Annals of Economic and Social Measurement, 3/1, 1974

A PRACTICAL METHOD FOR CONTROLLING A LARGE NONLINEAR STOCHASTIC SYSTEM

BY ROBERT S. HOLBROOK*

The paper reviews a practical technique for optimizing a quadratic objective function under the constraint of a large nonlinear econometric model, and extends that technique to the stochastic case. Several of the difficulties which will accompany any attempt to control a large model are discussed, and the optimization technique (for the deterministic case) is illustrated with an application to the Michigan Model of the U.S. economy.

I

Economic applications of optimal control theory, first in its deterministic and now in its stochastic form, are attracting great attention in the professional literature, and even in the press. In no area of economics is the anticipated payoff from this activity greater than in that of macroeconomic policy making.

Developments in macro-model building during the last decade would almost appear to have been designed with the aim of facilitating the use of control theoretic techniques. Except for some relatively simple cases, however, model builders did not utilize the techniques necessary to enable them to select an "optimal" policy. The array of forecasts, simulations, multipliers, and model evaluations they presented were certainly of great relevance to the needs of policy-making. But a policy-maker typically is concerned with a multidimensional policy decision which takes into account both current and future goals, a decision requiring some form of optimal control technique rather than the standard fare of simulation results commonly provided. Thus, it is not surprising that many economists are devoting their attention to these problems, and that some excellent papers dealing with them have appeared in the past few years (and this conference is the occasion for the appearance of several more).

My own interest in this area arose as a result of my attempt to devise optimal "rules of thumb" using RDX2, the Bank of Canada's model of the Canadian economy[1] (ultimately I found it easier simply to control the model in an optimal fashion than to develop such rules of thumb). Thus, my approach was—and is—from the point of view of a real world policy-maker, and not that of a control theoretician, as I'm sure will be evident in the course of this paper.

I was searching for a practical, simple, and cheap means of selecting an optimal path for a large nonlinear model. I believe that the method I developed while at the Bank of Canada satisfies these criteria.[2] The original report [5] dealt only with a deterministic version of the optimizing method. This paper extends it to the stochastic case.

[1] The model is described by Helliwell, *et al.* [4], and my experiments with it are reported in [5] and [6].

* I would like to thank Gregory Chow and my colleagues Saul Hymans and Harold Shapiro for their helpful comments. I also wish to thank Lockwood Lyon for his valuable programming assistance, and Ms. Kris Maki for her speedy and expert typing. Part of this research was supported by the NSF under grant GS-36932X.

In Section II I describe some of the difficulties confronting anyone who attempts to use a nonlinear model for policy making. I describe my optimizing procedure in its deterministic version briefly in Section III, and cite some new results from its use with the Michigan Model [8] in Section IV. Section V contains a theoretical extension to the stochastic case, but there is as yet no empirical evidence as to the method's success under those conditions. Section VI concludes the paper.

II

Many difficulties are presented by the problem of selecting an optimal macroeconomic policy for an actual economy. First, the selection of the objective function is usually subjective and often quite arbitrary. Although I doubt if any policy-maker could describe the function controlling his policy decisions, such a function is absolutely necessary before we can speak of or hope to derive an "optimal" policy. I have described elsewhere an approach which could be used in an attempt to decipher the policy-maker's views, and to capture them in functional form,[3] but at best the function finally chosen will be only a rough approximation to the truth. Since we are doomed to great uncertainty regarding the appropriate form, variables, and parameter values for this function, I believe that the many computational advantages of the quadratic form make it the obvious choice. Though its faults are well known, its advantages are so great as to outweigh them, in the absence of substantial evidence that some other form is more nearly correct.

Second, the actual economy for which policy is to be chosen is highly complex and is only imperfectly represented by even the largest of our macroeconomic models. This conference is concerned with one aspect of this imperfection, namely that the models are deterministic representations of a stochastic system. But there are other difficulties, due not to stochastic but to systematic errors in the models. When the model is used to simulate much beyond the fitting period, it usually will get off track rather quickly. Various adjustments, dummy variables, etc., may be used for repair purposes, but this can be done only after the actual data become available. Our primary interest is in planning policy for the future, however, and in this context no such corrections are possible until it is too late. For this reason (as well as others perhaps less well founded) no policy-maker is likely ever to base his actual decisions solely upon an optimal control calculation using a model of the economy. Instead he is likely to use the optimal control results as signals, noting their sign and order of magnitude, but ignoring everything beyond the first couple of significant digits. If this is true, it suggests that practical economic policy making does not require that the control problem be solved exactly. A cheaply and easily

[2] I was not surprised when I was informed by others more familiar with the control literature that the method I had developed was, in fact, a gradient method related to but not identical with the Newton–Raphson method as described by Polak [10].

[3] In [5] I suggested simply that the staff economist explore the policy-maker's preference map by means of a series of questions comparing hypothetical situations. Ann Friedlaender [2] has shown how, under certain assumptions, one can infer from historical data the coefficients of the policy-maker's objective function. Unfortunately, one of the assumptions required is that policy-makers aim for what they get. This may be appropriate for historical analysis (the use to which Friedlaender puts it) but it is not likely to be very useful in an actual policy-making context.

obtainable exact solution would be best, of course, but lacking that, a close approximation to the exact solution will probably be entirely adequate.

Third, leaving aside the question of how well the model represents the real world, there is an additional problem within the model itself. With few exceptions, most of the important models in use today consist in part of a set of simultaneous, nonlinear, equations. The solution of this system of equations is usually carried out by a computer program in an iterative fashion, continuing until some preset convergence criteria are met. This is not an exact solution, and a given set of values for the exogenous and lagged endogenous variables can give rise to an indefinitely large number of different solutions, depending on the initial values from which the iterative procedure begins. Much of this imprecision is hidden from view, as only a few significant digits are usually printed on the computer output, but this degree of inaccuracy in the solution of the model equations limits the degree of accuracy of the control calculations. Although one can tighten the convergence criteria used in the model solution program, this is costly in terms of increased solution time, and it will probably result in the inference of a degree of accuracy which is entirely spurious. If the model builders believe their results to three decimals and choose the convergence criteria accordingly, tightening these criteria will apparently yield more exact answers, but these answers can contain no more information than before.

With these cautionary comments as preface, I turn to the problem of actually selecting an optimal path for a set of macroeconomic policy variables with the aid of a nonlinear model. Among a variety of approaches that could be taken, perhaps the most aesthetically pleasing would be simply to treat the problem in a straightforward control theory manner: set up the Hamiltonian, solve the necessary conditions, and obtain an exact analytic solution. This would be enormously difficult to carry out for most of the macro models now in use, however, and it also has the major disadvantage that the entire process would have to be carried out separately for each model; there would be no standardized procedure which could be applied easily to all models.[4]

A dynamic programming framework also suggests itself, but in any realistic context, with several instruments to be used to control several targets over several time periods, the "curse of dimensionality" is likely to result in too massive a demand for computer storage space.

