7 Solution

7.1 Definition of Terms

Once the stochastic equations of a model have been estimated and the identities have been written down, the next step is to solve the model. There are various meanings to the word "solve," and it will be useful to begin this discussion with some definitions. "Solve" and "simulate" mean the same thing. A "static" solution or simulation is one in which the actual values of the predetermined variables are used for the solution each period. Predetermined variables include both exogenous and lagged endogenous variables. A "dynamic" simulation is one in which the predicted values of the endogenous variables from the solutions for the previous periods are used for the values of the lagged endogenous variables for the solution for the current period.

"Forecast" and "prediction" are generally used to mean the same thing, and they are so used here. They mean the same thing as solution and simulation. An "outside-sample" forecast or prediction is one for a period that is not included within the estimation period; otherwise the forecast is "within-sample." An "ex post" forecast is one in which the actual values of the exogenous variables are used. An ex post forecast can be outside sample, but it must be within the period for which there are data on the exogenous variables. An "ex ante" forecast is made for a period beyond the period for which data exist; it is a forecast in which guessed values of the exogenous variables are used. In other words, ex ante forecasts are for a period that is truly unknown. Ex ante forecasts must be outside sample and (if the forecast is for more than one period ahead) dynamic. The forecasts must be dynamic because the values of the lagged endogenous variables are only known for the initial period.

In order to solve a model some assumption must be made about the error terms in the stochastic equations. If only one set of values of the error terms is used, the simulation is said to be "deterministic." The expected values of most error terms in most models are zero, and for most deterministic simulations the error terms are set to zero. For linear models the procedure of setting the error terms equal to their expected values and solving the model

results in the predicted values of the endogenous variables being equal to their expected values. This is not the case, however, for nonlinear models (see, for example, Howrey and Kelejian 1971), which is simply due to the fact that a nonlinear function of expected values is not equal to the expected value of the nonlinear function. A "stochastic" simulation is one in which many draws of the error terms are made in the process of solving the model. This procedure is discussed in Section 7.3. Aside from sampling error and a few other approximations, solving a nonlinear model by means of stochastic simulation does result in the predicted values being equal to the expected values. As will be seen in Chapters 8 and 9, stochastic simulation is useful for other purposes as well.

7.2 The Gauss-Seidel Technique

Most macroeconometric models are solved using the Gauss-Seidel technique. It is a remarkably simple technique and in most cases works remarkably well. This technique is used for all of the main procedures discussed in the rest of this book. The vast majority of computer time used for any of these procedures is spent solving the model using the Gauss-Seidel technique, and thus the technique is obviously of crucial importance. The technique is easiest to describe by means of an example.

Assume that the model (6.1) consists of three equations, and let x_{it} denote the vector of predetermined variables in equation i. The model is as follows:

$$f_1(y_{1t}, y_{2t}, y_{3t}, x_{1t}, \alpha_1) = u_{1t}, \tag{7.1}$$

$$f_2(y_{1t}, y_{2t}, y_{3t}, x_{2t}, \alpha_2) = u_{2t}, \tag{7.2}$$

$$f_3(y_{1t}, y_{2t}, y_{3t}, x_{3t}, \alpha_3) = u_{3t}, \tag{7.3}$$

where y_{1t}, y_{2t}, and y_{3t} are scalars. The technique requires that the equations be rewritten with each endogenous variable on the LHS of one equation. This is usually quite easy for macroeconometric models, since most equations have an obvious LHS variable. If, say, the LHS variable for (7.2) is $\log(y_{2t}/y_{3t})$, then y_{2t} can be written on the LHS by taking exponents and multiplying the resulting expression by y_{3t}. The technique does not require that each endogenous variable be isolated on the LHS; the LHS variable can also appear on the RHS. It is almost always possible in macroeconometric work, however, to isolate the variable, and this will be assumed in the following example.

