A RATIONAL EXPECTATIONS FRAMEWORK
FOR SHORT RUN POLICY ANALYSIS

by

Christopher A. Sims

Discussion Paper No. 223, October, 1985

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minn 55455

A RATIONAL EXPECTATIONS FRAMEWORK FOR SHORT RUN POLICY ANALYSIS

This paper is a revision of one prepared for presentation at the "New Approaches to Monetary Economics" conference, May 23-24 1985, at the IC^2 Institute of the University of Texas at Austin. The research for the paper was supported in part by NSF grant SES 8309329.

---

1. Introduction

There is increasing recognition that Lucas's [1976] critique of econometric policy evaluation, at least under its usual interpretation, is logically flawed. The point has been made forcefully recently by Sargent [1984] and by Cooley, Leroy and Rahman [1984] (henceforth referred to as CLR), as well as in my own paper [1982]. The problem is that if the parameters of the policy "rule" are subject to change, as they must be if it makes sense to evaluate changes in them, then the public must recognize this fact and have a probability distribution over the parameters of the rule. But then these "parameters" are themselves "policy variables", taking on a time series of values drawn from some probability law. Predicting how the economy will behave if we set the parameters of the rule at some value and keep them there is logically equivalent to predicting the behavior of the economy conditional on a certain path of a policy variable. Yet this is just the kind of exercise which Lucas claimed to be meaningless.

It is also evident that the methods of policy evaluation which Lucas criticized are still in wide use nine years after the appearance of his paper. During discussions of monetary and fiscal policy, statistical models prepared by the Congressional Budget Office, the Federal Reserve Board, numerous other agencies, and by private entities are used to prepare predictions of the likely future path of the economy conditional on various possible paths for policy variables. These conditional projections influence policy-makers' views of the likely consequences of the choices they must make. Though Lucas suggested an alternative paradigm for policy analysis, it is still little used.

Nonetheless, we are not quite to the point where well-trained young macroeconomists collaborate in preparing and interpreting conditional projections of the effects of alternative paths for policy variables without feeling queasy. Sargent, while clearly explaining the problems with the rational expectations paradigm for policy analysis and finding no way around them, claims we must ignore them if we are to avoid the conclusion that policy recommendations of any kind are meaningless. CLR show that projections of the future path of the economy conditional on paths of policy variables are not meaningless, even in the presence of shifts in policy rule. But they never explicitly address the central question of whether such conditional projections could ever be used in the process of policy choice without invalidating them. My own paper [1982] asserts that policy choice based on conditional projection from models which do not identify the parameters of tastes and technology can be logically coherent, but it does not support the assertion with formal modeling and many remain unconvinced.

The remainder of this paper attempts to make that assertion more convincing by analyzing more closely how properly constructed and interpreted conditional projections sidestep the Lucas critique and by presenting examples of model economies in which policy makers steadily make good use of conditional projections from "loosely identified" statistical models.

2. Finding a Coherent Interpretation of the Lucas Critique

Lucas formulates his critique most baldly and succinctly at the beginning of section 6 of his paper, where he writes

> ...there are compelling empirical and theoretical reasons for believing that a structure of the form
>
> 1) $y_{t+1} = F(y_t, x_t, \theta, \epsilon_t)$
>
> (F known, θ fixed, x_t "arbitrary") will not be of use for forecasting and policy evaluation in actual economies.

He goes on to observe that for short-term forecasting the problem can avoided, indeed has been avoided in practice, by allowing the "parameters" to drift in time. But for policy evaluation he argues that a completely different approach is required, in which models are given the form

2) $x_t = G(y_t, \gamma, \eta_t)$

3) $y_{t+1} = F(y_t, x_t, \theta(\gamma), \epsilon_t)$,

where G and F are known, γ is a fixed parameter vector, and η_t and ϵ_t are vectors of disturbances. The econometric problem is now that of estimating the function $\theta(\gamma)$, not a fixed vector of real numbers θ, and policy evaluation is performed by considering the effects of alternative settings for γ, not by comparing choices of paths for x.