In response to these problems, the tendency has been to utilize some kind of linear approximation rather than the true model, and thus to avoid the difficulties associated with nonlinearity. This is the approach which I have taken, and which I will describe in the next part of the paper.

III

The optimization procedure I will describe has several advantages, not the least of which is that it is relatively inexpensive to use with any model already prepared for computerized simulation experiments. With the exception of one series of matrix manipulations, all necessary calculations can be performed by the

[4] An additional drawback is that an apparently minor change in the structure of the model could require that the entire problem be re-analyzed.

existing solution program. Another advantage is that the extra storage requirement for the optimization procedure depends only on the number of variables in the objective function, and not on the size of the model. And finally, although the true model is approximated by a set of linear relations, this is not a once-for-all linearization, but is repeated at each step in the iterative process.

At this stage all problems associated with uncertainty and the presence of stochastic elements will be ignored. I assume the existence of a known loss function which is quadratic in certain target and instrument variables, and of a known, nonstochastic (and in general, nonlinear) relationship between the instruments and the targets. These assumptions will now be spelled out in greater detail.

Targets will ordinarily be related to concepts such as unemployment, inflation, growth, balance of payments, etc. For example, the unemployment target might be defined as the unemployment rate, or as the number of unemployed workers, and it could be defined in terms of the entire work force or of some sub-category (or several unemployment concepts might be used as separate targets). Any of these alternatives can easily be handled, provided only that two requirements are met. The first is that the target variables so defined (or functions of them) must be appropriate for inclusion in a quadratic loss function, and the second is that their values must be generated by the model.

The second of these requirements is trivial, since the presence in the loss function of a variable which is not at the same time in the model would have no operational significance. The first is not trivial, as one of the difficulties with a quadratic loss function is that it treats deviations of a variable from its target value as equally undesirable, regardless of sign. While this may be quite appropriate for some target variables, for others positive deviations may be viewed quite differently from negative ones. To deal with this problem, simply devise a function of the variable in question which will be smooth and will at the same time capture the essential characteristics of the policy-maker's attitude toward the original variable's behavior. For example, suppose that we dislike values of unemployment greater than 4 percent but are nearly indifferent to values less than 4 percent. In this case we can create a new variable (Y) related to unemployment (U) such that Y is zero when U is 4 percent, dY/dU is positive and large when U is greater than four percent, and dY/dU is positive but small when U is less than four percent. Such a relationship could be approximated by a single function or by pieces of several functions, so long as care is taken to avoid sharp corners where the functions join.[5] With some care and ingenuity, this technique could be used to fit most target variables into a quadratic function.[6] It is implicit in the above example, and I will continue to assume, without loss of generality, that each target variable is defined such that its "desired" value is zero.[7]

Much of what was just said about target variables will be equally true of instrument variables. Each instrument must be a variable whose value can reasonably

[5] Such an approach is similar to the use of a "penalty function" as a means of turning a constrained into an unconstrained minimization problem [10].

[6] Friedman [3] suggests that the problem should be solved by the use of an objective function which is piecewise quadratic, but I believe that the procedure I have described will be equally effective, and more easily implemented. In the event that we must use a non-quadratic objective function, an extension of my optimizing method described in [5] could probably be used.

[7] This assumption will usually eliminate the necessity for linear terms in the loss function.

be assumed to be chosen by the policy-maker. It must appear in the model as an exogenous variable, and be defined so that it makes sense when included as an argument of a quadratic loss function with a desired value of zero. The incorporation of policy instruments in the loss function has been viewed as a questionable device [11], but it seems clear from casual observation that policy-makers are not indifferent to the values they select for their policy instruments. Their attitude may reflect political considerations, uncertainty about the future, or merely a desire for the quiet life, but whatever the reason, I do not believe that reality is violated when we include policy instruments in the loss function.

As an example of a policy instrument, consider government purchases of goods and services. Clearly, it is nonsense to assume that the government would be willing to spend at whatever level is dictated by simple macroeconomic considerations. Many of the other goals that government spending is designed to achieve (most of them not even represented in the typical macroeconomic model) are likely to be poorly served if spending is determined only on the basis of macroeconomic goals. It is more sensible to assume that there is some desired level (or rate of growth) of government expenditure, and that progressive deviations from this desired value (in either direction) are viewed as increasingly undesirable. And, of course, this is precisely the attitude captured when we include the appropriate function of government spending as an argument in the loss function.[8]

I can now write the loss function as

$$L = (Y' \mid X')H\left(\frac{Y}{X}\right) \tag{1}$$

where Y and X are column vectors of targets and instruments, respectively, as described above, and H is a symmetric matrix of coefficients.[9] Although the only operational information emerging from the optimization procedure will have to do with the current period, a policy-maker would probably wish to take account of the path of the economy for some time into the future; if that is true the Y and X vectors must contain both current and future values of the target and instrument variables, for as many periods as are necessary. If there are n targets and m instruments, and the planning horizon is T periods, then

$$Y = \begin{bmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{n1} \\ y_{12} \\ \vdots \\ y_{nT} \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \\ x_{12} \\ \vdots \\ x_{mT} \end{bmatrix}$$

[8] More complex assumptions regarding the policy-maker's view of the behavior of his instrument variable can be easily handled by the device described earlier in the discussion of an unemployment rate target.

[9] As noted earlier, it will usually be unnecessary to employ linear terms in the loss function, and I will omit them here in the interest of simplicity. See [6] for a derivation which includes the linear terms.

where the lower case letters refer to the individual target or instrument values. The matrix H is then $(n + m)T$ by $(n + m)T$, and must be positive definite (implying that loss takes on its minimum value of zero when, and only when, Y and X are both zero vectors).

The model describing the behavior of the economic system is typically both large and nonlinear, containing many non-target endogenous variables and non-instrument exogenous variables. In fact, it may be that none of the original variables in the model precisely fit the definitions chosen for target and instrument variables, in which case we would have to augment the model by adding the necessary equations, as described in Section II.

This (augmented) model can be written as

$$F(Y, Y_1, X, X_1) = 0 \tag{2}$$

where Y and X are as before, Y_1 is a vector of the non-target endogenous variables, and X_1 is a vector of the non-instrument exogenous variables. The vectors Y and X contain all the arguments of the loss function, so if T (tne number of periods in the planning horizon) is greater than one, F is not simply the set of equations in our econometric model, but rather is T sets of these equations, one for the current period, and one for each of the $T - 1$ future periods over the planning horizon, with appropriate time subscripts.

