deviation of this sample distribution, which will give us an estimated value for the standard error of the forecast. We can then use this standard error of forecast just as we did in Chapter 8 to determine forecast confidence intervals.

If the model is linear, then as the number of simulations included in the sample becomes large, the sample mean will approach the deterministic forecast (the forecast corresponding to all the random parameters set at 0). If the model is nonlinear, however, there is no guarantee that the sample mean will approach the deterministic forecast as the sample size increases, and in fact it usually will not. In addition, it may be necessary for an unacceptably large number of simulations to be performed before the sample means for each variable converge at all. We therefore center our confidence intervals on the *deterministic forecast* rather than on the sample mean of the stochastic simulations.

The process would be exactly the same if we wanted to forecast over a time horizon longer than one period. For each simulation we simply select a *different* random value for ε_{1t}, ε_{2t}, and ε_{3t} for *each period*, but we use the *same* random value for v_{11}, v_{12}, etc., during the entire simulation (since the equations of the model were specified and estimated under the assumption that the coefficients are constant over time). Furthermore, if the future values of the exogenous variables G_t and M_t were not known with certainty but had to be forecasted themselves, standard errors could be associated with their forecasts. The exogenous variables could then be treated as normally distributed random variables (with means equal to their forecasted values and standard deviations equal to their standard errors of forecast) in our stochastic simulation. For example, Eq. (13.40) could be rewritten to include an error term associated with G_t (which now must itself be forecasted):

$$Y_t = C_t + I_t + (\hat{G}_t + \eta_t) \tag{13.41}$$

Here $\hat{G}_t$ is a forecast of G_t, and η_t is a normally distributed random variable (defined for each period t) with mean 0 and standard deviation equal to the standard error of the forecast $\hat{G}_t$. Note that our forecast is now a *conditional* forecast.

APPENDIX 13.1 A Small Macroeconomic Model

As a more detailed example of a simulation model, we estimate and simulate a small model of the United States economy.[11] The model is intended to be illustrative and is therefore simple and highly aggregated, representing the economy in a way that is consistent with most introductory Keynesian macroeconomics textbooks. It does not represent the state of the art of macroeconomic modeling and forecasting, but it should be a good starting point for students interested in the development and use of multi-equation models.

[11] Our thanks to John Simpson and Randall Marris for their help in the development of this model.

TABLE A13.1
VARIABLES IN THE MODEL

Variables		Equation number
Endogenous:		
C	Consumption	(A13.4)
GNP	Gross national product	(A13.1)
IIN	Inventory investment	(A13.7)
INF	Inflation rate	(A13.10)
INR	Nonresidential investment	(A13.5)
IR	Residential investment	(A13.6)
RL	Long-term interest rate	(A13.9)
RS	Short-term interest rate	(A13.8)
TAX	Tax revenue	(A13.3)
UR	Unemployment rate	(A13.12)
WINF	Rate of growth of wage rate	(A13.11)
YD	Disposable income	(A13.2)
Exogenous:		
G	Government spending	
GNPP	Potential GNP	
RGM	Growth rate of the money supply	
POIL	Growth rate of oil price	
TR	Transfer payments	
WLTH	Household wealth	
XM	Net exports	

Model Specification and Estimation

The model contains 12 equations, of which 9 are behavioral, 2 are identities (for GNP and disposable income), and the last estimates a relationship between taxes and GNP. Each equation in the multi-equation model is estimated using some form of two-stage least squares. In particular, when there are lagged dependent variables as well as serial correlation in the error terms, Fair's method is used for estimation. (This method is discussed in Chapter 11.) We will describe the model and explain the theory underlying the specification on an equation-by-equation basis. A list of all the endogenous and exogenous variables of the model (with equation references) is given in Table A13.1.[12]

GNP and its components, consumption C, total investment I, government spending G, and net exports (exports minus imports) XM, are all in real (1982 dollar) terms. Total investment is disaggregated, and separate equations are estimated to explain nonresidential investment (fixed plant and equipment)

[12] All the data series are listed in the instructor's manual that accompanies this text. All data are seasonally adjusted when appropriate.

