YALE
INTERNATIONAL
CENTER
FOR FINANCE

Yale ICF Working Paper No. 00-41

OPTIMAL CONTROL AND STOCHASTIC SIMULATION OF LARGE NONLINEAR MODELS WITH RATIONAL EXPECTATIONS

Ray C. Fair
Yale University

May 2001

Yale SCHOOL of MANAGEMENT

Optimal Control and Stochastic Simulation of Large Nonlinear Models with Rational Expectations

Ray C. Fair*

May 2001

Abstract

This paper presents a computationally feasible procedure for the optimal control and stochastic simulation of large nonlinear models with rational expectations under the assumption of certainty equivalence.

1 Introduction

There is a large literature on examining the stabilization effectiveness of different interest rate rules. The general approach in this literature is to choose a rule and then use a model of the economy to examine how the economy would have behaved under the rule.[1] Sometimes optimal rules are derived by solving optimal control problems,[2] and these rules are compared to other rules.

*Cowles Foundation and International Center for Finance, Yale University, New Haven, CT 06520-8281. Voice: 203-432-3715; Fax: 203-432-6167; e-mail: ray.fair@yale.edu; website: http://fairmodel.econ.yale.edu.

[1]See, for example, Feldstein and Stock (1993), Hall and Mankiw (1993), Judd and Motley (1993), Clark (1994), Croushore and Stark (1994), Thorton (1995), Fair and Howrey (1996), Rudebusch (1999), Fair (2000), and Clarida, Galí, and Gertler (2000). Taylor (1985, fn. 1, p. 61) cites much of the literature prior to 1985.

[2]See, for example, Feldstein and Stock (1993), Fair and Howrey (1996), Rudebusch (1999), and Fair (2000).

This literature requires that the stochastic features of the economy be accounted for and that optimal control problems be solved. In addition, many of the economic models used have rational expectations. Solving optimal control problems for stochastic models with rational expectations is fairly involved, and almost all the recent studies have used small linear models. For example, only one of the studies in Taylor (1999)—Levin, Wieland, and Williams (1999) (LWW)—uses large scale models, and LWW do not solve optimal control problems. They use linearizations of the Federal Reserve model and the Taylor multicountry model to compute unconditional second moments of the variables in the models. In the recent study of Clarida, Galí, and Gertler (2000) a four equation calibrated model is used. Finan and Tetlow (1999) discuss the optimal control of large models with rational expectations, but their method is limited to linear models.

This paper discusses methods for the stochastic simulation and optimal control of large nonlinear models with rational expectations. The methods are computer intensive but computationally feasible given the current speed of computers. The results show that the analysis of interest rate rules and optimal policy need not be limited to the use of small linear models, even when the models have rational expectations.

The key approximation in this paper is the use of certainty equivalence for nonlinear models. It is argued below that this approximation seems good for most macroeconometric models, and some ways of examining the accuracy of the approximation are suggested. There undoubtedly are, however, applications in which the use of certainty equivalence gives poor results, and for these applications the methods in this paper are of little use.

This paper is also based on the assumption of known coefficients. It does not consider, for example, the possibility of unknown coefficients and learning. Amman and Kendrick (1999) consider this case within the context of the linear quadratic optimization problem for models with rational expectations. It would be interesting in future work to consider the case of unknown coefficients with learning in the more general setting of this paper.

For ease of reference, Table 1 lists some of the notation used in this paper.

2 The Model

The model will be written as:[3]

$$f_i(y_t, y_{t-1}, \ldots, y_{t-p}, E_{t-1}y_t, E_{t-1}y_{t+1}, \ldots, E_{t-1}y_{t+h}, x_t, \alpha_i) = u_{it} \quad (i = 1, \ldots, n) \tag{1}$$

where y_t is an n–dimensional vector of endogenous variables, x_t is a vector of exogenous variables, E_{t-1} is the conditional expectations operator based on the model and on information through period $t-1$, α_i is a vector of parameters, and u_{it} is an error term with mean zero that may be correlated across equations but not across time. The first m equations are assumed to be stochastic, with the remaining equations identities. The function f_i may be nonlinear in variables, parameters, and expectations.

[3]It is straightforward to generalize the model to include serially correlated errors: $u_{it} = \rho_i u_{it-1} + \epsilon_{it}$. See Fair and Taylor (1983, 1990) for a discussion of this case.