The variables in Y_1 can be ignored, as their values are of no consequence to the policy-maker. He may wish to know about their behavior along the optimal path, but, by assumption, there is no feedback from that behavior affecting the choice of optimal policy. I also drop explicit mention of X_1 since, although the values of its elements much be chosen (or predicted) by some means or other, once they have been chosen they can be taken as parameters rather than variables with respect to the optimization problem. In principle, then, the system can be simplified to

$$Y = G(X). \tag{3}$$

While it may be difficult or even impossible actually to write out the equations explicitly in this way, the typical econometric model has an associated computer program which can readily provide numerical solutions, and this is all that is necessary.

The problem is simply to select that value of X which will minimize L, subject to the constraint imposed by the relationship in (3). This can be done as follows:

Let X^*, Y^*, and L^* be some initial mutually consistent values such that

$$L^* = (Y^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} Y^* \\ \hdashline X^* \end{pmatrix}$$

and

$$Y^* = G(X^*).$$

Then define ΔL, ΔY, and ΔX such that

$$L^* + \Delta L = (Y^{*\prime} + \Delta Y' \,\vdots\, X^{*\prime} + \Delta X')H\begin{pmatrix} Y^* + \Delta Y \\ \hdashline X^* + \Delta X \end{pmatrix}$$

and

$$Y^* + \Delta Y = G(X^* + \Delta X) = G(X^*) + K(X^*, \Delta X).$$

Even if G is a nonlinear function, it will ordinarily be safe to assume that, for sufficiently small ΔX, K is reasonably linear in ΔX. We thus approximate K by the nT by mT matrix U, such that

$$\Delta Y = K(X^*, \Delta X) = U\Delta X$$

where each element of U is an approximation to the partial derivative of a particular y with respect to an x, evaluated at $X = X^*$. Then

$$\frac{\partial(L^* + \Delta L)}{\partial \Delta X} = 2\left[\left(\frac{\partial K(X^*, \Delta X)}{\partial \Delta X}\right)' \vdots\; I\right] H \left(\frac{Y^* + K(X^*, \Delta X)}{X^* + \Delta X}\right)$$

$$0 = (U' \,\vdots\, I)H\left(\frac{Y^*}{X^*}\right) + (U' \,\vdots\, I)H\left(\frac{U}{I}\right)\Delta\hat{X}$$

where I is an mT by mT identity matrix. Then

$$\Delta\hat{X} = -\left[(U' \,\vdots\, I)H\left(\frac{U}{I}\right)\right]^{-1}(U' \,\vdots\, I)H\left(\frac{Y^*}{X^*}\right) \qquad (4)$$

where $\Delta\hat{X}$ is the loss minimizing value of ΔX, given Y^*, X^*, and U, and is an approximation to the loss minimizing value of ΔX, given Y^*, X^*, and G. This will be valid only to the extent that the approximation of K by $U\Delta X$ is valid, but if $\Delta\hat{X}$ is very small, the error in the approximation is likely also to be small and the solution will be almost correct.

The matrix U can most easily be estimated by a series of mT simulations of the model over the planning horizon. In each simulation all elements of X but one are set equal to the values in X^*, and that one (say x_{it}) differs from its value in X^* by a small amount (Δx_{it}). The values of the elements of Y in this perturbed simulation will differ slightly from those in Y^*, and it is these differences (each divided by the size of the perturbation that caused them) that are used as the elements of U. When x_{it} is perturbed, a vector of results is obtained of the form

$$\begin{bmatrix} \dfrac{y_{11} - y_{11}^*}{\Delta x_{it}} \\ \dfrac{y_{21} - y_{21}^*}{\Delta x_{it}} \\ \vdots \\ \dfrac{y_{n1} - y_{n1}^*}{\Delta x_{it}} \\ \dfrac{y_{12} - y_{12}^*}{\Delta x_{it}} \\ \vdots \\ \dfrac{y_{nT} - y_{nT}^*}{\Delta x_{it}} \end{bmatrix}$$

and this vector will be used as the $[n(t - 1) + i]$-th column of U.

The initial value of X^* is likely to be rather far from its optimal value, so the solution of (4) will yield large values for the elements of $\Delta\hat{X}$. If G is not linear, these will not be the truly optimal changes in X^*, but it is simple to make the indicated changes, solve for new values of Y, both control and perturbed, and solve (4) once again. My experiments indicate that this iterative procedure converges rapidly even when the model is rather large and quite nonlinear.

Before presenting some recent results obtained with the use of this technique, I will discuss its relation to one of the standard gradient methods of function minimization. As described by Polak [10], the Newton–Raphson iteration procedure can be written as

$$X_{i+1} - X_i = -\left(\frac{\partial^2 L(X_i)}{\partial X^2}\right)^{-1} \nabla L(X_i) \tag{5}$$

where X_i is a p-vector of values of the independent variables as of the i-th iteration, and $\nabla L(X_i)$ is the gradient of L at X_i, written as a column vector.

Using (3), I can rewrite (1), evaluated at X_i, as

$$L(X_i) = (G(X_i)' \mid X_i')H\left(\frac{G(X_i)}{X_i}\right). \tag{6}$$

The first derivative of L with respect to X, written as a column vector (i.e., the gradient of L) is then

$$\frac{\partial L(X_i)}{\partial X} = \nabla L(X_i) = 2\left(\frac{\partial G(X_i)'}{\partial X} \,\middle|\, I\right)H\left(\frac{G(X_i)}{X_i}\right) \tag{7}$$

and the second derivative is

$$\frac{\partial^2 L(X_i)}{\partial X^2} = 2\left[\left(\frac{\partial G(X_i)}{\partial X}\,\middle|\, I\right) H \left(\frac{\;\frac{\partial G(X_i)}{\partial X}\;}{I}\right) + D\right] \tag{8}$$

where

$$D = \left[\frac{\partial\left(\frac{\partial G(X_i)}{\partial X}\right)'}{\partial x_1} \,\middle|\, 0 \,\middle|\, \frac{\partial\left(\frac{\partial G(X_i)}{\partial X}\right)'}{\partial x_2} \,\middle|\, 0 \,\middle|\, \cdots \,\middle|\, \frac{\partial\left(\frac{\partial G(X_i)}{\partial X}\right)'}{\partial x_p} \,\middle|\, 0\right] \tag{9}$$

$$\times \left[\begin{array}{c|c|c|c} H & 0 & \cdots & 0 \\ \hline 0 & H & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline 0 & 0 & \cdots & H \end{array}\right] \left[\begin{array}{c|c|c|c} G(X_i) & 0 & \cdots & 0 \\ \hline X_i & 0 & \cdots & 0 \\ \hline 0 & G(X_i) & \cdots & 0 \\ \hline 0 & X_i & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline 0 & 0 & \cdots & G(X_i) \\ \hline 0 & 0 & \cdots & X_i \end{array}\right]$$

and x_j is the j-th element of X. Using (7) and (8), I can now rewrite (5) as

$$(X^i)$$

$$(10)\quad X_{i+1} - X_i = -\left[\left(\frac{\partial G(X_i)'}{\partial X} \;\vdots\; I\right) H \begin{pmatrix} \dfrac{\partial G(X_i)}{\partial X} \\ \hdashline I \end{pmatrix} + D\right]^{-1} \left(\frac{\partial G(X_i)'}{\partial X} \;\vdots\; I\right) H \begin{pmatrix} G(X_i) \\ \hdashline X_i \end{pmatrix}.$$

This equation is exactly equivalent to (4), except for the second derivative term, D, which has no counterpart in the earlier equation. My optimizing procedure is based on the assumption that G is nearly linear within a small neighborhood, and thus that D can safely be set equal to zero.