These assertions make no sense if taken at face value. The most widely used method for allowing for parameter drift in forecasting models takes the "parameters" to be a time series evolving according to

4) $\theta(t) = H(\theta(t-1), \alpha, v(t))$,

where v is a vector of random disturbances, H is a known function, and α is a fixed vector of parameters. By recursively substituting lagged versions of (4) into itself, we can obtain

5) $\theta(t) = H_1(\alpha, V(t), \theta(0))$,

where $V(t) = \{v(1),...,v(t)\}$. If the influence of $\theta(0)$ on H_1 is non-negligible, we can merge it with the parameter vector α, to form the new parameter vector $\alpha^*=(\alpha,\theta(0))$, so that

6) $\theta(t) = H^*(\alpha^*, V(t))$.

Substituting (6) into (1) produces

7) $y_{t+1} = F(y_t, x_t, H^*(\alpha^*, V(t)), \epsilon_t) =$

$$F^*(y_t, x_t, \alpha^*, \epsilon^*(t)) ,$$

where $\epsilon^*(t)=(V(t),\epsilon(t))$. Now obviously F^* in (7) has exactly the same form as F in (1). If a "time-varying parameter" model like that described by F^* can in fact provide good forecasts, which Lucas claims it can, then it would seem he cannot consistently claim also that a structure of the form (1) "will not be of use for forecasting... in actual economies." Of course, if F in (1) is linear, or in some other sense "simple", F^* in (7) will generally not be. But Lucas says only that F is known, not that it is linear, and he undoubtedly meant to be saying something stronger than that we need nonlinear models.

There is a similar problem with the claim that evaluating policy by using (2) and (3) to gauge the consequences of changing γ is different from using (1) to gauge the consequences of various choices of time path for x. The "fixed" parameter γ is itself necessarily a time series with at least two possible values (the value before we change policy and the value after). Putting the appropriate subscript on γ and substituting (2) into (3) gives us

$$8)\quad y_{t+1} = F(y_t, G(y_t, \gamma_t, \eta_t), \theta(\gamma_t), \epsilon_t)$$

$$= F^{**}(y_t, \gamma_t, \epsilon_t^{**}) \ ,$$

where $\epsilon_t^{**} = (\epsilon_t, \eta_t)$. Now we have transformed (2) and (3) into (8), wherein F^{**} has exactly the functional form of F, with playing the role of x. There are no unknown parameters in (8) at all, but of course in practice there would be unknown parameters in the functional form of the $\theta(\gamma_t)$ function, to take the place of the original fixed θ parameters.

Again, if F were linear, F^{**} would probably not be, but surely most economists interpret Lucas's argument as asserting more than that we will need nonlinear models to do good policy analysis.

It is also true that Lucas wants us to take γ as "fixed". Of course if we contemplate changing γ it cannot really be fixed, but the spirit of the argument is that we ought to consider once-and-for-all changes in γ, i.e. "paths" of γ_t which are constant up to some date T and constant thereafter, with a discontinuity at T. Further according to this interpretation, we ought to concentrate on predicting the long run effects of the change, after has been at its new level long enough for behavior to have completely adjusted. From the point of view of equation (8), this is the suggestion that we should limit ourselves to comparative statics based on the long run properties of the model. But this recommendation, if we admit that F^{**} is the same form as displayed by standard econometric models, is not at all revolutionary. And it is a dubious recommendation -- in practice econometric models are probably less reliable in their long run properties than in their short run properties.