The model (7.1)–(7.3) will be written

$$(7.1)' \qquad y_{1t} = g_1(y_{2t}, y_{3t}, x_{1t}, \alpha_1, u_{1t}).$$

$$(7.2)' \qquad y_{2t} = g_2(y_{1t}, y_{3t}, x_{2t}, \alpha_2, u_{2t}),$$

$$(7.3)' \qquad y_{3t} = g_3(y_{1t}, y_{2t}, x_{3t}, \alpha_3, u_{3t}).$$

In order to solve the model, values of the coefficients and the error terms are needed. It is unimportant for now what values are used, as long as some values are available. Given these values and given values of the predetermined variables, the solution proceeds as follows. Initial values of the endogenous variables are guessed. These are usually either actual values or predicted values from the previous period. Given these values, (7.1)′–(7.3)′ can be solved for a new set of values. This requires one "pass" through the model: each equation is solved once. One pass through the model is also called an "iteration." Given this new set of values, the model can be solved again to get another set, and so on. Convergence is reached if for each endogenous variable the values on successive iterations are within some prescribed tolerance level.

There are two main options that can be used when passing through the model. One is to use the values from the previous iteration for all the computations for the current iteration, and the other is to use, whenever possible, the values from the current iteration in solving the remaining equations. Following the second option in the example just given would mean using the current solution for y_{1t} in the solution of y_{2t} and y_{3t} and using the current solutions for y_{1t} and y_{2t} in the solution of y_{3t}. In most cases convergence is somewhat faster using the second option. If the second option is used, the order of the equations obviously matters in terms of the likely speed of convergence. The first option is sometimes called the Jacobi technique rather than the Gauss-Seidel technique, but for present purposes both options will be referred to as the Gauss-Seidel technique.

There is no guarantee that the Gauss-Seidel technique will converge. It is easy to construct examples in which it does not, and I have seen many examples in practice where it did not. The advantage of the technique, however, is that it can usually be made to converge (assuming an actual solution exists) with sufficient damping. By "damping" is meant the following. Let $\hat{y}_{1t}^{(n-1)}$ denote the solution value of y_{1t} for iteration $n-1$ (or the initial value if n is 1), and let $\hat{\hat{y}}_{1t}^{(n)}$ denote the value computed by solving (7.1) on iteration n. Instead of using $\hat{\hat{y}}_{1t}^{(n)}$ as the solution value for iteration n, one can instead adjust $\hat{y}_{1t}^{(n-1)}$ only partway toward $\hat{\hat{y}}_{1t}^{(n)}$:

$$(7.4) \qquad \hat{y}_{1t}^{(n)} = \hat{y}_{1t}^{(n-1)} + \lambda(\hat{\hat{y}}_{1t}^{(n)} - \hat{y}_{1t}^{(n-1)}), \qquad 0 < \lambda \leq 1.$$

If λ is 1, there is no damping, but otherwise there is. Damping can be done for any or all of the endogenous variables, and different values of λ can be used for different variables.

My experience is that one can usually make λ small enough to achieve convergence. The cost of damping is, of course, slow convergence. In some cases I have seen values as low as .05 needed. In the vast majority of the problems that I have solved, however, no damping at all was needed. Two other ways in which one can deal with problems of convergence are to try different starting values and to reorder the equations. This involves, however, more work than merely rerunning the problem with lower values of λ, and I have generally not found it necessary to experiment with starting values and the order of the equations.

Note that nothing is changed in the foregoing discussion if, say, y_{1t} is also on the RHS of (7.1)′. One still passes though the model in the same way. This generally means, however, that it takes longer to converge, and more damping may be required than if y_{1t} is only on the LHS; thus it is better to isolate variables on the LHS whenever possible.

The question of what to use for a stopping rule is not as easy at it might sound. The stopping rule can either be in absolute or percentage terms. In absolute terms it is

$$|\hat{y}_{it}^{(n)} - \hat{y}_{it}^{(n-1)}| < \epsilon_i \tag{7.5}$$

and in percentage terms it is

$$\left|\frac{\hat{y}_{it}^{(n)} - \hat{y}_{it}^{(n-1)}}{\hat{y}_{it}^{(n-1)}}\right| < \epsilon_i, \tag{7.6}$$

where ϵ_i is the tolerance criterion for variable i. (If damping is used, $\hat{y}_{it}^{(n)}$ in (7.5) and (7.6) should be replaced with $\hat{\hat{y}}_{it}^{(n)}$.)