In fact, D is ordinarily not exactly zero, and omitting it will have an effect on the size of the step $(X_{i+1} - X_i)$, and thus on the convergence properties of the method. It remains true, however, that if the sequence generated by the repeated application of (4) converges, the gradient at the point of convergence must be zero (this follows from the fact that H is assumed to be positive definite). I must leave unanswered the question of the conditions under which the procedure necessarily converges to a solution. I can only report that it has converged to within a small neighborhood very quickly for two large, nonlinear models; whether it will do so for others remains to be determined.

There are several other general points that should be made regarding the use of this optimization method on a large model of the standard type:

1. Assuming that the procedure converges, it can at best find only a local minimum for the loss function.

The only ways I know to be sure of finding the global minimum are to search over a very fine grid, or to solve the problem analytically. I am comforted, however, by the fact that in the many runs I have made with two different models there has never been the slightest hint of multiple minima.

2. Although the model is nonlinear, we approximate it by a linear function and ignore the second derivatives. Using that linear approximation, we could satisfy all the requirements for minimizing the loss function, and yet actually be at a maximum for the true model.[10]

This is not a serious problem, for two reasons. First, as stated earlier, the solution programs do not solve the model exactly. Thus, the optimizing procedure cannot converge to a point, but only to a neighborhood (the size of which will be discussed in point 3, below), and once it has reached that neighborhood the solution will tend simply to wander about, moving in first one direction, then another, but always being pushed back toward the loss minimizing point. If we were really at a maximum, a point of inflection, or a saddle point, such wandering would almost certainly discover it, and the procedure would immediately move us off in the appropriate direction. A second factor reducing the probability that this problem could arise is that we are not taking derivatives of the model at a point, but finite differences, and these will also tend to move the solution around enough so that the fact that we had reached a maximum rather than a minimum would quickly become apparent.

3. As just noted, when the solution arrives in the neighborhood of the optimal value, the nonzero convergence criteria used in the solution of the model

[10] Richard Kopcke was the first to point out this possibility to me.

and the finite perturbations used in estimating the U matrix prevent us from ever reaching an exact solution.

In principle, we could reach any desired degree of accuracy simply by reducing both the convergence criteria and the size of the perturbations as we approach minimum loss, but there are limits to this, as mentioned in Section II. First, the cost of solving the model goes up rapidly. Second, the precision of the computer programs is limited. And third, in most models the effects of changes in policy instruments on certain targets is likely to be quite small, so the smallest perturbation we can use may have to be 1,000 or 10,000 times the size of the convergence criterion, in order to ensure that we obtain sufficient significant digits to produce a dependable answer. Given the level of accuracy of the econometric model and of the forecasts of exogenous variables, however, an attempt greatly to increase the accuracy of the optimization procedure beyond the level now easily attained cannot be justified on practical grounds.

Some of these problems are illustrated in the next section of the paper, where I discuss a recent application of the method to the Michigan Model.

IV

I first applied my optimization procedure just described to RDX2, the Canadian model mentioned earlier. I have just begun to experiment with the Michigan Model, but it will be useful to examine some of the preliminary results. These are of interest not so much as policy prescriptions (the instruments, targets, and loss function coefficients are all chosen rather arbitrarily) but because they illustrate both the ease with which the technique can yield a solution, and the limitations caused by some of the problems described in Section II.

The Michigan Model consists of 61 equations, of which about 45 form a set of simultaneous relationships while the remainder are recursive. The simultaneous block is solved by the Gauss–Seidel method, with a convergence criterion for all variables of 0.001 (i.e., one-tenth of one percent).

I chose three target variables: unemployment, inflation, and the trade balance; and three instrument variables: federal spending, personal income tax rates, and the reserve base. The precise definition of each of these variables follows:

In the model, UG % is the global unemployment rate, in percent. I assume that the desired value of this target is zero, so the first target is defined simply as the global unemployment rate, in percent,

$$Y_{1t} = \text{UNEMP}_t = \text{UG}\%_{0t}. \tag{11}$$

This variable is referred to as Y_{1t} when it is utilized in the optimizing equation (4), but in the text and tables I will use the mnemonic UNEMP_t. In this case there is an apparently needless proliferation of variable names, but it is desirable to differentiate between those variables already defined in the model and the special variables defined for the purpose of policy selection.

The model defines PGNP as the gross national product implicit deflator, but does not provide an explicit rate of inflation. I take the desired rate of inflation to be zero, and define the second target as the annual percentage rate of change in the GNP deflator,

$$Y_{2t} = \text{INFL}_t = \left[\frac{\text{PGNP}_t}{\text{PGNP}_{t-1}} - 1.0\right]400.0. \tag{12}$$

I assume that it is the nominal surplus or deficit in the balance of trade which is of concern to the policy-maker, and that he prefers a zero balance. This assumption is subject to objection, and can easily be modified in future experiments. The model defines exports in billions of current dollars (X\$), imports in billions of 1958 dollars (M), and the import implicit deflator (PM). The third target is then simply the difference between exports and imports in billions of current dollars,

$$Y_{3t} = \text{BOFT}_t = \text{X\$}_t - (\text{M}_t)\frac{\text{PM}_t}{100.0}. \tag{13}$$

Federal nondefense purchases of goods and services in current dollars (GFO\$) is an exogenous variable in the model, and the implicit deflator for these purchases (PG) is an endogenous variable. I assume that, in the absence of stabilization problems, the desired level of federal purchases would grow, in real terms, about as fast as the economy (say, 0.9 percent per quarter). The policy-maker is assumed to inflate constant dollar purchases by the price deflator in the preceding quarter, because of the simultaneity problem (i.e., he doesn't know what PG will be next quarter). Given an acceptable figure for these expenditures in period t_0, it is easy to calculate their "desired" value in period t as

$$\text{“Desired”}\left(\frac{\text{GFO\$}_t}{\text{PG}_{t-1}}\right) = (1.009)^{t-t_0}\left(\frac{\text{GFO}_{t_0}}{\text{PG}_{t_0}}\right). \tag{14}$$

The expenditure policy instrument is the difference between actual and desired expenditure, expressed as a percentage deviation at an annual rate,

$$X_{1t} = \text{GEXP}_t = \left[\frac{\dfrac{\text{GFO\$}_t}{\text{PG}_{t-1}}}{(1.009)^{t-t_0}\left(\dfrac{\text{GFO\$}_{t_0}}{\text{PG}_{t_0}}\right)} - 1.0\right]400.0. \tag{15}$$