Some readers undoubtedly have given Lucas's critique the interpretation outlined above. Sargent's paper is one example. Under this interpretation, equation (1) is interpreted as general enough to correspond to any statistical model relating endogenous variables y to lagged endogenous variables, policy variables x, some parameters θ, and random disturbances ϵ. The claim is that any use of such a model to make policy will have to change the model, making the model invalid. This leads to skepticism toward all use of data in forming policy, since there is no way to do so without using a model in the general form (1). Concentration on choice of the parameters of a policy rule is of no help, since there is no logical distinction between such "parameters" and "policy variables". Thus on this interpretation, if one takes the rational expectations critique of standard econometric policy evaluation seriously it applies with equal force to the rational expectations program for "correct" econometric evaluation of policy.

As we will see in the next section through an example, the argument in this form is simply incorrect. It is possible for an optimizing, benevolent, immortal policy maker to make policy every period forever by choosing among conditional projections from a statistical model while the model remains accurate.

However, this version of the argument distorts the original. When Lucas writes down (1) in section 6 of his paper, he introduces it as a "structure" of the form (1). By this he probably means to imply that F is not some general statistical model, but the kind of entity he described in more detail when using (1) earlier in the paper. There he requires that "the function F and the parameter vector θ are derived from decision rules (demand and supply functions) of agents in the economy, and these decisions are, theoretically, optimal given the situation in which each agent is placed." Furthermore, in the examples Lucas considers there is in every case one or more functions contained in the model which represent or are directly affected by agents' expectation-formation rules.

The lesson of rational expectations is that when we use a model in whose functional form is embedded agents' expectational rules, we are likely to make errors even in forecasting if we insist that those expectational rules are fixed through time, and that we will make even more serious errors in policy evaluation if we pretend that those rules will remain fixed despite changes in policy which make them clearly suboptimal. The difference between (1) and the system (2-3) as frameworks for policy analysis is not the superficial one that in the latter we think of ourselves as choosing a "fixed parameter" while in the former we are choosing "arbitrary" values of policy "variables" x. The difference is that in (1) the parameters in fact depend on the hidden policy variable γ. If we try to use (1) to guess the effects of various x paths which are in fact accompanied by changes in γ, we will make errors. The advantage of (2-3), or the equivalent model (8), is that either of them takes proper account of the effect of γ on θ.

Thus we should not necessarily expect a different mathematical form or a different probabilistic treatment of policy variables for models which take proper account of rational expectations. It is even in principle possible that such models could have F^{**} functions which turned out to be a set of linear stochastic difference equations, with the corresponding mistaken F function being of complicated nonlinear form.

Most important, the problems of identification for rational expectations models are not fundamentally different from those for what used to be standard models. In the happy circumstance where the historically observed data contains exogenous variation in policy along the lines we are currently contemplating, we can estimate the effects of our policy choices by reduced form modeling. We can estimate regressions of current data on current and past policy variables and correctly use the results to project the likely effects of our policy choices. To do this we need not separate the effects of policy occuring directly from those occuring indirectly through modifications in expectation-formation rules of the public. It is exactly this point which is so neatly laid out by CLR.

This is not to deny that identification is a hard problem. Identification for purposes of policy evaluation is roughly the same hard problem whether or not we take account of rational expectations. Historical data on policy variables will usually reflect some systematic pattern of response of policy to disturbances originating elsewhere in the economy. Therefore conditional distributions of other variables given policy variables do not necessarily correspond to conditional distributions of those variables given autonomously induced changes in policy variables.

3. Optimal Policy Using Conditional Projections

If policy is made optimally, it is always reacting correctly to all available information about the state of the economy. Presumably then it does not display capricious or arbitrary variation. This suggests that it should contain no autonomous randomness, so that policy variables should be exact functions of past data. In this case, there would be no way to separate the effects of policy variables from the effects of the variables policy depends on, except by use of strong auxiliary assumptions. On the other hand, if policy is made suboptimally, it might contain a lot of capricious variation; we might then easily obtain estimates of the effect of this variation. Yet if policy makers use our estimates to improve policy, the amount of autonomous variation in policy will shrink, the probability structure of the economy is likely to change, and our estimates may quickly become obsolete.