The problem comes in choosing the values for ϵ_i. It is inconvenient to have to choose different values of the tolerance criterion for different variables, and one would like to use just one value of ϵ throughout. This is not, however, a sensible procedure if the units of the variables differ and if the absolute criterion is used. Setting the value of ϵ small enough for the required accuracy of the variable with the smallest units is likely to lead to an excess number of iterations, since a large number of iterations are likely to be needed to satisfy the criterion for the variables with the largest units. Setting ϵ greater than this value, on the other hand, runs the risk of not achieving the desired accuracy for some variables. This problem is lessened if the percentage criterion is used,

but in this case one must be concerned with variables, like the level of savings of a sector, that can be zero or close to zero.

My experience is that the number of iterations needed for convergence is quite sensitive to the stopping rule. It does not seem to be the case, for example, that once one has converged for most variables, one or two additional iterations increase the accuracy for the remaining variables very much. There is no real answer to this problem. One must do some initial experimentation to decide how many different values of ϵ are needed and whether to use the absolute or percentage criterion for a given variable.

7.3 Stochastic Simulation

7.3.1 The Basic Procedure

Stochastic simulation can be either with respect to the error terms or the coefficient estimates, or both. It requires that an assumption be made about the distributions of the error terms and/or coefficient estimates. In practice these distributions are almost always assumed to be normal, although in principle other assumptions can be made. For the present discussion the normality assumption will be used. In particular, it is assumed that $u_t = (u_{1t}, \ldots, u_{mt})'$ is independently and identically distributed as multivariate $N(0, S)$. This is the same assumption that was used for the FIML estimates in Chapter 6. Given an estimation technique and the data, one can estimate the coefficients, the covariance matrix of the coefficient estimates, and the covariance matrix of the error terms. Denote the estimates of the two covariance matrices $\hat{V}$ and $\hat{S}$ respectively. The dimension of $\hat{S}$ is $m \times m$, and the dimension of $\hat{V}$ is $k \times k$. $\hat{S}$ can be computed as $\frac{1}{T}\hat{U}\hat{U}'$, where $\hat{U}$ is the $m \times T$ matrix of values of the estimated error terms. The computation of $\hat{V}$ depends on the particular estimation technique used. Given $\hat{V}$ and given the normality assumption, an estimate of the distribution of the coefficient estimates is $N(\hat{\alpha}, \hat{V})$, where $\hat{\alpha}$ is the $k \times 1$ vector of coefficient estimates.

Let u_t^* denote a particular draw of the m error terms for period t from the $N(0, \hat{S})$ distribution, and let α^* denote a particular draw of the k coefficients from the $N(\hat{\alpha}, \hat{V})$ distribution. Given u_t^* for each period t of the simulation and given α^*, one can solve the model. This is merely a deterministic simulation for the given values of the error terms and coefficients. Call this simulation a "trial." Another trial can be made by drawing a new set of values of u_t^* for each period t and a new set of values of α^*. This can be done as many

times as desired. From each trial one obtains a prediction of each endogenous variable for each period. Let $\tilde{y}^j_{itk}$ denote the value on the jth trial of the k-period-ahead prediction of variable i from a simulation beginning in period t. For J trials, the estimate of the expected value of the variable, denoted $\tilde{\bar{y}}_{itk}$, is

$$\tilde{\bar{y}}_{itk} = \frac{1}{J}\sum_{j=1}^{J} \tilde{y}^j_{itk}. \tag{7.7}$$

Let σ^2_{itk} denote the variance of the forecast error for a k-period-ahead forecast of variable i from a simulation beginning in period t. Given the J trials, a stochastic-simulation estimate of σ^2_{itk} (denoted $\tilde{\sigma}^2_{itk}$) is

$$\tilde{\sigma}^2_{itk} = \frac{1}{J}\sum_{j=1}^{J} (\tilde{y}^j_{itk} - \tilde{\bar{y}}_{itk})^2, \tag{7.8}$$

where $\tilde{\bar{y}}_{itk}$ is determined in (7.7).