Table 1
Notation in Alphabetical Order

h	maximum lead
I	number of DFP iterations needed for convergence
J	number of stochastic simulation repetitions
k	extra periods beyond h needed for convergence
L	number of function evaluations needed for line searching
M	number of entire path comptations needed for convergence
N	number of one-period passes needed for convergence
q	number of control variables
Q	length of simulation period
R	length of optimal control horizon needed for first-period convergence
S	length of stochastic simulation period
T	length of optimal control period

3 Solution

Consider the solution of the model for period t. Assume that estimates of α_i are available, that current and future values of the exogenous variables are available, and that all values for periods $t-1$ and back are known. If the current and future values of the u_{it} error terms are set to zero (their expected values), the solution of (1) is straightforward. A popular method is the extended path (EP) method in Fair and Taylor (1983), which has been programmed into a number of computer packages. The method iterates over solution *paths*. Values of the expectations for period t through period $t+h+k+h$ are first guessed, where h is the maximum lead in the model and k is chosen as discussed below. Given these guesses, the model can be solved for periods t through $t+h+k$ in the usual ways (usually period by period using the Gauss-Seidel technique). This solution provides new values for the expectations through period $t+h+k$, namely the solution values. Given these

new values, the model can be solved again for periods t through $t+h+k$, which provides new values for the expectations, and so on. Convergence is reached when the predicted values for periods t through $t+h$ from one iteration to the next are within some tolerance level of each other. (There is no guarantee of convergence, but in most applications convergence is not a problem.)

In this process the guessed values of the expectations for periods $t+h+k+1$ through $t+h+k+h$ (the h periods beyond the last period solved) have not been changed. If the solution values for periods t through $t+h$ depend in a nontrivial way on these guesses, then overall convergence has not been achieved. To check for this, the entire process can be repeated for k one larger. If increasing k by one has a trivial effect (based on a tolerance criterion) on the solution values for t through $t+h$, then overall convergence has been achieved; otherwise k must continue to be increased until the criterion is met. In practice what is usually done is to experiment to find the value of k that is large enough to make it unlikely that further increases are necessary for any experiment that might be run and then do no further checking using larger values of k.

The solution requires values for x_t through x_{t+h+k}, the current and future values of the exogenous variables. These values are what the agents are assumed to know or expect at the beginning of period t. If agents are assumed not to have perfect foresight regarding x_t, then after convergence as described above has been achieved, one more step is needed. This step is to solve the model for period t using the computed expectations and the *actual* value of x_t, not the value that the agents expected. This is just a standard Gauss-Seidel solution for period t. To the extent that the expected value of x_t differs from the actual value, $E_{t-1}y_t$ will differ

from the final solution value for y_t. The final solution value for y_t is conditional on 1) the use of zero errors, 2) the actual value of x_t, and 3) the values of x_t through x_{t+h+k} that are used by the agents.

So far only the solution for period t has been described. In many cases one is interested in a dynamic simulation over a number of periods, say the Q periods t through $t+Q-1$. If it is assumed that all exogenous variable values are known by the agents, this simulation can be performed with just one use of the EP method, where the path is from t through $t+Q-1+h+k$ rather than just t through $t+h+k$. With known exogenous variables, the solution values for the expectations are the same as the overall solution values, and so if convergence is reached for the expectations for periods t through $t+Q-1+h$, the model has been solved for periods t through $t+Q-1$.

If the actual values of the exogenous variables differ from those used by the agents, then Q separate uses of the EP method are required to solve for t through $t+Q-1$. It is no longer the case, for example, that $E_{t-1}y_{t+1}$ equals $E_t y_{t+1}$ because the information sets through periods $t-1$ and t differ. The latter includes knowledge of x_t and the former does not. For simplicity the rest of this paper will only consider the case in which agents know the exogenous variables. It is straightforward but somewhat tedious to incorporate the case in which the exogenous variables are not known.

A useful way of estimating the computational cost of the EP method is to calculate the number of "passes" through the model that are used. A pass using the Gauss-Seidel technique is going through the equations of the model once for a given period and computing the values of the left hand side variables given the

values of the right hand side variables. Let N denote the number of passes that are needed to obtain Gauss-Seidel convergence for a given period, and let M denote the number of times the entire path has to be computed to obtain overall convergence (assuming that k has been chosen large enough ahead of time). Then the total number of passes that are needed to solve the model for the Q periods t through $t+Q-1$ is $N \cdot M \cdot (Q+h+k)$, since the path consists of $Q+h+k$ periods. If the model does not have rational expectations, the total number of passes is just $N \cdot Q$.