Since expenditure has now become an endogenous variable, it must have a defining equation,

$$\text{GFO\$}_t = \left[\frac{\text{GEXP}_t}{400.0} + 1.0\right](1.009)^{t-t_0}\text{GFO\$}_{t_0}\left(\frac{\text{PG}_{t-1}}{\text{PG}_{t_0}}\right). \tag{16}$$

The tax instrument is a one-time percentage surcharge (positive or negative) applied to the average tax rate on personal incomes. Given total personal income tax liability in current dollars (TP\$) in some initial period, t_0, the model defines the difference between tax liability in that period and in some later period t as

$$\text{TP\$}_t - \text{TP\$}_{t_0} = 0.20[(\text{YP\$}_t - \text{GTRP\$}_t) - (\text{YP\$}_{t_0} - \text{GTRP\$}_{t_0})] \tag{17}$$

where YP\$ is personal income and GTRP\$ is government transfer payments to persons, both in current dollars. The tax instrument TAX_t, or X_{2t}, is then the one period surcharge as a percent of the total tax liability calculated on the basis of (17), so the actual tax liability in any period, including the surcharge, is

$$\text{TP\$}_t = \left(1.0 + \frac{\text{TAX}_t}{100.0}\right)(\text{TP\$}_{t_0} + 0.20[(\text{YP\$}_t - \text{GTRP}_t) - (\text{YP\$}_{t_0} - \text{GTRP\$}_{t_0})]).$$

(18)

The monetary instrument is defined as the difference between the actual annual rate of growth in unborrowed reserves in current dollars (UR\$) and 6 percent,

$$\text{RES}_t = X_{3t} = \left(\frac{\text{UR\$}_t}{\text{UR\$}_{t-1}} - 1.0\right)400.0 - 6.0 \tag{19}$$

so actual unborrowed reserves in any period are

$$\text{UR\$}_t = \left[\frac{\text{RES}_t + 6.0}{400.0} + 1.0\right]\text{UR\$}_{t-1}. \tag{20}$$

I have made three optimizing experiments with the Michigan Model using these definitions for the targets and instruments. The results are presented in Table 1. In each case the loss function coefficients for all targets in all periods were assigned the same value (100.0). This assumption is made not for its realism but because it simplifies the presentation of and comparisons among the outcomes. The loss function coefficients for the instruments are also the same in all periods, but differ between instruments and between experiments, as shown in the table. The same six quarter planning period (1968.3 to 1969.4) was used for all experiments, and all non-instrument exogenous variables were assigned their actual historical values.

Part A of Table 1 shows the paths of the target variables when the instruments are set at zero for each period.[11] The loss function coefficient for each target in each period is shown to be 100.0. The loss associated with each target is shown separately

TABLE 1

A Initial Path

Quarter	UNEMP	INFL	BOFT	GEXP	TAX	RES	Loss
1968: 3	3.52	4.82	5.21	0.0	0.0	0.0	6,276.69
4	3.66	4.80	0.97	0.0	0.0	0.0	3,737.65
1969: 1	3.84	4.64	1.50	0.0	0.0	0.0	3,852.52
2	4.00	4.97	1.73	0.0	0.0	0.0	4,369.38
3	4.13	5.02	2.18	0.0	0.0	0.0	4,700.97
4	4.27	4.23	2.48	0.0	0.0	0.0	4,227.62
Loss coefficient	100	100	100	—	—	—	
Loss	9,182.14	13,559.62	4,423.07				27,164.83

[11] Setting an instrument at zero implies that the underlying policy variable is following what I have assumed to be its "desired" path, as described in equations (16), (18), and (20). It should not be confused with a zero value for the underlying policy variable (e.g., government spending) or with its historical path.

TABLE 1—continued

B Experiment 1

Quarter	UNEMP	INFL	BOFT	GEXP	TAX	RES	Loss
1968: 3	3.52	4.80	5.17	−0.11	−4.41	−0.71	6,315.75
4	3.63	4.76	0.83	0.37	−4.79	−0.35	3,768.36
1969: 1	3.75	4.59	1.23	0.43	−6.04	0.04	3,847.69
2	3.82	4.93	1.31	0.59	−6.35	0.00	4,264.69
3	3.87	4.99	1.64	0.47	−5.33	0.01	4,399.81
4	3.96	4.40	1.89	0.02	−2.20	−0.01	3,885.57
Loss coefficient	100	100	100	5	5	5	
Loss	8,488.07	13,533.07	3,690.85	4.52	762.23	3.14	26,481.88
% decline in loss from initial path	7.6%	0.2%	16.6%				2.5%

C Experiment 2

Quarter	UNEMP	INFL	BOFT	GEXP	TAX	RES	Loss
1968: 3	3.52	4.80	5.17	0.86	−4.15	−4.06	6,310.65
4	3.63	4.75	0.84	2.55	−4.46	−2.21	3,757.84
1969: 1	3.75	4.58	1.23	4.70	−5.91	−2.00	3,842.87
2	3.81	4.93	1.30	5.56	−6.28	−0.06	4,263.75
3	3.85	4.99	1.62	5.62	−5.30	0.56	4,391.10
4	3.95	4.40	1.87	0.65	−2.06	0.16	3,867.38
Loss coefficient	100	100	100	0.5	5	0.5	
Loss	8,457.09	13,514.39	3,675.87	46.13	727.26	12.86	26,433.59
% decline in loss from initial path	7.9%	0.3%	16.9%				2.7%

D Experiment 3

Quarter	UNEMP	INFL	BOFT	GEXP	TAX	RES	Loss
1968: 3	3.48	4.78	5.10	14.08	−7.88	−10.94	6,143.82
4	3.55	4.66	0.66	4.49	−6.01	7.16	3,497.00
1969: 1	3.57	4.57	0.88	22.71	−15.84	8.68	3,595.43
2	3.46	4.82	0.56	24.62	−18.88	10.12	3,765.42
3	3.39	4.91	0.40	22.54	−23.79	−2.95	3,884.84
4	3.40	4.63	0.35	−6.45	−11.13	0.23	3,375.96
Loss coefficient	100	100	100	0.05	0.5	0.05	
Loss	7,248.15	13,422.63	2,781.61	94.50	697.71	17.87	24,262.47
% decline in loss from initial path	21.1%	1.0%	37.1%				10.7%

at the bottom of each column, and the combined loss associated with each period is shown at the right end of each row.

Part B displays the results obtained when each instrument is assigned a loss function coefficient of 5.0. Total loss has declined about 2.5 percent, but this improvement is not reflected equally in all targets. The rate of inflation has hardly been affected at all, but the increase in unemployment has been significantly reduced, and the average trade balance has been reduced by about a third of a billion dollars per quarter. Examination of the trends in the variables will reveal that the improvement is small at the beginning, but tends to increase later in the planning period, reflecting the presence of lags in the effects of the policy instruments. In fact, we get almost no improvement at all in the initial period, and it isn't until the third period that the combined loss (shown in the last column) begins to decline.