INR, residential investment IR, and investment in business inventories IIN. Thus the GNP identity is written as

$$\mathrm{GNP}_t = C_t + \mathrm{INR}_t + \mathrm{IR}_t + \mathrm{IIN}_t + G_t + \mathrm{XM}_t \tag{A13.1}$$

Disposable income is given by

$$\mathrm{YD}_t = \mathrm{GNP}_t - \mathrm{TAX}_t + \mathrm{TR}_t \tag{A13.2}$$

where TAX is the total tax flow, and TR is transfer payments. The tax mechanism in the model is extremely simple; the total tax flow is estimated as

$$\mathrm{TAX}_t = \underset{(-7.54)}{-137.074} + \underset{(41.29)}{.272}\mathrm{GNP}_t \tag{A13.3}$$

$$R^2 = .995 \qquad s = 12.087 \qquad \mathrm{DW} = 2.203 \qquad \hat{\rho} = .748$$

Let us now turn to the consumption function. As one would expect, consumption can be largely explained by disposable income. However, we would expect that the marginal propensity to consume is higher for income received from transfer payments than for the rest of disposable income. (The majority of transfer payments go to lower-income individuals who have a higher marginal propensity to consume.) We therefore separate these two components of disposable income. In addition, we expect consumption to depend on household wealth WLTH, as well as the short-term interest rate RS. (Higher interest rates increase the return from saving and also raise the cost of borrowing for large expenditures, thereby reducing consumption.) Our estimated equation for consumption is

$$\begin{aligned} C_t = & \underset{(-1.01)}{-11.161} + \underset{(4.94)}{.155}(\mathrm{YD}_t - \mathrm{TR}_t) + \underset{(5.15)}{.262}\mathrm{TR}_t + \underset{(3.40)}{9.716}\mathrm{WLTH}_t \\ & - \underset{(-5.06)}{2.590}\mathrm{RS}_t + \underset{(15.34)}{.739}C_{t-1} \end{aligned} \tag{A13.4}$$

$$R^2 = .999 \qquad s = 10.839 \qquad \mathrm{DW} = 1.86$$

Note that the estimated propensity to consume out of transfer payments is indeed considerably larger than out of the remainder of disposable income. Also, household wealth has a significant and positive impact on consumption, and the interest rate has a significant and negative impact.

We expect nonresidential investment to depend positively on aggregate economic activity, and negatively on long-term interest rates. Our estimated equation is

$$\mathrm{INR}_t = \underset{(-5.18)}{-22.418} + \underset{(6.24)}{.037}\mathrm{GNP}_t - \underset{(-2.82)}{1.174}\mathrm{RL}_{t-4} + \underset{(18.68)}{.776}\mathrm{INR}_{t-1} \tag{A13.5}$$

$$R^2 = .994 \qquad s = 7.186 \qquad \mathrm{DW} = 1.237$$

Residential investment is related to disposable income and the short-term interest rate (which serves as a proxy for the mortgage rate). Our estimated equation is

$$\mathrm{IR}_t = \underset{(-.47)}{-23.046} + \underset{(3.85)}{.069}\mathrm{YD}_t - \underset{(-2.11)}{1.005}\mathrm{RS}_{t-1} \tag{A13.6}$$

$$R^2 = .960 \qquad s = 6.735 \qquad \mathrm{DW} = 1.209 \qquad \hat{\rho} = .952$$

Inventory investment should depend on changes in output relative to changes in consumption. (Inventories accumulate when unanticipated increases in GNP exceed unanticipated increases in consumption.) The estimated equation is

$$\mathrm{IIN}_t = \underset{(2.25)}{3.582} + \underset{(10.97)}{.509}\Delta\mathrm{GNP}_t - \underset{(-6.39)}{.647}\Delta C_t + \underset{(15.73)}{.726}\mathrm{IIN}_{t-1} \tag{A13.7}$$

$$R^2 = .770 \qquad s = 9.374 \qquad \mathrm{DW} = 2.234$$

Note that both the change in GNP and the change in consumption have the correct signs and are highly significant.

Our equation for the short-term interest rate represents a standard LM curve. The demand for money increases when disposable income increases, but decreases when the interest rate (the price of holding money) rises. Hence the real (net of inflation) interest rate rises with disposable income and falls with increases in the money supply. The nominal interest rate is the real rate plus the expected rate of inflation. We therefore estimate the following equation:

$$\begin{aligned} \mathrm{RS}_t = & \underset{(-.49)}{-.805} + \underset{(3.51)}{.002}\mathrm{YD}_t - \underset{(-4.38)}{.116}(\Delta M_t/M_{t-1}) \\ & + \underset{(2.58)}{.105}(\mathrm{INF}_{t-1} + \mathrm{INF}_{t-2} + \mathrm{INF}_{t-3}) \end{aligned} \tag{A13.8}$$

$$R^2 = .856 \qquad s = 1.166 \qquad \mathrm{DW} = 2.371 \qquad \hat{\rho} = .745$$

Note that we use a moving average of inflation over the past three quarters as a proxy for expected future inflation. All three explanatory variables are statistically significant and have the expected signs.