4 Optimal Control

The solution of optimal control problems for large scale models is fairly easy using certainty equivalence. Assume that the period of interest is t through $t+T-1$ (a horizon of length T) and that the objective is to maximize the expected value of W, where W is

$$W = g(y_t, \ldots, y_{t+T-1}, x_t, \ldots, x_{t+T-1}) \tag{2}$$

In most applications the objective function is assumed to be additive across time, which means that (2) can be written

$$W = \sum_{s=t}^{t+T-1} g_s(y_s, x_s) \tag{3}$$

Let z_t be a q–dimensional vector of control variables, where z_t is a subset of x_t, and let z be the $q \cdot (T+h+k)$–dimensional vector of all the control values: $z = (z_t, \ldots, z_{t+T+h+k-1})$, where k is taken to be large enough for solution

convergence through period $t + T - 1$.[4] If all the error terms are set to zero, then for each value of z one can compute a value of W by first solving the model for $y_t, \ldots, y_{t+T-1}$ and then using these values along with the values for $x_t, \ldots, x_{t+T-1}$ to compute W in (2) or (3). Stated this way, the optimal control problem is choosing variables (the elements of z) to maximize an *unconstrained* nonlinear function. By substitution, the constrained maximization problem is transformed into the problem of maximizing an unconstrained function of the control variables:

$$W = \Phi(z) \tag{4}$$

where Φ stands for the mapping $z \longrightarrow y_t, \ldots, y_{t+T-1}, x_t, \ldots, x_{t+T-1} \longrightarrow W$. Given this setup, the problem can be turned over to a nonlinear maximization algorithm like DFP. For each iteration of the algorithm, the derivatives of Φ with respect to the elements of z, which are needed by the algorithm, can be computed numerically. An algorithm like DFP is generally quite good at finding the optimum for a typical control problem.[5]

Once the problem is solved, z_t^*, the optimal vector of control values for period t, is implemented. If, for example, the Fed is solving the control problem and there is one control variable—the interest rate—then the Fed would implement through open market operations the optimal value of the interest rate for period t. In the process of computing z_t^* the optimal values for periods $t + 1$ through $t + T + h + k - 1$ are also computed. Agents are assumed to know these values when they solve the model to form their expectations. For the Fed example, one

[4]Remember that the guessed values of the expectations for periods $t + T + h + k$ through $t + T + h + k + h - 1$ are never changed in the solution. k has to be large enough so that increasing it by one has a trivial effect on the relevant solution values.

[5]See Fair (1974) for various applications of this procedure.

can think of the Fed implementing the period t value of the interest rate and at the same time announcing the planned future values.

After z_t^* is implemented and period t passes, the entire process can be repeated beginning in $t+1$. In the present deterministic case, however, the optimal value of z_{t+1} chosen at the beginning of $t+1$ would be the same as the value chosen at the beginning of t, and so there is no need to reoptimize. Reoptimization is needed in the stochastic case, which is discussed in Section 6.

Each evaluation of W requires $N \cdot M \cdot (T+h+k)$ passes, since the path is of length $T+h+k$. Each iteration of the DFP algorithm requires $q \cdot (T+h+k)$ evaluations of W to compute the derivatives numerically and then a few more evaluations to do the line searching. Let L denote the number of evaluations that are needed for the line searching after the derivatives have been computed, and let I denote the total number of iterations of the DFP algorithm that are needed for convergence to the maximum. The total number of evaluations of W is thus $I \cdot (q \cdot (T+h+k)+L)$. Since from Section 3 the number of passes needed to solve a model for T periods is $N \cdot M \cdot (T+h+k)$, the total number of passes needed to compute z_t^* is $N \cdot M \cdot (T+h+k) \cdot I \cdot (q \cdot (T+h+k)+L)$.