The small values of the expenditure and monetary instruments in the first experiment led me to reduce their loss coefficients in the second experiment, as shown in Part C. The overall results are not substantially different from those in Part B, except that GEXP and RES are used more vigorously than before.

Part D shows the optimal result when the weights on the instruments are reduced to one tenth their size in Part C. Here we see a major improvement in the unemployment situation, together with some measurable reduction of inflation.[12] The average trade balance has now been reduced by about a billion dollars per quarter. All these improvements have been achieved through substantial increases in federal spending, reductions in income taxes, and rather wild gyrations in the reserve base, but the behavior of total loss indicates that it was worthwhile. Total loss over the six periods has declined by almost 11 percent from the initial path, and the reduction of loss in the fourth quarter of 1969 alone is more than 20 percent.

Experiment 1 required six iterations to achieve a minimum loss configuration, experiment 2 required 14, and experiment 3 required 10. As described earlier, these were not exact solutions, and continued iteration might eventually have produced slightly lower values of loss. Any further improvement, however, would almost certainly come from a rearrangement of the instrument values, rather than from any reduction in the loss due to the target values. In fact, the optimization procedure usually requires only four or five iterations to achieve values of the target variables that are within a very narrow range (typically ± 0.02) of the ultimate optimal values. Further iteration only explores improved ways of arranging the instruments in order to achieve those target values.

The optimal solutions in Table 1 were derived with a convergence criterion of 0.0001, one tenth of the value usually used for solving the Michigan Model. I tried even smaller values, but the number of iterations necessary to solve the model increased so fast that I decided to forego the higher degree of precision.

In most cases the smallest perturbation I used was 0.5. The reason for this should be apparent from an examination of Table 2, which shows the effects on the target variables of a perturbation of 1.0 for each of the instrument variables in 1968.3. Many of the effects are so small as to be quite unreliable, even with the new convergence criterion, but the optimization routine takes them all to be equally significant, and produces an optimal path based upon them. It is not surprising, then, that the resulting path fails to be quite optimal, given the faulty information from which it was derived.[13]

[12] Except in the fourth quarter of 1969, the third experiment managed to reduce both unemployment and the rate of inflation in every period (in comparison with their initial values). There is obviously a limit to this, and we may be running into it just at the end of the planning period. If this is true we would probably wish to lengthen the planning horizon in order to temper somewhat our near-term policy action.

[13] The data in Table 2 also explain why the tax variable is used so vigorously in the first experiment; it is far more effective (per unit) than either of the other two instruments.

TABLE 2

EFFECTS OF UNIT PERTURBATIONS ON TARGET VARIABLES*

	A Perturbation of 1.0 for GEXP in 1968.3					
	1968 3	1968 4	1969 1	1969 2	1969 3	1969 4
UNEMP	−0.00292	−0.00335	−0.00334	−0.00267	−0.00196	−0.00147
INFL	0.0	−0.00267	0.00343	0.00343	0.00305	0.00229
BOFT	−0.00145	−0.00328	−0.00133	−0.00121	−0.00148	−0.00125
	B Perturbation of 1.0 for TAX in 1968.3					
	1968 3	1968 4	1969 1	1969 2	1969 3	1969 4
UNEMP	0.00416	0.01205	0.01338	0.01073	0.00740	0.00402
INFL	0.00648	0.01183	−0.01411	−0.01717	−0.01564	−0.01907
BOFT	0.01508	0.03140	0.02769	0.01775	0.01593	0.01640
	C Perturbation of 1.0 for RES in 1968.3					
	1968 3	1968 4	1969 1	1969 2	1969 3	1969 4
UNEMP	−0.00199	−0.00179	−0.00203	−0.00176	−0.00123	−0.00078
INFL	−0.00076	0.0	0.00191	0.00153	0.00114	0.00114
BOFT	−0.00050	−0.00050	0.00055	0.00090	0.00133	0.00191

* The figures shown are the differences between the perturbed path and the initial path as given in Part A of Table 1.

Many other experiments are suggested by the results just described, and they will be undertaken as part of the continuing research effort in the Research Seminar in Quantitative Economics at the University of Michigan.

V

In Sections III and IV I assumed the existence both of a fully deterministic model and of errorless forecasts of all exogenous variables over the planning horizon. These assumptions will now be replaced by more reasonable ones regarding the nature of the model and of a policy-maker's ability to forecast the future.

Many of the non-instrument exogenous variables are actually random variables with means, variances, and covariances which would have to be estimated from information outside the model (probably from the forecaster's subjective views regarding the reliability of his data and projections). The estimated coefficients in the equations of the model are also random variables, and each fitted equation has an additive error term as well. For these reasons I now assume a stochastic relationship between the selected values for the policy instruments and the resulting values of the target variables.

Under these modified assumptions, a particular set of values for the instrument variables (X^*) will result in an outcome (Y^*) which is the sum of its expected value (W^*) and an error term (ε) with an expected value of zero,

$$Y^* = G(X^*) = E(Y^*) + \varepsilon = W^* + \varepsilon \tag{21}$$

$$E(\varepsilon) = 0.$$

At the same time, the matrix of effects from a set of perturbations (given X^*) is also a random variable (U) which is the sum of its expected value (V) and an error term (Q) with an expected value of zero,

$$\Delta Y = U\Delta X = E(U)\Delta X + Q\Delta X = V\Delta X + Q\Delta X \qquad (22)$$
$$E(Q) = 0.$$

Estimates of W^*, V, and the variances and covariances of the elements of ε and Q will be obtained by means of stochastic simulation. A complete set of control and perturbation runs will be made for each set of values assigned to the stochastic elements in the model. After a suitable number of such runs the indicated calculations can easily be performed.

Given the successful completion of these stochastic simulations, two problems will remain. First, what if the estimated values of W^* and V are not equal to the deterministic values of Y^* and U we used in Section III? And second, how should we incorporate our knowledge of ε and Q into our policy solution?