The long-term interest rate responds with a geometric lag to the short-term rate:

$$\underset{}{RL_t} = \underset{(2.21)}{.249} + \underset{(7.03)}{.214RS_t} + \underset{(26.89)}{.793RL_{t-1}} \qquad \text{(A13.9)}$$

$$R^2 = .975 \qquad s = .475 \qquad DW = 2.103$$

We also estimate equations for the rates of price and wage inflation. We expect each of these variables to depend positively on the other. In addition, price inflation should depend on disposable income (reflecting demand pressure) as well as past rates of inflation. Finally, we introduce rate of growth of the price of oil as an exogenous variable; oil price inflation has been an important determinant of inflation in the United States and elsewhere. Our estimated equation is thus

$$INF_t = \underset{(-2.03)}{-1.231} + \underset{(5.58)}{.521WINF_t} + \underset{(2.21)}{.001YD_t} + \underset{(2.26)}{.011POIL_t}$$

$$+ \underset{(3.95)}{.187(INF_{t-1} + INF_{t-2})} \qquad \text{(A13.10)}$$

$$R^2 = .723 \qquad s = 1.540 \qquad DW = 2.294$$

In addition to price inflation, we expect the rate of wage inflation to depend on the unemployment rate. (A higher unemployment rate should reduce the bargaining power of workers for higher wage increases.) Our estimated equation is

$$WINF_t = \underset{(2.30)}{1.793} + \underset{(11.36)}{.795INF_t} - \underset{(-.34)}{.039UR_{t-3}} \qquad \text{(A13.11)}$$

$$R^2 = .537 \qquad s = 2.047 \qquad DW = 2.054$$

Note that the unemployment rate is statistically insignificant in this equation.

Finally, an equation is estimated for the unemployment rate (measured as a percent). This variable is related to the change in disposable income (and thus output) and to the difference between actual and potential GNP.

$$UR_t = \underset{(14.39)}{6.394} + \underset{(2.95)}{.003\Delta YD_t} - \underset{(-10.12)}{.013(GNP_t - GNPP_t)} \qquad \text{(A13.12)}$$

$$R^2 = .967 \qquad s = .296 \qquad DW = 1.833 \qquad \hat{\rho} = .938$$

Most of the explanation in this equation comes from the difference between actual and potential GNP. (The change in disposable income, however, has the wrong sign.)

TABLE A13.2
RESULTS OF SIMULATIONS

	1957-4 to 1987-4		1979-1 to 1987-4		1986-1 to 1987-4	
Variable	rms error	rms % error	rms error	rms % error	rms error	rms % error
C	16.541	.0101	17.979	.0082	55.84	.0224
INR	15.056	.0471	19.607	.0456	41.06	.0929
TAX	17.922	.0337	19.626	.0246	17.663	.0195
IR	20.759	.173	32.238	.275	8.255	.0412
IIN	11.91	11.661	13.453	21.12	11.825	50.930
RS	1.688	.304	2.684	.2728	3.050	.534
RL	.704	.0963	.960	.0935	1.180	.139
INF	1.508	.174	1.173	.564	1.845	1.061
WINF	2.021	.439	1.798	2.019	2.765	4.554
UR	.890	.175	.715	.0922	.596	.0987
GNP	37.828	.0138	3,414.4	.998	89.128	.0233

Simulation of the Model

The model can now be simulated as a complete system. Two historical simulations and an *ex post* forecast were performed in order to evaluate the model's ability to replicate the actual data.[13] The first simulation covered the entire estimation period (1957-4 to 1987-4). The second runs from 1979-1 to 1987-4. Finally, to perform an *ex post* forecast, the model was reestimated using data from 1957-4 through only 1985-4, and was then simulated from 1986-1 to 1987-4. The results of these simulations are summarized in Table A13.2. For each simulation, rms errors and rms percent errors are shown for all the endogenous variables. In addition, actual and simulated values for the endogenous variables are plotted in Figs. A13.1 to A13.9 for the historical simulation that covers the entire estimation period, and in Figs. A13.10 to A13.12 for the *ex post* forecast.

The model reproduces the general movements of most variables but does not fully capture many of the sharp fluctuations that occurred over the business cycle. For example, the general movement of nonresidential investment is captured, but the extent of the downturns that occurred during the 1975 and 1982 recessions is not reproduced. And the model fails almost completely to capture the sharp fluctuations that occurred in residential investment. Fortunately, the historical simulation reproduces much of the movement in inventory investment, wage and price inflation, and the unemployment rate. (However, the *ex post* forecast of unemployment misses the drop that occurred in 1986–1987.) Finally, although it does not reproduce the sharp increases that occurred in

[13] All simulations are *dynamic*, in the sense that simulated (rather than actual) values for the endogenous variables in a given period are used as inputs when the model is solved in future periods.