Is there any way that policy could both be chosen optimally, on the basis of a correct model, and at the same time contain enough autonomous variation to allow accurate estimation of the effects of deliberately induced variations in policy?

It is not hard to see that the answer must be yes. All that is required is the existence of some source of variation in policy choice which, as far as the public is concerned, is indistinguishable from an error or a capricious shift in policy choice. One obvious possibility arises when we recognize that macroeconomic policy is in fact set through a political process, in which groups with varying knowledge and objectives contest to influence policy. The public does not know the identity, the objectives or the relative political strength of these groups with certainty. Hence actual policy always contains an unpredictable element from this source. The public has no way of distinguishing an error by one of the political groups in choosing its target policy from a random disturbance in policy from the political process. Hence members of a such a group can accurately project the effects of various settings of policy they might aim for by using historically observed reactions to random shifts in policy induced by the political process. The group will itself, if it behaves optimally according to its own objective function, make its target policy a deterministic function of data it observes. But it can implement that function correctly by, at each date, using a statistical model to make conditional projections of the effects of alternative policy variable paths and choosing the projected path it likes best.

CLR include regime switches in their model, but provide no explanation for why policy differs under the two regimes. It should be apparent, though, that something at least very much like their model could emerge if their two regimes were generated by two optimizing political coalitions._/

_/ Recently Roberds [1985] has produced an explicit model which goes some way toward capturing optimal behavior of stochastically alternating regimes of policy makers. Like CLR, he has the regime switch a purely exogenous random process, but instead of arbitrarily changing policy rules, regime switches in Roberds' model change policy makers' objective functions. His model is not ready to be tacked on to CLR's, however, because it does not follow CLR in giving the public an inference problem in trying to determine the current regime. This means also that, unlike CLR, Roberds does not make the stochastic regime switches a source of identifying variation in policy; the public sees regime switches directly in Roberds' model, so they are an extra observable policy variable rather than an underlying source of variation in observable policy variables. Roberds also ignores the interesting problem of strategic interaction between the two regimes. Each regime takes account of the fact that the public is modeling its behavior, but neither attempts to model the behavior of the other regime or consider the possibility that its own behavior could affect the behavior of the other regime. It would be interesting to see work along the lines Roberds has begun which tied more directly to the setting CLR dealt with.

While policy randomness due to political struggles is probably the most realistic and important source of identifying variation in policy, it leads to analytically challenging models. A simpler way to approach the question of how to design good policy posits a unitary policy authority. In such a framework, the most plausible source of identifying variation in policy is noisy information available to the policy authority which is visible to the econometrician, if at all, only with a delay. The optimizing policy maker will use the noisy information, but because it is noisy his use of it will introduce into policy a "random error". The reaction of the economy to this random error will provide a statistical basis for determining the reaction of the economy to hypothetical optimization errors. To some it may be apparent immediately that this setup will "work". But since I am not aware of a closely similar construction in the literature, we work out an explicit model here.

The typical agent in this economy chooses the vector C_t (consumption) at t. The government chooses G_t at t, government activity per capita. Utility of the agent is

$$9)\qquad E[\sum_{t=1}^{\infty} R^t .5(C_t'AC_t + G_t'BG_t + K_t'FK_t)$$

The technology imposes the constraint

$$10)\qquad K_t = HC_t + MG_t + NK_{t-1} + X_t \quad ,$$

where the stochastic process X is exogenous to both private agents and the government. Though this framework is very general and looks like the canonical quadratic-linear control problem, it is unconventional in that it does not assert that X_t is i.i.d., only that its evolution is unaffected by choices of G, C, or K.

The government and the agents both try to maximize the same objective function (9). There is an "information process" containing three subvectors, Q_t, W_t and Z_t. When private agents choose C_t, they do so with knowledge of all values of Q_s, W_s and Z_s for s t. The government must choose G_t based on knowledge only of these variables for s t-1, plus an observation on Z_t.