5 Stochastic Simulation

Forget optimal control for now and assume that some (not necessarily optimal) control rule is postulated. The stabilization features of a rule can be examined using stochastic simulation. One first needs an estimate of typical shocks to the economy. Shocks can be estimated in one of two ways. The first is to use the estimated error

terms from the econometric model. If, for example, the estimation period is 160 quarters, there are 160 vectors of error terms. The stochastic simulation can be set up so there is a probability of 1/160 of drawing any particular vector for any particular period.

The second way is to draw error terms from an estimated distribution. Let $\hat{V}$ be an estimate of the covariance matrix V of the u_{it} error terms (V and $\hat{V}$ are $m \times m$). If the error terms are assumed to be multivariate normal with zero means, one can draw errors from the $N(0, \hat{V})$ distribution. For large models there may not be enough observations to estimate all the nonzero elements of V, and so zero restrictions may have to be imposed. The advantage of drawing the historical error vectors directly is that no distributional assumption has to be made and no zero restrictions have to be imposed.

Assume that the periods of interest are t through $t+S-1$. The steps to estimate the variances of the endogenous variables for these periods under the rule are as follows:

1. Let u_t^*, an m-dimensional vector, denote a particular draw of the m error terms for period t. This draw can either be from a set of historically estimated vectors or from an estimated distribution. Assume that agents know this draw but use zero values of the errors for periods $t+1$ and beyond. (This means that the certainty equivalence assumption is still being used for agents for future periods.) Then solve the model (with the rule included) for period t using the EP method. Record the solution values for period t.

2. Draw a vector of error terms for period $t+1$, u_{t+1}^*, and use these errors and the solution values for period t to solve the model for period $t+1$ using the EP method. For this solution agents are assumed to use zero values of the errors for periods $t+2$ and beyond. Record the solution values for period $t+1$.

3. Repeat step 2 for periods $t + 2$ through $t + S - 1$. This set of solution values is one repetition. From this repetition one obtains a prediction of each endogenous variable for periods t through $t + S - 1$.

4. Repeat steps 1 through 3 J times for J repetitions.

5. Let y_{it}^j denote the value on the jth repetition of variable i for period t. For J repetitions, the stochastic simulation estimate of the expected value of variable i for period t, denoted $\tilde{\mu}_{it}$, is

$$\tilde{\mu}_{it} = \frac{1}{J} \sum_{j=1}^{J} y_{it}^j \tag{5}$$

Let

$$\sigma_{it}^{2j} = (y_{it}^j - \tilde{\mu}_{it})^2 \tag{6}$$

The stochastic simulation estimate of the variance of variable i for period t, denoted $\tilde{\sigma}_{it}^2$, is then

$$\tilde{\sigma}_{it}^2 = \frac{1}{J} \sum_{j=1}^{J} \sigma_{it}^{2j} \tag{7}$$

In practice it is usually the case with macroeconometric models that $\tilde{\mu}_{it}$, the estimate of the expected value of y_{it}, is quite close to the predicted value of y_{it} based on setting all the error terms to zero (no stochastic simulation).[6] (The main reason for doing stochastic simulation is not to improve on the estimates of the expected values but to compute variances.) If this is true for a particular model, it suggests that the use of certainty equivalence may not be a bad approximation. In other words, the expectations that agents compute using certainty equivalence may not be too far from the expectations that they would compute if they did a complete stochastic simulation. The closeness of $\tilde{\mu}_{it}$ to the certainty equivalence

[6] See, for example, Fair (1984), Section 7.3.4.

prediction may thus serve as a rough guide for how much confidence to place in the use of certainty equivalence, although this is by no means a rigorous test.[7]

In the above steps agents are assumed to know the draw u_t^* when solving the model beginning in period t, to know the draw u_{t+1}^* when solving the model beginning in period $t+1$, and so on. The steps could be set up so that agents do not know these draws and use zero errors instead. In this case the expectations would be computed using all zero errors, and after this the model would be solved using these computed expectations and the drawn error vector. For reasons that will be clear in the next section, the focus here is on the case where the current period draw is known.

The total number of passes that are needed for the J repetitions is $J \cdot S \cdot N \cdot M \cdot (h+k)$, since each path is of length $h+k$ and there are $J \cdot S$ paths solved.

