5 Other Econometric Models

5.1 An Autoregressive Model

5.1.1. The United States Model (ARUS)

An easy model to work with for comparison purposes is one in which each endogenous variable is simply a function of its own lagged values. This model, which will be called an autoregressive model, consists of a set of completely unrelated equations. For the U.S. data I have used a lag length of 8 and have added a constant term and a time trend to the equation. Ten equations were estimated, one each for real GNP (*GNPR*), the GNP deflator (*GNPD*), the unemployment rate (*UR*), the bill rate (*RS*), the money supply ($M1$), the wage rate (W_f), profits (π_f), the savings rate (*SR*), the savings of the federal government (S_g), and the savings of the foreign sector (S_r).

The estimated equations are presented in Table 5-1. The first lag provides most of the explanatory power in these equations, which is typically the case with macro time series data. All the lags of length 1 are significant. Of the other lags, five of length 2 are significant (out of ten), one of length 3, two of length 4, two of length 5, one of length 6, two of length 7, and three of length 8. Five of the coefficient estimates of the time trend are significant.

5.1.2 The Multicountry Model (ARMC)

An autoregressive model was also estimated for the variables in the multicountry model. Each of the variables that appears on the LHS of a stochastic equation in the regular model was regressed on a constant, a time trend, three seasonal dummy variables, and the first four lagged values. The same estimation periods were used for these equations as were used for the equations in the regular model. Equations were not estimated for variables explained by definitions in the regular model. The accuracy of the MC and ARMC models is compared in Section 8.6.

TABLE 5-1. Estimated equations for the ARUS model

Explanatory variables	GNPR	GNPD	UR	RS	M1	W_f	π_f	SR	S_g	S_r
constant	47.5	-.00568	.00171	.0560	.916	-.0000122	.189	.0142	1.06	-.497
	(3.23)	(2.09)	(1.25)	(0.37)	(1.03)	(1.25)	(0.42)	(2.48)	(1.42)	(2.13)
t	.798	.0000406	.0000226	.0131	.0465	.000000298	.0274	.00000571	-.0255	.00470
	(2.90)	(1.42)	(2.15)	(2.85)	(2.13)	(1.45)	(2.06)	(0.29)	(1.79)	(1.54)
Lags:										
-1	1.213	1.542	1.640	1.274	.694	1.113	.901	.645	.945	.614
	(12.92)	(16.83)	(18.22)	(13.61)	(7.28)	(11.77)	(10.75)	(6.98)	(9.35)	(6.01)
-2	-.180	-.394	-.795	-.892	.463	.025	.091	.278	.097	.252
	(1.22)	(2.37)	(4.53)	(5.82)	(4.02)	(0.17)	(0.80)	(2.55)	(0.72)	(1.87)
-3	-.154	.128	.064	1.052	.127	-.037	-.150	-.052	-.132	-.164
	(1.02)	(0.75)	(0.33)	(6.05)	(1.04)	(0.26)	(1.20)	(0.46)	(0.98)	(1.29)
-4	.048	-.337	-.152	-.909	-.125	-.111	.361	-.122	.007	.019
	(0.32)	(1.89)	(0.76)	(4.42)	(1.02)	(0.78)	(2.82)	(1.07)	(0.05)	(0.15)
-5	-.011	-.146	.303	.823	-.208	.100	-.286	.033	-.163	-.178
	(0.07)	(0.81)	(1.47)	(3.69)	(1.64)	(0.69)	(2.27)	(0.29)	(1.16)	(1.40)
-6	.076	.107	-.009	-.733	.197	-.082	.141	.015	.188	.193
	(0.48)	(0.58)	(0.04)	(3.53)	(1.53)	(0.55)	(1.07)	(0.14)	(1.33)	(1.49)
-7	-.056	.104	-.217	.242	.211	.094	.424	-.170	-.038	.329
	(0.35)	(0.56)	(1.18)	(1.29)	(1.57)	(0.61)	(3.03)	(1.59)	(0.27)	(2.34)
-8	-.029	.004	.111	-.031	-.372	-.095	-.572	.166	.031	-.567
	(0.29)	(0.04)	(1.20)	(0.27)	(3.25)	(0.85)	(5.35)	(1.82)	(0.28)	(5.25)
SE	10.6	.00397	.00296	.703	2.44	.0000241	2.04	.00682	3.03	1.06
R^2	.999	.999	.958	.951	.999	.999	.981	.624	.845	.633
DW	1.98	1.99	1.97	2.01	1.85	1.94	1.80	2.04	1.90	1.88