We assume special relationships among Q, W, X and Z to generate our example: Q and W are Granger causally prior to Z, X is a linear function of current and past Q, and W alone, and Z_t has the form of a noisy measurement of W_t. I.e., assuming Q, W, and Z form a linear process with an autoregressive representation,

$$11)\qquad E_t\begin{bmatrix} Q_{t+1} \\ W_{t+1} \end{bmatrix} = a * \begin{bmatrix} Q_t \\ W_t \end{bmatrix}$$

12) $$E[Z_t \mid W_s, Q_s, Z_{s-1}, \text{ all } s \leq t] = W_t$$

13) $$X_t = b * \begin{bmatrix} Q_t \\ W_t \end{bmatrix} ,$$

where $a(s)=b(s)=0$ for $s<0$ (a and b are one-sided) and "$E_t[.]$" stands for "$E[.|Q_s,W_s,Z_s$, all s t]".

The Granger causal priority assumption means that the information the government obtains at t is strictly redundant from the point of view of agents at t. This means in turn that the solution to this problem can be found as if there were a single optimizing agent who must choose G_t at a stage where his information set is smaller than it is at the next stage, when he chooses C_t. This would be impossible if, say, private agents also observed only noisy information on X_t at t, but with a noise different from that facing the government.

Because K_{t-1} enters the constraint (10), the public cares about future values of G, leaving a role for the rational expectations hypothesis.

Because the model is quadratic-linear, we can solve it using dynamic certainty-equivalence.

The main point of the example does not depend on the explicit solution of the model. Consider the government's problem at t, which is to choose G_t. Dynamic certainty equivalence tells us that a correct approach to this problem can begin by forming forecasts of the paths X using the data available to the government at this point, i.e. values of Q_s, W_s for $s<t$ together with the current observation on Z_t. The objective function (9) can then be maximized with respect to the future paths of C, G and K, subject to the constraint (10), replacing future X's by their expectations. The G_t which emerges from this computation

will be the correct period-t decision for the government.

However, suppose historical data on C, G, K, and W are publicly available, the government has access to an econometrician capable of estimating a vector autoregression (VAR) for these variables, and current and past values of these variables constitute an information set equivalent to current and past Q, W, and Z. In this case the government can simplify its problem by having the econometrician estimate the vector autoregression and generate predictions with it conditional on various possible paths for G.

In particular, the government should use its observation on Z_t to form an estimate of W_t. It should ask the econometrician to put his vector autoregression in triangular form, in the order W, G, then K and C (I.e., the regression equations for C, G, and K should all contain contemporaneous G and W, and the regression equation for G should contain contemporaneous W.) Then the conditional projections should be formed with all disturbances in the VAR set to zero, except for those in the G equation (which are varied to generate various paths for G) and for the contemporaneous disturbance in the W equation (which is determined by the government's estimate of W_t). If the government evaluates the paths generated this way using the utility function (9), ignoring uncertainty, and chooses the value of G_t associated with the projected path maximizing utility, it will make exactly the optimal choice of G_t.

In proceeding this way the government avoids any need to know the structure of the constraints (10), trusting that the private sector has systematically taken care of optimizing relative to these constraints and will continue to do so.

The reason this procedure works is that a class of paths for C, G, and K in which only the part of the innovation in G which is orthogonal to W varies is a class in which the path of X remains fixed. This is true because the only contemporaneous innovation on which G_t depends is that of Z_t, which in turn depends only on the W_t-innovation and e_t. So the part of G_t orthogonal to the W_t-innovation is a funciton of e_t alone, which is independent of X_t. But if in this class of paths X is being fixed at its projected value given Z_t, while G varies, the paths are all satisfying (10) with this fixed X. Thus the paths being considered are a subset of those considered in the full certainty-equivalent solution when C, K and G are all varied while (10) is imposed and the path of X held fixed at its predicted value. The only question remaining is whether we can be sure that this subset of the feasible paths actually contains the optimum. But it does, because another way to characterize the certainty-equivalent solution is that at t it sets the paths of G, C, K and W equal to their conditional expectations given information available to the government, i.e. data up to t-1 together with Z_t. But this is just what is obtained from the VAR when the innovations in X_t and G_t are set at their conditional expectations given Z_t and past data, while all other innovations in the triangularized system at t and later are set to zero.