Regarding the first of these two issues, Howrey and Kelejian [7] have shown that in a nonlinear system the values of the endogenous variables obtained by solving the model with all stochastic variables set equal to their expected values (as we did in Sections III and IV) will not in general be equal to the expected values of those endogenous variables. Thus, we would expect that the value of W^* calculated as the mean outcome from the stochastic simulations would differ systematically from the value of Y^* we employed before in the deterministic version (the same can be said for V and U). Although annoying, this problem presents no serious difficulty, as we can simply replace our earlier values of Y^* and U with W^* and V, and proceed.[14]

The question regarding the use to be made of ε and Q is the primary subject of this section. We proceed very much as in Section III, but this time we minimize *expected* loss:

$$E(L^* + \Delta L) = E\left[(Y^{*\prime} + \Delta Y' \,\vdots\, X^{*\prime} + \Delta X')H\begin{pmatrix} Y^* + \Delta Y \\ \hdashline X^* + \Delta X \end{pmatrix}\right].$$

This can be rewritten as

$$E(L^* + \Delta L) = E\Bigg[(W^{*\prime} + \varepsilon' + \Delta X'V'$$
$$+ \Delta X'Q' \,\vdots\, X^{*\prime} + \Delta X')H\begin{pmatrix} W^* + \varepsilon + V\Delta X + Q\Delta X \\ \hdashline X^* + \Delta X \end{pmatrix}\Bigg].$$

[14] Fair [1] also suggests the use of stochastic simulation as a device for estimating the mean value of W^* in a control context. Muench *et al.* [9] made some stochastic simulation experiments using an earlier version of the Michigan Model. They found that, while there were systematic differences between the point estimates of Y based on the deterministic model and the mean outcome of the stochastic simulations, the differences in most cases were quite small. It is yet to be seen whether this remains true for the current version of the model.

Now expand, dropping those terms with an expected value equal to zero.

$$E(L^* + \Delta L) = (W^{*\prime} \vdots X^{*\prime})H\left(\frac{W^*}{X^*}\right) + (W^{*\prime} \vdots X^{*\prime})H\left(\frac{V}{I}\right)\Delta X$$
$$+ \Delta X'(V' \vdots I)H\left(\frac{W^*}{X^*}\right) + \Delta X'(V' \vdots I)H\left(\frac{V}{I}\right)\Delta X$$
$$+ E\left[(\varepsilon' \vdots 0)H\left(\frac{\varepsilon}{0}\right) + (\varepsilon' \vdots 0)H\left(\frac{Q}{0}\right)\Delta X\right.$$
$$\left. + \Delta X'(Q' \vdots 0)H\left(\frac{\varepsilon}{0}\right) + \Delta X'(Q' \vdots 0)H\left(\frac{Q}{0}\right)\Delta X\right].$$

Make the following definitions:

$$\phi = E\left[(\varepsilon' \vdots 0)H\left(\frac{\varepsilon}{0}\right)\right]$$
$$\theta = E\left[(Q' \vdots 0)H\left(\frac{\varepsilon}{0}\right)\right]$$
$$\Sigma = E\left[(Q' \vdots 0)H\left(\frac{Q}{0}\right)\right].$$

Now we can solve for $\Delta\hat{X}$, the value of ΔX which will minimize the expected loss:

$$\frac{d(E(L^* + \Delta L))}{d\Delta X} = 2(V' \vdots I)H\left(\frac{W^*}{X^*}\right) + 2(V' \vdots I)H\left(\frac{V}{I}\right)\Delta X + 2\theta + 2\Sigma\Delta X$$

$$\Delta\hat{X} = -\left[(V' \vdots I)H\left(\frac{V}{I}\right) + \Sigma\right]^{-1}\left[(V' \vdots I)H\left(\frac{W^*}{X^*}\right) + \theta\right]. \tag{23}$$

The solution of this matrix equation is readily obtainable, given the results of the stochastic simulations already described. The two new factors, Σ and θ, are composed of elements each of which is a weighted sum of variances and covariances of some components of Q or of Q and ε. For example, the element in the ith row and jth column of Σ is equal to

$$\sum_{g=1}^{nT}\sum_{k=1}^{nT} h_{gk} \operatorname{cov} q_{gi}q_{kj}$$

where h_{gk} is an element of H and q_{gi} is an element of Q.

The effect of including Σ and θ can more easily be seen if (23) is rewritten as

$$\begin{aligned}\Delta\hat{X} = &-\left[I + \left[(V' \vdots I)H\left(\frac{V}{I}\right)\right]^{-1}\Sigma\right]^{-1}\left[(V' \vdots I)H\left(\frac{V}{I}\right)\right]^{-1}(V' \vdots I)H\left(\frac{W^*}{X^*}\right)\\ &-\left[I + \left[(V' \vdots I)H\left(\frac{V}{I}\right)\right]^{-1}\Sigma\right]^{-1}\left[(V' \vdots I)H\left(\frac{V}{I}\right)\right]^{-1}\theta.\end{aligned} \tag{24}$$

It can be observed that the first term in (24) is simply the optimizing value of ΔX if Σ and θ were assumed to be zero, multiplied by a factor which is inversely related to Σ, and which becomes the identity matrix if Σ equals zero. The second term in (24) is an additive correction factor whose size depends on the covariation between the elements of Q and those of ε (it also contains the same function of Σ as a multiplicative factor).

With the aid of this form of the optimizing equation, it would be simple to calculate an optimal value for ΔX on the assumption of zero variances and covariances, and then to calculate the effects on that optimal value due to the estimates of Σ and θ. It is possible that the size of the effect is so small relative to our overall degree of confidence in our results that we would then decide to omit the adjustment.[15]

It may be of some interest to examine the difference in loss associated with the alternative approaches suggested in the preceding paragraph.[16] If the calculation of ΔX takes account of Σ and θ as in (23) or (24), the expected value of total loss is

$$(25)\quad E(\hat{L}) = (W^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix} + \phi - \left[(W^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} V \\ \hdashline I \end{pmatrix} + \theta'\right] \times\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix} + \Sigma\right]^{-1}\left[(V' \,\vdots\, I)H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix} + \theta\right].$$

If we ignore Σ and θ (i.e., assume when calculating ΔX that their values are zero), then the value of the expected loss (call it $E(\hat{\hat{L}})$ to distinguish it from the former case) will be

$$(26)\quad E(\hat{\hat{L}}) = (W^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix} + \phi - \left[(W^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} V \\ \hdashline I \end{pmatrix} + \theta'\right] \times\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\right]^{-1}\left[(V' \,\vdots\, I)H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix} + \theta\right]$$
$$+(W^{*\prime} \,\vdots\, X^{*\prime})H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\right]^{-1} \times\Sigma\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\right]^{-1}(V' \,\vdots\, I)H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix}$$
$$+\theta'\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\right]^{-1}\theta.$$

Although one would expect $E(\hat{\hat{L}})$ to be greater than $E(\hat{L})$, that relationship does not emerge easily from a comparison of equations (25) and (26). However, if

[15] In fact, if θ happens to have the value

$$\Sigma\left[(V' \,\vdots\, I)H\begin{pmatrix} V \\ \hdashline I \end{pmatrix}\right]^{-1}(V' \,\vdots\, I)H\begin{pmatrix} W^* \\ \hdashline X^* \end{pmatrix}.$$

Then the correction due to Σ is just offset by that due to θ, and optimal policy will be unaffected by the inclusion of these two terms.