It is also possible to treat the coefficients as known and draw only from the distribution of the error terms. For a one-period-ahead forecast and known coefficients, the estimated variance is merely the estimated variance of the reduced form error term.

It should be stressed that these stochastic-simulation estimates of the means and variances are not exact. There are two reasons for this. The first is that the true distributions of the error terms and coefficient estimates are not known; one must always draw from estimated distributions. The second is sampling error that results from taking only a finite number of draws.

7.3.2 The Possible Nonexistence of Moments

It may be the case that the forecast means and variances do not exist, and this problem requires some discussion. For linear models Sargan (1976) has shown that for most overidentified models the 2SLS and 3SLS reduced form estimators have no moments of positive integral order. (A general theorem regarding the nonexistence of moments is given in Phillips 1984, theorem 3.9.1.) For linear models Sargan (1973) has also shown that the FIML reduced form estimates have finite moments of up to order T-K-G, where T is the number of observations, K is the number of exogenous variables in the model, and G is the number of endogenous variables in the model.

In practice, the possible nonexistence of moments is generally ignored: means and variances are estimated as if they always exist. One reason the nonexistence of moments does not appear to arise in practice is that extreme draws of the error terms and coefficient estimates are generally not used. By

"extreme" in this case is meant a draw that results in the failure of the Gauss-Seidel technique to find a solution of the model. In many of these cases it may be that with further damping and experimenting with the technique the solution could be found, but in some cases it may be that a solution truly does not exist. By throwing away the extreme draws, one is effectively sampling from truncated distributions, where the moments are likely to exist.

It is possible to compute more robust measures of central tendency and dispersion, such as the median, range, and interquartile range, and for some of the results in Chapter 8 I have reported measures like this. The measure of dispersion that I have used (denoted $\tilde{\delta}_{itk}$) is the following:

$$\tilde{\delta}_{itk} = \frac{\tilde{y}^b_{itk} - \tilde{y}^a_{itk}}{2}. \tag{7.9}$$

$\tilde{y}^a_{itk}$ is the value for which 34.135 percent of the J trial values lie *above* it and *below* the median, and $\tilde{y}^b_{itk}$ is the value for which 34.135 percent of the J trial values lie *below* it and *above* the median. For the normal distribution $\tilde{\sigma}_{itk}$ equals $\tilde{\delta}_{itk}$ except for sampling error, and thus the size of $\tilde{\delta}_{itk}$ is something that one may have some feeling for. Its size is similar to the size of the square root of the variance if the variance exists and if the true error distribution is close to being normal. Another way of looking at $\tilde{\delta}_{itk}$ is that it is like, say, the interquartile range except that $\tilde{y}^b_{itk} - \tilde{y}^a_{itk}$ encompasses 68.270 percent of the values rather than 50.0 percent of the values. If the variance does not exist for a particular problem and if the number of trials is large, one might expect $\tilde{\sigma}_{itk}$ to be considerably larger than $\tilde{\delta}_{itk}$. Therefore, by computing both measures one has at least a loose check on the possible nonexistence of moments.

Another approach to the problem of the possible nonexistence of moments is to modify an estimator in such a way that it is guaranteed to have moments. For linear models, for example, Maasoumi (1978) has proposed an estimator of the reduced form coefficients that is a weighted average of the unrestricted least squares estimator and the 3SLS estimator. The weight on the least squares estimator, which has finite moments, is nonzero when the two sets of estimates are far from each other according to a certain criterion. This way of truncating the 3SLS estimator is enough to ensure that the modified version has finite moments of up to order T-K-G, where T is the number of observations, K is the number of exogenous variables, and G is the number of endogenous variables.

It is not clear whether an approach like Maasoumi's can be extended to nonlinear models and whether it will be practical if it can. It may be that the main way in which this problem is dealt with in practice for large nonlinear

models is merely to truncate the distributions by not using extreme draws that occur during the stochastic simulations.