Notes:
- Sample period is 1954 II - 1982 III (114 observations).
- Estimation technique is OLS.
- t-statistics in absolute value are in parentheses.

5.2 Two Vector Autoregressive Models (VAR1US and VAR2US)

Vector autoregressive models are also useful for comparison purposes, and two have been considered here. Both consist of five equations, explaining respectively the log of real GNP (log *GNPR*), the log of the GNP deflator (log *GNPD*), the unemployment rate (*UR*), the bill rate (*RS*), and the log of the money supply (log *M*1). For the first model the explanatory variables in each equation consist of a constant, a time trend, and the first six lagged values of each of the five variables, for a total of 32 coefficients to estimate per equation. For the second model the explanatory variables in each equation consist of a constant, a time trend, the first six lagged values of the own variable, and the first two lagged values of each of the other four variables, for a total of 16 coefficients to estimate per equation. For the second model each equation has a different set of RHS variables.

TABLE 5-2. Summary statistics for the VAR1US and VAR2US models

LHS variable	SE	R^2	DW	SE^a
VAR1US model:				
log GNPR	.00731	.9993	2.02	.00861
log GNPD	.00270	.9999	1.82	.00318
UR	.00238	.9730	1.98	.00280
RS	.544	.9709	1.98	.640
log M1	.00661	.9997	2.06	.00778
VAR2US model:				
log GNPR	.00804	.9992	1.95	.00867
log GNPD	.00310	.9999	1.79	.00334
UR	.00271	.9649	2.07	.00292
RS	.616	.9626	2.01	.664
log M1	.00790	.9996	2.01	.00851

Notes: a. Adjusted for degrees of freedom.
- Sample period is 1954 I - 1982 III (115 observations).
- Estimation technique is OLS.

The summary statistics for the two models are presented in Table 5-2. The SE's for VAR1US are only slightly lower than the SE's for VAR2US, and thus little explanatory power has been lost by excluding lags 3 through 6 of the variables other than the own variable. VAR2US has the advantage that many fewer coefficients are estimated per equation, and thus the degrees of freedom problem is considerably reduced. Vector autoregressive models in general have the problem of rapidly decreasing degrees of freedom as the number of variables is increased, and one way of dealing with this problem is to exclude all but the first two or so lags of the non-own variables in each equation. As just seen, little explanatory power is lost by following this approach. Another way of dealing with the degrees of freedom problem, which has not been pursued here, is to impose various constraints on the coefficients, either within or across equations.

5.3 A Twelve-Equation Linear Model (LINUS)

The twelve-equation linear model has eight stochastic equations and four identities. With respect to the use of economic theory in the model, it is somewhere between the US model and the autoregressive models; there is some theory behind the specifications, but it is very crude. The model is of interest in providing another basis of comparison for the US model. By comparing it to the US model, one can get an idea of how much gain there is (if any) in going from a simple theory to a more sophisticated one. It is also of interest to see how a model like this compares to the autoregressive models.

The equations are as follows.

