 b_0 X_{t-2} + \cdots + u_t$$

Lagging by one period, we have

$$Y_{t-1} = a + b_0 X_{t-1} + \lambda b_0 X_{t-2} + \lambda^2 b_0 X_{t-3} + \cdots + u_{t-1}$$

Multiplying by λ yields

$$\lambda Y_{t-1} = \lambda a + \lambda b_0 X_{t-1} + \lambda^2 b_0 X_{t-2} + \cdots + \lambda u_{t-1}$$

and subtracting from Eq. (*8.9*) yields

$$Y_t - \lambda Y_{t-1} = a - \lambda a + b_0 X_t + \lambda b_0 X_{t-1} - \lambda b_0 X_{t-1} + \lambda^2 b_0 X_{t-2} - \lambda^2 b_0 X_{t-2} + \cdots + u_t - \lambda u_{t-1}$$

$$Y_t - \lambda Y_{t-1} = a(1-\lambda) + b_0 X_t + u_t - \lambda u_{t-1}$$

$$Y_t = a(1-\lambda) + b_0 X_t + \lambda Y_{t-1} + v_t \qquad (8.10)$$

where $v_t = u_t - \lambda u_{t-1}$. Note that in Eq. (*8.10*) the number of regressors has been reduced to only two, with only one X.

(*b*) Two serious problems arise in the estimation of a Koyck distributed lag model. First, if u_t in Eq. (*8.9*) satisfies all the OLS assumptions (see Prob. 6.4), then $v_t = u_t - \lambda u_{t-1}$ in Eq. (*8.10*) does not. Specifically, $E(v_t v_{t-1}) \neq 0$ because v_t and v_{t-1} are both defined with u_{t-1} in common (i.e., $v_t = u_t - \lambda u_{t-1}$ and $v_{t-1} = u_{t-1} - \lambda u_{t-2}$). In addition, $E(v_t Y_{t-1}) \neq 0$. Violations of these OLS assumptions result in biased and inconsistent estimators for the Koyck lag model [Eq. (*8.10*)], requiring elaborate correction procedures (some of which are discussed in Sec. 9.3). The second serious problem is that the Koyck model rigidly assumes geometrically declining weights. This may seldom be the case in the real world, thus requiring a more flexible lag scheme (see Prob. 8.13).

8.12 Table 8.15 gives the level of inventories Y and sales X (in billions of dollars) in U.S. manufacturing from 1981 to 1999. (*a*) Fit the Koyck model to the data in Table 8.15. (*b*) What is the value of $\hat{\lambda}$ and $\hat{\alpha}$?

Table 8.15 Inventories and Sales in U.S. Manufacturing, 1981–1999 (in Billions of Dollars)

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Y	546	574	590	650	664	663	710	767	815	841
X	345	344	396	417	428	445	473	522	533	542
Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	
Y	835	843	870	935	996	1014	1062	1100	1151	
X	542	585	609	672	701	730	769	797	872	

Source: St. Louis Federal Reserve (U.S. Department of Commerce, Census Bureau).

(*a*)
$$\hat{Y}_t = 88{,}426.14 + \underset{(4.49)}{0.60}\ X_t + \underset{(4.22)}{0.50}\ Y_{t-1} \qquad R^2 = 0.99$$

(*b*)
$$\hat{\lambda} = 0.50 \quad \text{and} \quad \hat{\alpha}(1-0.50) = 88{,}426.14, \quad \text{so} \quad \hat{\alpha} = 176{,}852.28$$

8.13 (*a*) What is the lag structure in the Almon lag model? (*b*) What are the advantages and disadvantages of the Almon lag model with respect to the Koyck model?

(*a*) While the Koyck lag model assumes geometrically declining weights, the Almon lag model allows for any lag structure, to be approximated empirically by a polynomial of degree at least one more than the number of turning points in the function. For example, a lag structure of the form of an inverted U (i.e., with $b_1 > b_0$) can be approximated by a polynomial of at least the second degree. This may arise, as in the case of an investment function, when because of delays in recognition and in making decisions, the level of investment in the current period is more responsive to demand conditions in a few earlier periods than in the current period.

(*b*) The Almon lag model has at least two important advantages with respect to the Koyck lag model. First (and as pointed out earlier), the Almon model has a flexible lag structure as opposed to the rigid lag structure of the Koyck model. Second, since the Almon lag model does not replace the lagged independent variables (the Xs) with the lagged dependent variable, it does not violate any of the OLS assumptions (as does the Koyck model). One disadvantage of the Almon model is that the number of coefficients to be estimated is not reduced by as much as in the Koyck model. Another disadvantage is that in actual empirical work, neither the period nor the form of the lag may be suggested by theory or be known a priori.

8.14 Derive the Almon transformation for (*a*) a three-period lag taking the form of a second-degree polynomial and (*b*) a four-period lag taking the form of a third-degree polynomial.