7.3.3 Numerical Procedures for Drawing Values

A standard way of drawing values of α^* from the $N(\hat{\alpha}, \hat{V})$ distribution is to (1) factor numerically (using a subroutine package) $\hat{V}$ into PP', (2) draw (again using a subroutine package) k values of a standard normal random variable with mean 0 and variance 1, and (3) compute α^* as $\hat{\alpha} + Pe$, where e is the $k \times 1$ vector of the standard normal draws. Since $Eee' = I$, then $E(\alpha^* - \hat{\alpha})(\alpha^* - \hat{\alpha})' = EPee'P' = \hat{V}$, which is as desired for the distribution of α^*. A similar procedure can be used to draw values of u_t^* from the $N(0, \hat{S})$ distribution: $\hat{S}$ is factored into PP', and u_t^* is computed as Pe, where e is a $m \times 1$ vector of standard normal draws.

An alternative procedure for drawing values of the error terms, derived from McCarthy (1972), has also been used in practice. For this procedure one begins with the $m \times T$ matrix of estimated error terms, $\hat{U}$. T standard normal random variables are then drawn, and u_t^* is computed as $T^{-\frac{1}{2}}\hat{U}e$, where e is a $T \times 1$ vector of the standard normal draws. It is easy to show that the covariance matrix of u_t^* is $\hat{S}$, where, as earlier, $\hat{S}$ is $\frac{1}{T}\hat{U}\hat{U}'$.

An alternative procedure is also available for drawing values of the coefficients. Given the estimation period (say, 1 through T) and given $\hat{S}$, one can draw T values of u_t^* ($t = 1, \ldots, T$). One can then add these errors to the model and solve the model over the estimation period (static simulation, using the original values of the coefficient estimates). The predicted values of the endogenous variables from this solution can be taken to be a new data base, from which a new set of coefficients can be estimated. This set can then be taken to be one draw of the coefficients. This procedure is more expensive than drawing from the $N(\hat{\alpha}, \hat{V})$ distribution, since reestimation is required for each draw, but it has the advantage of not being based on a fixed estimate of the distribution of the coefficient estimates. It is, of course, based on a fixed value of $\hat{S}$ and a fixed set of original coefficient estimates.

7.3.4 Previous Studies and Results

Stochastic simulation has not been widely used in practice, but a few studies do exist. Studies in which only draws from the distribution of the error terms have been made include Nagar (1969); Evans, Klein, and Saito (1972);

Fromm, Klein, and Schink (1972); Green, Leibenberg, and Hirsch (1972); Cooper and Fischer (1972); Sowey (1973); Cooper (1974); Garbade (1975); Bianchi, Calzolari, and Corsi (1976); and Calzolari and Corsi (1977). Studies in which draws from both the distribution of the error terms and the distribution of the coefficient estimates have been made include Schink (1971), (1974); Haitovsky and Wallace (1972); Cooper and Fischer (1974); Muench, Rolnick, Wallace, and Weiler (1974); and Fair (1980a).

One important empirical conclusion that can be drawn from these stochastic simulation studies is that the values computed from deterministic simulations are quite close to the mean predicted values computed from stochastic simulations. In other words, the bias that results from using deterministic simulation to solve nonlinear models appears to be small. This conclusion has been reached by Nagar (1969); Sowey (1973); Cooper (1974); Bianchi, Calzolari, and Corsi (1976); and Calzolari and Corsi (1977) for stochastic simulation with respect to the error terms only and by Fair (1980a) for stochastic simulation with respect to both error terms and coefficients. The results reported in Section 7.5.1 for the US model also confirm this conclusion.

7.4 Subjective Adjustment of Models

In actual forecasting situations most models are "subjectively adjusted" before the forecasts are computed. The adjustments take the form of either using values other than zero for the future error terms or using values other than the estimated values for the coefficients. Different values of the same coefficient are sometimes used for different periods. Adjusting the values of constant terms is equivalent to adjusting the values of the error terms, given that a different value of the constant term can be used each period. Adjustments of this type are sometimes called "add factors." One interpretation of add factors, which is stressed by Intriligator (1978, p. 516), is that they are the user's estimates of the future values of the error terms. With enough add factors it is possible to have the forecasts from a model be whatever the user wants, subject to the restriction that the identities must be satisfied. Most add factors are subjective in that the procedure by which they were chosen cannot be replicated by others. A few add factors are objective; for example, the procedure of setting the future values of the error terms equal to the average of the past two estimated values is an objective one. This procedure, along with another type of mechanical adjustment procedure, is used for some of the results in Haitovsky, Treyz, and Su (1974). (See Green, Liebenberg, and Hirsch 1972 for other examples.)