The foregoing verbal argument is complete except for the absence of any proof that it is possible for current and past G, C, K, and W to be an information set equivalent to current and past Q, W, and Z. Since the foregoing argument is intricate and possibly hard to follow, and since the general conditions for equivalence of the information sets appear difficult to set down, we will work out a specific example in some detail below. First, though, we display the general form of the first order conditions:

14) C: $AC_t = -H\lambda_t$

15) G: $BG_t = -ME_{Zt}[\lambda_t]$

16) K: $FK_t = \lambda_t - RNE_t[\lambda_{t+1}]$,

where λ_t is a stochastic Lagrange multiplier vector and "$E_{Z\,t}[.]$" means "$E[.|Z_s, Q_{s-1}, W_{s-1}$, all s t]".

Because this is a quadratic-linear problem, we could solve it in closed form by assuming some simple form for the autoregressive representation of the Q, W, Z process, then deriving the decision rules. Even for simple cases, however, the parameters of the decision rules and therefore of the AR representation of the system are fairly complicated nonlinear functions of the problem's original parameterization. To keep the algebra simpler, we will derive a solution "backwards" (see my paper [1984]. We proceed by not using (13), instead assuming directly

$$17) \qquad {}_t = f* \begin{bmatrix} Q_t \\ W_t \\ Z_t \end{bmatrix} ,$$

treating (10) now as defining X_t after (14)-(16) have been used to determine C, G, and K from Q, W, and Z. We cannot use the most straightforward backwards solution method, because we cannot choose f arbitrarily and still guarantee that the constraint in (13) -- that X_t not depend on current Z_t -- be satisfied. Nonetheless, it proves to be easy to adjust our choice of f to impose this constraint.

To make the example really simple, we assume that C, K and G all have dimension one. W and Z must then be scalar also. (Since Z's innovation affects only G, we will not be able to recover Z from data on C, K, G, and W unless G's dimension at least matches Z's; and W and Z are assumed to be of the same dimension.) If C, K, G and W are not to be jointly singular, we must have the dimension of C and K jointly no greater than the dimension of Q, and if we are to be able to recover the history of Q, W, and Z from that of C, K, G, and W, we must have C and K jointly no smaller in dimension than Q. So Q is 2x1.

We assume A=B=F=H=M=N=1. We take a in (11) to be such that

$$18) \qquad E_t\begin{bmatrix} Q_{t+1} \\ W_{t+1} \end{bmatrix} = \text{diag}(\delta)\begin{bmatrix} Q_t \\ W_t \end{bmatrix} ,$$

where δ is a vector with distinct elements all less than one in absolute value. We use lower case letters to refer to the innovations in the corresponding upper case stochastic processes. We define π as the vector of coefficients projecting the innovations in Q and W onto the innovation in Z, i.e.

$$19) \qquad E_t\begin{bmatrix} q_t \\ w_t \end{bmatrix} = \pi z_t .$$

Obviously π is determined by the covariance matrix of the innovations in the information process Q, W, Z. We choose f in (17) so that

$$20) \qquad -\lambda_t = 1'Q_t + W_t + \{\phi/(1-\gamma L)\}Z_t .$$

It will turn out that we can keep Z from affecting X (maintain the validity of (13)) by setting ϕ and γ properly.