1. $$CS = \underset{(3.05)}{-.447} + \underset{(106.37)}{.989}\ CS_{-1} + \underset{(3.24)}{.00945}\ GNPR - \underset{(8.19)}{.111}\ RS$$

[consumption of services]

$$SE = .260,\ R^2 = .999,\ DW = 2.13,\ \hat{\rho} = \underset{(2.58)}{-.229}$$

2. $$CN = \underset{(2.54)}{2.69} + \underset{(11.09)}{.800}\ CN_{-1} + \underset{(3.05)}{.0439}\ GNPR - \underset{(2.03)}{.0772}\ RS_{-1}$$

[consumption of nondurables]

$$SE = .493,\ R^2 = .999,\ DW = 1.94,\ \hat{\rho} = \underset{(2.03)}{.206}$$

3. $$CD = \underset{(3.83)}{-2.45} + \underset{(13.34)}{.760}\ CD_{-1} + \underset{(4.83)}{.0369}\ GNPR - \underset{(4.29)}{.210}\ RM_{-1}$$

[consumption of durables]

$$SE = .768,\ R^2 = .993,\ DW = 2.01$$

4. $$IH_h = \underset{(1.98)}{1.97} + \underset{(4.17)}{.505}\ IH_{h-1} + \underset{(4.37)}{.0259}\ GNPR - \underset{(4.75)}{.442}\ RM_{-1}$$

[housing investment, h]

$$SE = .395,\ R^2 = .975,\ DW = 1.96,\ \hat{\rho} = \underset{(9.10)}{.816}$$

5. $$Y = \underset{(4.35)}{9.93} + \underset{(3.64)}{.177}\ Y_{-1} + \underset{(17.20)}{.972}\ X - \underset{(4.32)}{.166}\ V_{-1}$$

[production]

$$SE = 1.16,\ R^2 = .999,\ DW = 2.19,\ \hat{\rho} = \underset{(5.82)}{.535}$$

6. $$IK_f = \underset{(4.53)}{-1.21} + \underset{(17.14)}{.822}\ IK_{f-1}\ \underset{(4.21)}{.00760}\ KK_{-1} + \underset{(2.88)}{.0592}\ Y - \underset{(0.79)}{.0200}\ Y_{-1}$$

[investment, f]

$$SE = .424,\ R^2 = .996,\ DW = 1.90$$

7. $$RM = .329 + .842\ RM_{-1} + .276\ RS - .066\ RS_{-1} - .025\ RS_{-2}$$
$$\quad (3.20)\quad (28.60)\qquad (7.32)\qquad (1.31)\qquad (0.72)$$
[mortgage rate]

$SE = .261,\ R^2 = .992,\ DW = 2.11$

8. $$RS = -.310 + .852\ RS_{-1} + .0557\ GNPR - .0527\ GNPR_{-1} + .0387\ \dot{M1}_{-1} + .132\ DD793 \cdot \dot{M1}_{-1}$$
$$\quad (0.89)\quad (14.24)\qquad (1.55)\qquad (1.41)\qquad (1.76)\qquad (3.92)$$
[bill rate]

$SE = .732,\ R^2 = .947,\ DW = 1.71$

9. $X = CS + CN + CD + IH_h + IK_f + Q_1$ [total sales]

10. $V = V_{-1} + Y - X$ [stock of inventories]

11. $GNPR = Y + Q_2$ [real GNP]

12. $KK = (1 - \delta_K)KK_{-1} + IK_f$ [capital stock]

Equations 1–4 are expenditure equations of the household sector. Each expenditure item is a function of its lagged value, real GNP, and either the short-term or the long-term interest rate. These equations differ from the expenditure equations in the US model in including real GNP and in excluding the price level, the wage rate, the initial value of assets, nonlabor income, and the labor constraint variable. The equations are also not in per-capita terms, and the housing investment equation does not include the lagged stock of housing. The GNP variable in these equations may capture some of the effects of the wage rate and the labor constraint variable in the US model. As discussed in Section 4.1.4, in periods of loose labor markets, when the labor constraint variable is not zero, the wage rate and the labor constraint variable are highly correlated with income.

The production equation, Eq. 5, is the same as Eq. 11 in the US model except for the exclusion here of the strike dummy variables. The investment equation, Eq. 6, is a simplified version of Eq. 12 in the US model. Investment is a function of its lagged value, the lagged value of the capital stock, and current and lagged output. No consideration is given here to the treatment of excess capital, which played an important role in the US model.