[7] It is difficult to find in the literature comparisons of truly optimal and certainty equivalent solutions. One example is in Binder, Pesaran, and Samiei (2000), who examine the finite horizon life cycle model of consumption under uncertainty. They consider the simple case of a negative exponential utility function, a constant rate of interest, and labor income following an arithmetic random walk. The following computations are based on the values: interest rate = .04, discount factor = .98, negative exponential utility parameter = .01, initial and terminal values of wealth = 500, initial value of income = 200, standard deviation of random walk error = 5. For these values the truly optimal and certainty equivalence solutions were computed. (I am indebted to Michael Binder for providing me with the solution code.) Let c_1^* denote the truly optimal first-period value of consumption, and let c_1^{**} denote the value computed under the assumption of certainty equivalence. For a life cycle horizon of 12 years, c_1^* is .30 percent below c_1^{**}. For 24 years it is .60 percent below; for 36 years it is .87 percent below, and for 48 years it is 1.09 percent below. Although these differences seem modest, it is not clear how much they can be generalized, given the specialized nature of the model, and so it is really an open question as to the restrictiveness of the certainty equivalence assumption for the general model in this paper.

6 Stochastic Simulation and Optimal Control

In the optimal control case the control rule is dropped and an optimal control problem is solved to determine the values of the control variables. The steps that are needed to estimate the variances of the endogenous variables in this case are similar to those in the previous section. The difference is that after each draw of the error vector an optimal control problem has to be solved. As in Section 5, assume that the periods of interest are t through $t + S - 1$. The steps are:

1. Draw u_t^* as in Section 5. Assume that both the control authority and the agents know this draw but use zero values of the errors for periods $t + 1$ and beyond. Given this draw and the zero future errors, solve the (deterministic) control problem beginning in period t as in Section 4. This solution produces z_t^*, the optimal value of the control vector for period t, which is implemented. Record the solution values for period t.

2. Draw a vector of error terms for period $t + 1$, u_{t+1}^*, and use these errors and the solution values for period t to solve the control problem beginning in period $t + 1$. For this problem the control authority and the agents are assumed to use zero values of the errors for periods $t + 2$ and beyond. This solution produces z_{t+1}^*, the optimal value of the control vector for period $t + 1$, which is implemented. Record the solution values for period $t + 1$.

3. Repeat step 2 for periods $t + 2$ through $t + S - 1$. This set of solution values is one repetition. From this repetition one obtains the implemented optimal values, $z_t^*,\ldots,z_{t+S-1}^*$, and a prediction of each endogenous variable for periods t through $t + S - 1$ based on these values.

4. Repeat steps 1 through 3 J times for J repetitions. $\tilde{\mu}_{it}$ and $\tilde{\sigma}_{it}^2$ can then be computed as in Section 5.

The variances computed in this section using optimal control can be compared to the variances computed in Section 5 using other rules. The steps are set up so that both procedures assume that agents know the current period draw of the error

terms. In addition, any rule used in Section 5 in effect knows the draw, as does the control authority in this section. The information sets are thus the same for the comparisons.

In step 1 a control problem is solved beginning in period t. In Section 4 the horizon of the control authority regarding the objective function was taken to be length T and values of the control variables were computed for periods t through $t+T+h+k-1$. In step 1, however, it may be possible to shorten the horizon. What step 1 needs are only the solution values for period t (including z_t^*), and the horizon only needs to be taken long enough so that increasing it further has a trivial effect (based on a tolerance criterion) on the values for period t. One can initially experiment with different values of the horizon to see how large it has to be to meet the tolerance criterion. Let R denote this length. This value of R can be used in step 2 for the control problem beginning in period $t+1$, and so on.

The overall procedure requires that S control problems be solved per repetition, and so with J repetitions there are $J \cdot S$ control problem solved, each with a horizon of length R. The total number of passes in this case is thus: $J \cdot S \cdot N \cdot M \cdot (R + h + k) \cdot I \cdot (q \cdot (R + h + k) + L)$. If the number of trials (J) is 20, the number of periods (S) is 16, the number of passes needed for convergence (N) is 5, the number of solution paths needed for convergence (M) is 5, the necessary horizon for the control problem (R) is 10, the lead length (h) is 3, the necessary value of k is 8, the number of DFP iterations needed for convergence (I) is 8, the number of control variables (q) is 1, and the number of line searches needed per DFP iteration (L) is 10, then the total number of passes needed is 41,664,000. At 11,573 passes per second, this could be done in 1 hour; at 5,787 passes per second, it could be

done in 2 hours; and so on.