7.5 Computational Results

7.5.1 The US Model

The US model consists of 30 stochastic equations, 169 unrestricted coefficients, and 98 identities. The covariance matrix of the error terms *(S)* is thus 30×30, and the covariance matrix of the coefficient estimates *(V)* is 169×169.

For the solution of the model, the stopping rule for the Gauss-Seidel technique was taken to be in percentage terms and the tolerance value was chosen to be .001 percent. The first 30 equations, which are the stochastic equations, were used for the convergence check. If each of the successive predictions of the first 30 variables were within the tolerance value, convergence was taken to be achieved. Not checking the identities avoided the problem that some of the values of the variables determined by identities are close to zero. Experimentation with alternative (and more precise) stopping rules indicated that the procedure of checking only the first 30 variables provided sufficient accuracy. The number of iterations needed for convergence varied between about 7 and 13 for a typical job. The time taken to solve the model for one quarter was about .2 seconds on the IBM 4341 and about 1.5 seconds on the VAX. No damping was used for any of the variables for the Gauss-Seidel technique.

The results of solving the model for the 1978I–1979IV period are presented in Table 7-1. The 2SLS estimates were used for these results. The values in the 0 rows are predicted values from a deterministic simulation, where the error terms have been set equal to zero. The time for this simulation was about 1.6 seconds on the IBM 4341 (.2 seconds $\times$ 8 quarters) and about 12 seconds on the VAX (1.5 seconds $\times$ 8 quarters). The values in the a rows are predicted values from a stochastic simulation in which only error terms are drawn. Each trial for this simulation consists of 8 draws of 30 values each from the $N(0, \hat{S})$ distribution. A total of 250 trials were made. The cost of each trial is roughly the cost of solving the model once for the eight quarters. The total cost for the 250 trials, as noted at the bottom of Table 7-1, was about 6.7 minutes on the IBM 4341 and 49 minutes on the VAX.

The values in the b and b′ rows are predicted values from a stochastic simulation in which draws of both error terms and coefficients are made. The results in the two rows are based on the same simulation. The b-row values are mean values, and the b′-row values are median values. Each trial for this simulation consists of eight draws of 30 values each from the $N(0, \hat{S})$ distribution and one draw of 169 values from the $N(\hat{\alpha}, \hat{V})$ distribution. A total

TABLE 7-1. Predicted values for the US model: comparison of deterministic and stochastic simulation results

	1978				1979			
	I	II	III	IV	I	II	III	IV
GNPR: Real GNP (billions of 1972 dollars)								
0	1403.6	1425.0	1441.8	1458.6	1467.0	1472.4	1484.0	1490.9
a	1403.4	1427.7	1441.8	1458.1	1466.8	1472.1	1483.2	1489.5
b	1403.2	1424.4	1441.5	1458.6	1465.4	1469.8	1480.3	1487.4
b'	1404.0	1423.8	1441.9	1458.1	1464.1	1469.0	1481.3	1488.6
GNPD: GNP deflator (1972 = 1.0)								
0	1.4589	1.4928	1.5164	1.5494	1.5826	1.6153	1.6506	1.6825
a	1.4590	1.4927	1.5167	1.5499	1.5835	1.6163	1.6521	1.6838
b	1.4590	1.4926	1.5160	1.5489	1.5821	1.6148	1.6493	1.6811
b'	1.4590	1.4931	1.5166	1.5488	1.5823	1.6156	1.6495	1.6826
100·UR: Unemployment rate (percentage points)								
0	6.75	6.76	6.84	6.93	7.00	7.05	7.03	7.20
a	6.77	6.81	6.88	6.98	7.00	7.04	7.02	7.20
b	6.76	6.77	6.83	6.91	7.00	7.09	7.11	7.33
b'	6.75	6.77	6.86	6.92	7.00	7.08	7.05	7.32
RS: Bill rate (percentage points)								
0	6.48	7.06	7.45	7.77	8.13	8.24	10.11	11.01
a	6.42	7.03	7.39	7.79	8.13	8.29	10.21	11.02
b	6.50	7.15	7.58	7.92	8.28	8.36	10.11	10.91
b'	6.46	7.13	7.56	7.99	8.32	8.39	10.20	10.98
M1: Money supply (billions of current dollars)								
0	376.1	383.5	390.1	399.6	406.3	417.0	424.0	432.4
a	376.3	383.8	390.7	400.2	406.9	417.6	424.4	432.7
b	376.0	383.5	390.1	400.1	406.6	416.9	423.9	432.5
b'	376.1	383.1	389.6	399.7	406.0	416.9	423.8	432.5