Equation 7 is a term structure equation explaining the mortgage rate. It is the same as Eq. 24 in the US model. The coefficient estimates in the two equations differ slightly as a result of the use of different sets of first-stage regressors in the estimation of the equations. Equation 8 explains the short-term interest rate, and it can be interpreted as an interest rate reaction function. It is a simplified version of Eq. 30 in the US model.

Equation 9 defines final sales, X. The variable Q_1, which is taken to be exogenous, is the difference in the data between X and $CS + CN + CS + IH_h + IK_f$. In other words, Q_1 is simply defined to make the definition hold. Equation 10 defines the stock of inventories; it is the same as Eq. 63 in the US model. Equation 11 relates production, Y, to real GNP. Again, the variable Q_2, which is taken to be exogenous, is simply the difference in the data between real GNP and Y. Equation 12 defines the capital stock; it is the same as Eq. 92 in the US model. The depreciation rate δ_K is taken to be exogenous.

The exogenous variables in the model other than Q_1, Q_2, and δ_K are $\dot{M1}_{-1}$ and $DD793 \cdot \dot{M1}_{-1}$. These last two variables, the percentage change in the money supply lagged one quarter and the same variable for the period 1979III and beyond, appear only in the interest rate reaction function.

The equations were estimated by 2SLS for the 1954I–1982III period. Equations 1, 2, 4, and 5 were estimated under the assumption of first-order serial correlation of the error term. The same set of first-stage regressors was used for each equation. The variables in this set in alphabetical order are as follows: constant term, CD_{-1}, CD_{-2}, CN_{-1}, CN_{-2}, CS_{-1}, CS_{-2}, $DD793 \cdot \dot{M1}_{-1}$, $DD793_{-1} \cdot \dot{M1}_{-2}$, $GNPR_{-1}$, $GNPR_{-2}$, IH_{h-1}, IH_{h-2}, IK_{f-1}, IK_{f-2}, KK_{-1}, KK_{-2}, $\dot{M1}_{-1}$, $\dot{M1}_{-2}$, Q_1, Q_2, RM_{-1}, RM_{-2}, RS_{-1}, RS_{-2}, RS_{-3}, V_{-1}, V_{-2}, Y_{-1}, Y_{-2}.

5.4 Sargent's Classical Macroeconomic Model (SARUS)

Sargent's (1976) model is an econometric version of the class of rational expectations models that was discussed in Section 3.1.7. It is an interesting model to consider both because it is the main empirical model of this class and because it incorporates the assumption of rational expectations. The assumption of rational expectations imposes difficult econometric problems, and Sargent's model is good for illustrating the estimation and solution methods presented in Chapter 11.

The model as Sargent estimated it is presented in Table 5-3. Sargent made two econometric mistakes in estimating this model: the first was to include variables in the regression to obtain $E_{t-1}P_t$ and in the first-stage regressions of the 2SLS technique that are not in the model; the second was to fail to note

TABLE 5-3. Sargent's model as originally estimated

Equation number	LHS variable	RHS variables
(1)	Un_t	$1,\ t,\ p_t - E_{t-1}p_t,\ Un_{t-i}\ (i = 1, \ldots, 4)$
(2)	nf_t	$1,\ t,\ p_t - E_{t-1}p_t,\ Un_t,\ nf_{t-i}\ (i = 1, \ldots, 4)$
(3)	y_t	$1,\ t,\ n_t,\ n_{t-i}\ (i = 1, \ldots, 4)$; filter: $(1 - .6L)^2$
(4)	R_t	$1,\ t,\ R_{t-i}\ (i = 1, \ldots, 4)$
(5c)	$m_t - p_t$	$1,\ t,\ R_t,\ R_{t-i}\ (i = 1, \ldots, 7),\ y_t,\ y_{t-i}\ (i = 1, \ldots, 7)$; filter: $(1 - .8L)^2$
(6)	n_t	$nf_t - Un_t + pop_t$

Notes:

- $E_{t-1}p_t$ was obtained from a regression of p_t on 1, t, three seasonal dummies, p_{t-i} $(i = 1, \ldots, 4)$, w_{t-i} $(i = 1, \ldots, 4)$, nf_{t-i} $(i = 1, \ldots, 4)$, and Un_{t-i} $(i = 1, \ldots, 4)$.
- The equations were estimated by 2SLS. The explanatory variables used in the first-stage regressions were those variables listed in the above note plus pop_t, m_t, the log of government purchases of goods and services in real terms, government surplus in real terms, and the log of current government employment. The RHS endogenous variables in the structural equations are p_t in equation (1), p_t and Un_t in equation (2), n_t in equation (3), and R_t and y_t in equation (5c).
- The filter $(1 - .6L)^2$ means that each variable z_t in the equation was transformed into $z_t^* = z_t - 1.2z_{t-1} + .36z_{t-2}$ before estimation. For the filter $(1 - .8L)^2$, the transformation is $z_t^* = z_t - 1.6z_{t-1} + .64z_{t-2}$.
- Variables:
  - Un_t = unemployment rate
  - nf_t = log of labor force participation rate
  - y_t = log of real GNP
  - R_t = long-term interest rate (Moody's Baa rate)
  - m_t = log of the money supply
  - p_t = log of the GNP deflator
  - pop_t = log of population
  - n_t = log of employment (approximately)
  - w_t = log of an index of a straight-time manufacturing wage.

that Eq. (5c) is not identified unless one assumes that the error terms in Eqs. (4) and (5c) are uncorrelated. If this assumption is made, then R_t can be treated as predetermined in the estimation of Eq. (5c). Sargent did not treat R_t as predetermined, and he should not have been able to estimate Eq. (5c) by 2SLS. The reason he did not encounter any difficulties is that he used more variables in the first-stage regression for R_t than he should have.

One way of dealing with these mistakes would be to expand the model to

TABLE 5-4. Sargent's model as estimated in this book

i) In place of using the filters, equations (3) and (5c) were estimated under the assumption of first- and second-order serial correlation of the error terms.

ii) The error term in equation (4) was assumed to be uncorrelated with the other error terms in the model, and R_t was taken to be predetermined in the estimation of equation (5c).

iii) There are two exogenous variables in the model, m_t and pop_t. Each of these was regressed on 1, t, and its first eight lagged values, and predicted values, $\hat{m}_t$ and $\widehat{pop}_t$, from these two regressions were taken to be the expected values.

iv) The model was estimated using the method in Chapter 11.

v) Data:

Name in Table 5-3:	Variable(s) in the US model:
Un_t	UR
nf_t	$\log[(L1 + L2 + L3 - J_m)/(POP - J_m)]$
y_t	log GNPR
R_t	RB
m_t	log M1
p_t	log GNPD
pop_t	$\log(POP - J_m)$

include more variables. For those who are interested in this kind of model, this would be interesting work. For present purposes, however, I have not chosen to expand the model; I have instead concentrated on obtaining estimates under the assumption that the model as presented in Table 5-3 is correctly specified.

The model as I have estimated it is presented in Table 5-4. The changes are as follows. (1) The variables that Sargent used in the first-stage regressions that are not in the model were excluded from consideration. (2) The error term in Eq. (4) was assumed to be uncorrelated with the other error terms in the model, and R_t was taken to be predetermined in the estimation of Eq. (5c). (3) In place of using the filters for Eqs. (3) and (5c), the equations were estimated under the assumption of first-order and second-order serial correlation of the error terms. Sargent's use of the filters is equivalent to constraining the first-order and second-order serial correlation coefficients to particular numbers, and thus the approach followed here is less restrictive. (4) The expected values of the two exogenous variables in the model, m_t and pop_t, were taken to be the predicted values from two eighth-order autoregressive equations. (5) Finally, the model was estimated by the method described in Chapter 11. This method, full information maximum likelihood, takes account of all the nonlinear restrictions that are implied by the rational expectations assumption.

It is not convenient to discuss the coefficient estimates of Sargent's model until the method in Chapter 11 has been described, and therefore the estimates will be presented and explained in Chapter 11.