In term of speed it is obviously important that efficient code be written for passing through the model, since most of the time is spent passing through. A practical way to proceed after the code is written is to set limits on N, M, I, and J that are small enough to make the problem computationally feasible (like completion within an hour or two). Once the bugs are out and the (preliminary) results seem sensible, the limits can be gradually increased to gain more accuracy. If two cases are being compared using stochastic simulation, such as a simple rule versus an optimal control procedure, the same draws of the errors should be used for both cases. This can considerably lessen stochastic simulation error for the comparisons.

The number of passes needed is much smaller if the model is not a rational expectations model. In this case M is 1 and h and k are zero, and in the above example the number of passes is 2,560,000. This is about 1/16th the number of passes for the rational expectations case.

The appendix presents an example of the use of the procedure in this section for a large nonlinear model with rational expectations.

7 Conclusion

This study has shown that it is computationally feasible to solve stochastic simulation and optimal control problems for large nonlinear models with rational expectations if certainty equivalence is used. The analysis of monetary and fiscal policies need not be restricted to the use of small models or linear models.

What is lost by the use of the open loop procedure of certainty equivalence and reoptimization in Section 6? Agents know when they solve the model to form their expectations the current period values of the control variables that are implemented and the announced planned future values. They take the planned future values as deterministic rather than stochastic, and they take the future error terms to be deterministic, namely zero. Agents do not take into account the fact that everything will be redone at the beginning of each period after the error terms for that period are realized and known. The overall procedure is thus not fully optimal. As mentioned in Section 1, in some cases this may be a serious problem, and in these cases the procedure in Section 6 is of little use.

Appendix

Coding

As noted in the text, it is important that efficient code be written to pass through the equations of a model. Let PASS(r) denote a subroutine written to pass through the model once for period r. Let SOLVE(s,Q) denote a subroutine written to solve a rational expectations model for periods s through s + Q -1 using the extended path method. SOLVE(s,Q) calls PASS(r) many times for r equal to s through s + Q - 1 + h + k, where h is the maximum lead and k is chosen as discussed in the text. Let DFP(s,R) denote a subroutine written to solve an optimal control problem with beginning period s and necessary horizon R (as discussed in Section 6). DFP(s,R) calls SOLVE(s,R) one time per evaluation of the objective function W. Finally, let DRAW(s) denote a subroutine written to draw a vector of error terms for period s. The outline of the program to do stochastic simulation and optimal control as in Section 6 is:

```
      DO 100 j = 1, J
      DO 200 s = t, t+S-1
      CALL DRAW(s)
      CALL DFP(s,R)
         Calls SOLVE(s,R) once per evaluation of W.
             Calls PASS(r) many times for r = s, s+R-1+h+k.
      Record predicted values on trial j for period s.
  200 CONTINUE
  100 CONTINUE
```

An Example

One of the models used for the results in Fair (2000) is a model of the United States with rational expectations. There are 31 stochastic equations and about 100 identities in this model. The model is nonlinear, and the maximum lead length (h) is 7. The model is estimated for the 1954:1–1999:2 period, which gives 182 vectors of historical error terms. Stochastic simulation was done for this model in Fair (2000), but optimal control experiments were not performed. For the example below both stochastic simulation and optimal control were done using the procedure in Section 6. Stochastic simulation was done by drawing for a given period one of the 182 vectors of error terms with probability 1/182.

The period was 1993:1–1995:4 ($S = 12$); the number of trials (J) was 20; k was taken to be 12; and the DFP iteration limit (I) was taken to be 6. No limits were imposed on N and M. The tolerance criterion for a Gauss-Seidel iteration was 0.1 percent, and the tolerance criterion for extended path convergence was 0.2 percent.

The example was run using the Fair-Parke (1995) program. The program is written in FORTRAN and includes all the necessary subroutines. The computer used was a Dell Pro 200 (which uses an early Pentium chip), purchased in May 1996. Computer chips have gotten much faster since this purchase date, and so the time given below would be considerably smaller on newer computers.

The time taken for this example was about 11 hours. The average number of passes per Gauss-Seidel iteration was about 5, and the average number of passes for extended path convergence was about 2. The DFP iteration limit of 6 was

binding in that in most cases it appeared that more iterations would be needed to achieve acceptable accuracy. If the number of iterations were doubled, the time would also roughly double.