Notes: 0 = Error terms set equal to zero (no stochastic simulation).
a = Stochastic simulation with respect to error terms only.
b = Stochastic simulation with respect to error terms and coefficient estimates.
b' = Same as b except values are median values rather than mean values.
• Prediction period is 1978 I - 1979 IV. All simulations are dynamic. Number of trials = 250.
• The time for one eight-quarter stochastic simulation of 250 trials was about 6.7 minutes on the IBM 4341 and about 49 minutes on the VAX.

of 250 trials were also made for this simulation. The total cost for the 250 trials was about the same as the cost of the 250 trials for the a-row simulation.

The main conclusion to be drawn from the results in Table 7-1 is that the predicted values from the deterministic simulation are quite close to the corresponding predicted values from the stochastic simulations. This, as noted in Section 7.3, is a common result. The bias that results from solving nonlinear models deterministically appears to be small for most models.

The other important conclusion from the results is that the median values are quite close to the corresponding mean values. In other words, the results

are not sensitive to the use of a more robust measure of central tendency. For none of the draws for the results in the table did the Gauss-Seidel technique fail to find a solution, and therefore no draws had to be discarded as being too extreme.

7.5.2 The MC Model

The solution of the MC model is a fairly large computational problem. For each of the 42 countries for which there are estimated equations (not counting the United States), there are up to 11 stochastic equations and 9 identities. In addition, there are 2,388 estimated trade share equations. The model is solved in the following way.

1. Given exports, $X75\$_i$, and the import price index, PM_i, country i's model is solved using the Gauss-Seidel technique. Each model consists of all or some subset of the 20 equations in Table B-3 (Appendix B).

2. Given the solution of each country's model, the calculations in Table B-4 (Appendix B), including the calculations of the trade shares, are performed. Table B-4 takes from each country the predicted value of imports, $M75\$A_i$, the predicted value of the export price index, PX_i, and the predicted value of the exchange rate, e_i. It returns to each country the predicted value of its exports, $X75\$_i$, the predicted value of its import price index, PM_i, and the predicted value of the world price index, $PW\$_i$.

3. Given $X75\$_i$ and PM_i from step 2, each country's model is solved again. The Table B-4 calculations are then performed again. This process is repeated until the successive predicted values from one iteration to the next are within some prescribed tolerance level.

This procedure consists of two types of iterations. The first is the standard Gauss-Seidel type for each country's model separately (step 1), and the second is the iteration between Tables B-3 and B-4 (step 3). The tolerance criterion for the second type of iteration should be greater than that for the first, since otherwise sufficient accuracy may not be achieved for the first type of iteration to achieve the required accuracy for the second.

This procedure worked quite well for the MC model. The average number of iterations for each country's model was usually less than 10, and the number of iterations of the second type varied between about 3 and 15. The total time taken to solve the model for one quarter varied between about 20 and 40 seconds on the IBM 4341 and about 2 and 4 minutes on the VAX. As noted earlier, the times for the US model for one quarter are .2 seconds on the

IBM 4341 and 1.5 seconds on the VAX. The MC model is thus considerably more expensive to solve than the US model. For this reason, no stochastic simulation experiments were performed for the MC model. Deterministic simulations were used to examine both the model's predictive accuracy and its properties. The accuracy is examined in Section 8.6, and the properties are discussed in Section 9.5.