(*a*) Starting with Eqs. (*8.11*) and (*8.12*)

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + b_3 X_{t-3} + u_t \tag{8.11}$$

$$b_i = c_0 + c_1 i + c_2 i^2 \qquad \text{with } i = 0, 1, 2, 3 \tag{8.12}$$

and substituting Eq. (*8.12*) into Eq. (*8.11*), we get

$$Y_t = a + c_0 X_t + (c_0 + c_1 + c_2)X_{t-1} + (c_0 + 2c_1 + 4c_2)X_{t-2} + (c_0 + 3c_1 + 9c_2)X_{t-3} + u_t$$

Rearranging the terms in the last expression:

$$Y_t = a + c_0\left(\sum_{i=0}^{3} X_{t-i}\right) + c_1\left(\sum_{i=1}^{3} iX_{t-i}\right) + c_2\left(\sum_{i=1}^{3} i^2 X_{t-i}\right) + u_t$$

and letting $Z_{1t} = \sum_{i=0}^{3} X_{t-i}$, $Z_{2t} = \sum_{i=1}^{3} iX_{t-i}$, and $Z_{3t} = \sum_{i=1}^{3} i^2 X_{t-i}$, we get

$$Y_t = a + c_0 Z_{1t} + c_1 Z_{2t} + c_2 Z_{3t} + u_t \tag{8.13}$$

(*b*) With a four-period lag taking the form of a third-degree polynomial, we have

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + b_3 X_{t-3} + b_4 X_{t-4} + u_t$$

$$b_i = c_0 + c_1 i + c_2 i^2 + c_3 i^3 \qquad \text{with } i = 0, 1, 2, 3, 4$$

Substituting the second into the first, we get

$$\begin{aligned} Y_t = a + c_0 X_t &+ (c_0 + c_1 + c_2 + c_3)X_{t-1} + (c_0 + 2c_1 + 4c_2 + 8c_3)X_{t-2} \\ &+ (c_0 + 3c_1 + 9c_2 + 27c_3)X_{t-3} + (c_0 + 4c_1 + 16c_2 + 64c_3)X_{t-4} + u_t \end{aligned}$$

Rearranging the terms in the last expression, we have

$$Y_t = a + c_0\left(\sum_{i=0}^{4} X_{t-i}\right) + c_1\left(\sum_{i=1}^{4} iX_{t-i}\right) + c_2\left(\sum_{i=1}^{4} i^2 X_{t-i}\right) + c_3\left(\sum_{i=1}^{4} i^3 X_{t-i}\right) + u_t$$

and letting the terms in parentheses equal, respectively, Z_{1t}, Z_{2t}, Z_{3t}, and Z_{4t}, we get

$$Y_t = a + c_0 Z_{1t} + c_1 Z_{2t} + c_2 Z_{3t} + c_3 Z_{4t} + u_t$$

8.15 Using the data from Table 8.15 and assuming a three-period lag taking the form of a second-degree polynomial, (*a*) Prepare a table with the original variables and the calculated Z values to be used to estimate the Almon lag model. (*b*) Regress the level of inventories, Y, on the Z values in the table in part *a*, i.e., estimate regression Eq. (*8.13*). (*c*) Find the $\hat{b}$ values and write out estimated Eq. (*8.11*).

(*a*) The Z values given in Table 8.16 are calculated as follows:

Table 8.16 Inventories, Sales, and Z Values in U.S. Manufacturing, 1981–1999 (in Billions of Dollars)

Year	Y	X	Z_1	Z_2	Z_3
1981	546	345	—	—	—
1982	574	344	—	—	—
1983	590	396	—	—	—
1984	650	417	1502	2119	4877
1985	664	428	1585	2241	5097
1986	663	445	1686	2450	5660
1987	710	473	1763	2552	5910
1988	767	522	1868	2647	6105
1989	815	533	1973	2803	6419
1990	841	542	2070	2996	6878
1991	835	542	2139	3174	7372
1992	843	585	2202	3225	7507
1993	870	609	2278	3295	7631
1994	935	672	2408	3405	7827
1995	996	701	2567	3645	8373
1996	1014	730	2712	3872	8870
1997	1062	769	2872	4148	9582
1998	1100	797	2997	4332	9998
1999	1151	872	3168	4525	10,443

$$Z_{1t} = \sum_{i=0}^{3} X_{t-i} = (X_t + X_{t-1} + X_{t-2} + X_{t-3})$$

$$Z_{2t} = \sum_{i=1}^{3} iX_{t-i} = (X_{t-1} + 2X_{t-2} + 3X_{t-3})$$

$$Z_{3t} = \sum_{i=1}^{3} i^2 X_{t-i} = (X_{t-1} + 4X_{t-2} + 9X_{t-3})$$

(*b*) Regressing Y on the Zs, we get

$$\hat{Y}_t = 171.80 + \underset{(2.20)}{0.44}\, Z_{1t} + \underset{(0.56)}{0.27}\, Z_{2t} - \underset{(-0.99)}{0.15}\, Z_{3t} \qquad R^2 = 0.99$$

(*c*)

$$\hat{\alpha} = 171.80$$
$$\hat{b}_0 = \hat{c}_0 = 0.44$$
$$\hat{b}_1 = (\hat{c}_0 + \hat{c}_1 + \hat{c}_2) = (0.44 + 0.27 - 0.15) = 0.56$$
$$\hat{b}_2 = (\hat{c}_0 + 2\hat{c}_1 + 4\hat{c}_2) = (0.44 + 0.54 - 0.60) = 0.38$$
$$\hat{b}_3 = (\hat{c}_0 + 3\hat{c}_1 + 9\hat{c}_2) = (0.44 + 0.81 - 1.35) = -0.10$$

so that $\hat{Y}_t = 171.80 + \underset{(2.20)}{0.44}\, X_t + \underset{(3.41)}{0.56}\, X_{t-1} + \underset{(2.31)}{0.38}\, X_{t-2} - \underset{(-0.47)}{0.10}\, X_{t-3}$

where the standard errors of the lagged values of X have been found by

$$\sqrt{\operatorname{var} \hat{b}_i} = \sqrt{\operatorname{var}(\hat{c}_0 + \hat{c}_1 i + \hat{c}_2 i^2)} \qquad (8.17)$$