This time of 11 hours on a fairly old computer shows that the procedure in Section 6 is in the realm of computational feasibility even for a nonlinear model of over 100 equations with a nontrivial lead length (i.e., 7). As mentioned in the text, a good approach is to set fairly small limits on the relevant parameters and then increase the limits to gain more accuracy after the bugs are worked out. One programming issue that is important is setting the step size for the numeric derivatives used by the DFP algorithm. The step size must be larger than the solution tolerance criteria in order for the computed derivatives to be any good. Some experimentation is usually needed to get this right.

References

[1] Amman, Hans M., and David A. Kendrick, 1999, "Linear Quadratic Optimization for Models with Rational Expectations and Learning," *Macroeconomic Dynamics*,

[2] Binder, Michael, M. Hashem Pesaran, and S. Hossein Samiei, 2000, "Solution of Nonlinear Rational Expectations Models with Applications to Finite-Horizon Life-Cycle Models of Consumption," *Computational Economics*, 15, 25-57.

[3] Clarida, Richard, Jordi Galí, and Mark Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115, 147-180.

[4] Clark, Todd E., 1994, "Nominal GDP Targeting Rules: Can They Stabilize the Economy?" *Federal Reserve Bank of Kansas City Economic Review*, 3, 11-25.

[5] Croushore, Dean, and Tom Stark, 1994, "Evaluating McCallum's Rule for Monetary Policy," Federal Reserve Bank of Philadelphia, November.

[6] Fair, R. C., 1974, "On the Solution of Optimal Control Problems as Maximization Problems, *Annals of Economic and Social Measurement*, 3, 135-154.

[7] _____, 1984, *Specification, Estimation, and Analysis of Macroeconometric Models*, Cambridge, MA: Harvard University Press.

[8] _____, 1994, *Testing Macroeconometric Models*, Cambridge, MA: Harvard University Press.

[9] _____, 2000, "Estimated, Calibrated, and Optimal Interest Rate Rules," mimeo, May.

[10] _____, and E. Philip Howrey, 1996, "Evaluating Alternative Monetary Policy Rules," *Journal of Monetary Economics*, 38, 173-193.

[11] _____, and William R. Parke, *The Fair-Parke Program for the Estimation and Analysis of Nonlinear Econometric Models*, August 1995.

[12] _____, and John B. Taylor, 1983, "Solution and Maximum Likelihood Estimation of Dynamic Rational Expectations Models," *Econometrica*, 51, 1169-1185.

[13] _____, and John B. Taylor, 1990, "Full Information Estimation and Stochastic Simulation of Models with Rational Expectations," *Journal of Applied Econometrics*, 5, 381-392.

[14] Feldstein, Martin, and James H. Stock, 1993, "The Use of a Monetary Aggregate to Target Nominal GDP," NBER Working Paper No. 4304.

[15] Finan, Frederico S., and Robert Tetlow, "Optimal Control of Large, Forward-Looking Models," Board of Governors of the Federal Reserve System, October 1999.

[16] Hall, Robert E., and N. Gregory Mankiw, 1993, "Nominal Income Targeting," NBER Working Paper No. 4439.

[17] Judd, John P., and Brian Motley, 1993, "Using a Nominal GDP Rule to Guide Discretionary Monetary Policy," *Federal Reserve Bank of San Francisco Economic Review*, 3, 3-11.

[18] Levin, Andrew, Volker Wieland, and John C. Williams, 1999, "Robustness of Simple Monetary Policy Rules under Model Uncertainty," in John B. Taylor, ed., *Monetary Policy Rules*, Chicago: The University of Chicago Press, 263-316.

[19] Rudebusch, Glenn D., 1999, "Is the Fed Too Timid? Monetary Policy in an Uncertain World," March.

[20] Taylor, John B., 1985, "What Would Nominal GDP Targeting Do to the Business Cycle?" *Carnegie-Rochester Series on Public Policy*, 22, 61-84.

[21] _____, ed., 1999, *Monetary Policy Rules*, Chicago: The University of Chicago Press.

[22] Thornton, Saranna R., 1995, "Adaptive Monetary Policy Rules Utilizing Either MZM or M2 and the Effects of Positive Bias in Money Control Errors," for presentation at the AEA Meetings, Washington D.C., January 1995.