[16] In the following calculations I assume that the correct values of W^* and V are used in both cases.

we subtract (25) from (26) and carry out a great deal of manipulation, the result can be simplified to

$$
\begin{aligned}
(27)\quad E(\hat{\hat{L}}) - E(\hat{L}) = {} & \left[(W^{*\prime} \mid X^{*\prime}) H \begin{pmatrix} V \\ \hline I \end{pmatrix} \;\middle|\; \theta' \right] \\
& \times \left[\left(\begin{array}{c|c} \left[(V' \mid I) H \begin{pmatrix} V \\ \hline I \end{pmatrix} + \Sigma \right]^{-1} & 0 \\ \hline 0 & \left[(V' \mid I) H \begin{pmatrix} V \\ \hline I \end{pmatrix} + \Sigma \right]^{-1} \end{array} \right) \right. \\
& \times \left(\begin{array}{c|c} \Sigma \left[(V' \mid I) H \begin{pmatrix} V \\ \hline I \end{pmatrix} \right]^{-1} & 0 \\ \hline 0 & I \end{array} \right) \\
& \times \left. \left(\begin{array}{c|c} \Sigma \left[(V' \mid I) H \begin{pmatrix} V \\ \hline I \end{pmatrix} \right]^{-1} & -I \\ \hline -\Sigma \left[(V' \mid I) H \begin{pmatrix} V \\ \hline I \end{pmatrix} \right]^{-1} & I \end{array} \right) \right] \\
& \times \left(\begin{array}{c} (V' \mid I) H \begin{pmatrix} W^* \\ \hline X^* \end{pmatrix} \\ \hline \theta \end{array} \right)
\end{aligned}
$$

which is merely a very large quadratic form. As we might have anticipated, the improvement in loss is due to two effects, one of them a function of Σ alone, and the other a function of both Σ and θ. Further examination will reveal that all the matrices within the large square brackets are either positive definite or semidefinite (a zero difference in loss is possible if θ takes on the value mentioned in footnote 15, so that the correction in ΔX due to Σ is just offset by that due to θ). The implications of a value of 0 for either Σ or θ can easily be deduced from (27), and in general it can be shown that the improvement will be "small" if the variances and covariances contained in Σ and θ are small relative to the squares and cross-products of the variables themselves.

As stated earlier in the paper, I have not had the opportunity to experiment with this proposed solution to the stochastic control problem, as we have not yet done any stochastic simulation with the new model at the University of Michigan. Thus, it is impossible to say with any degree of certainty how it would work out in practice. When such an opportunity becomes available, I suggest the following procedure:

1. Solve the deterministic control problem in the iterative manner described in Section III.
2. On that path, conduct a sufficient number of stochastic simulations to provide reliable estimates of W^* and V, and of the weighted covariance matrices Σ and θ.

3. Reoptimize utilizing equation (23) or (24).

4. Return to step 2.

By utilizing this procedure we would soon learn whether a single pass through steps 2 and 3 is sufficient, or whether several iterations are necessary. If the difference between W^* and the deterministic Y^* is as small as some of the values obtained by Muench, *et al.* [9] and the difference between V and the deterministic U is correspondingly small, a second iteration may not be required. In principle, however, and ignoring costs, one could continue iterating until any convergence criteria were met, but we would soon run into the same problems as before, regarding the solution program convergence criteria and the loss of significance when the perturbations are very small.

VI

In this paper I have described a simple method for selecting, within either a deterministic or a stochastic framework, an optimal set of values for several policy instruments so as to minimize a quadratic loss function containing both those instruments and a set of target variables. The use of the method in the deterministic case was illustrated with an application to a medium-sized nonlinear model of the U.S. economy.

The procedure used does not yield an exact solution, partly because of its own nature, and partly because of the nature of the econometric model. Its accuracy can be increased at a cost, but I have questioned whether this increase in accuracy is real or only imagined, given the inaccuracies in the model itself. Optimal policy calculations can in general be no more exact than the model solution from which they are derived, and we should not overlook this fact in our quest for ever more accurate control techniques.

I have also raised the question of whether, in practice, the cost of carrying out the stochastic optimization procedure will exceed the value of the resulting improvement in control. This is an empirical question, however, and could easily be explored with the framework developed in Section V. I would not be surprised if we were to find that (given our current state of knowledge about the economy) for practical policy-making purposes a solution to the deterministic control problem provides a sufficiently high degree of accuracy.

Whatever the answer to the question raised in the preceding paragraph, the method I have described here has the advantage of being relatively simple, practical, and inexpensive to implement, and at the same time provides a mechanism for obtaining as exact a solution as the policy-maker is willing to trust.

University of Michigan

References

[1] Fair, Ray C., "On the Solution of Optimal Control Problems as Maximization Problems," this issue.

[2] Friedlaender, Ann F., "Macro Policy Goals in the Postwar Period: A Study in Revealed Preference," *The Quarterly Journal of Economics*, Vol. 87, February 1973, pp. 25–43.

[3] Friedman, Benjamin M., *Methods in Optimization for Economic Stabilization Policy*, Ph.D. thesis, Harvard University, 1971.

[4] Helliwell, John F., Harold T. Shapiro, Gordon R. Sparks, Ian A. Stewart, Frederick W. Gorbet, and Donald R. Stephenson, *The Structure of RDX2*, Ottawa, Bank of Canada, 1971. 2 Parts, 267 pp. and 128 pp. (Bank of Canada Research Studies, No. 7).
[5] Holbrook, Robert S., *An Approach to the Choice of Optimal Policy Using Large Econometric Models*, Ottawa, Bank of Canada, 1973. 105 pp. (Bank of Canada Staff Research Studies, No. 8).
[6] Holbrook, Robert S., "Optimal Policy Choice Under a Nonlinear Constraint: An Iterative Application of Linear Techniques," *Journal of Money, Credit, and Banking*, forthcoming.
[7] Howrey, E. Philip and H. H. Kelejian, "Simulation versus Analytical Solutions: The Case of Econometric Models," in *Computer Simulation Experiments with Models of Economic Systems*, edited by Thomas H. Naylor, John Wiley, New York, 1971.
[8] Hymans, Saul H. and Harold T. Shapiro, "The Michigan Quarterly Econometric Model of the U.S. Economy," in *The Economic Outlook for 1973*, Research Seminar in Quantitative Economics, Department of Economics, University of Michigan, 1973.
[9] Muench, T., A. Rolnick, N. Wallace, and W. Weiler, *Tests for Structural Change and Prediction Intervals for the Reduced Forms of Two Structural Models of the U.S.: The FRB-MIT and Michigan Quarterly Models*, Minneapolis, Federal Reserve Bank of Minneapolis (Working Paper No. 19).
[10] Polak, Elijah, *Computational Methods in Optimization*, Academic Press, New York, 1971.
[11] Shapiro, Harold T., "The Efficacy of Monetary and Fiscal Policy," *Journal of Money, Credit, and Banking*, Vol. 3, May 1971, pp. 550–554.