Now we can use our simplifying assumptions to rewrite the first order conditions (14)-(16) as

$$21)\qquad C_t = 1'Q_t + W_t + \{\phi/(1-\gamma L)\}Z_t$$

$$22)\qquad G_t = 1'\mathrm{diag}(\delta)\begin{bmatrix}Q_{t-1}\\ W_{t-1}\end{bmatrix} + \{\phi/(1-\delta L)\}Z_t + 1'\pi(Z_t - \delta_3 W_{t-1})$$

$$23)\qquad K_t = -1'Q_t - W_t - \{\phi/(1-\delta L)\}Z_t + RN\{1'\mathrm{diag}(\delta)\begin{bmatrix}Q_t\\ W_t\end{bmatrix} + \delta_3 W_t\} .$$

The simplified version of (10), written as solved for X, is

$$24)\qquad X_t = K_t - NK_{t-1} - C_t - G_t \quad .$$

It is easy to see that the terms in Z which result when (21)-(23) are substituted into (24) are given by

$$25)\qquad [-\{3\phi/(1-\delta L)\} - 1'\pi - NL\{\phi/(1-\gamma L)\}]\ Z_t \quad .$$

This term vanishes if and only if

$$26)\qquad (3+NL) = -1'\pi(1-\gamma L) \quad ,$$

which will be true if $\gamma=-N/3$ and $\phi=-1'\pi/3$. We can choose γ and ϕ to satisfy these relations so long as $|N|<3$.

To verify that the relation between C, K, G on the one hand and Q,Z on the other is invertible, we must check the characteristic roots of the matrix polynomial in the lag operator which makes up the right hand side of (21)-(23), i.e. of

$$27)\qquad \begin{bmatrix} 1 & 1 & \{\phi/(1-\gamma L)\} \\ \delta_1 L & \delta_2 L & \{\phi/(1-\delta L)\}+1'\pi \\ -1+RN\delta_1 & -1+RN\delta_2 & -\{\phi/(1-\gamma L)\} \end{bmatrix}$$

Some algebra leads to the conclusion that there is just one root to the determinant of this matrix, and it is less than one in absolute value so long as $|N|<4$.

Denoting the matrix polyonomial in the lag operator in (27) by (L), the full relation between C, G, K, W and Q, Z, W is

$$28) \quad \begin{bmatrix} C \\ G \\ K \\ W \end{bmatrix} = \left[\begin{array}{c|c} \Gamma(L) & 1 \\ & \delta_3(1-1'\pi)L \\ & -1+RN\delta_3(1+\varphi) \\ 0 & 1 \end{array}\right] \begin{bmatrix} Q \\ Z \\ W \end{bmatrix} ,$$

where the partition of the columns of the matrix on the right-hand side corresonds to a partition of the Q-Z-W vector between Q, Z and W. This means that the relation between innovations in the two processes is

$$29) \quad \begin{bmatrix} c \\ g \\ k \\ w \end{bmatrix} = \begin{bmatrix} 1 & 1 & \varphi & 1 \\ 0 & 0 & \varphi+1'\pi & 0 \\ -1+RN\delta_1 & -1+RN\delta_2 & -\varphi & -1+RN\delta_3(1+\varphi) \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ z \\ w \end{bmatrix} .$$

We see from (29) confirmation of what we have already argued for the general case: the innovation in G is a function of the innovation in Z alone. From (18) we can observe that in this example as in the general case Q and W can be expressed entirely in terms of current and lagged values of their own innovations (a consequence of their not being Granger-caused by Z). Thus a class of projections of future values of G, C, and K generated by varying only the part of g orthogonal to w will leave the corresponding (via the inverse of (28)) class of projections of future Q, W and Z unchanged in Q and W, with only Z varying. Optimal choice of G_t is described by the second row of (28) or (29), and the jointly optimal certainty-equivalent path for all variables conditional on the government's information at t is obtained by setting z_t to its actual value and q_t, w_t according to (19), generating future values for Q, W, Z by solving the equations of their VAR with these initial conditions, and translating these to future values for C, G, K, and W via (28). But as we can now see explicitly, this whole process can be described equivalently as setting w_t and g_t, then generating future values for C, G, K and W according to the VAR for these variables themselves.

4. Remarks

The example in the preceding section is set in a canonical quadratic-linear form, with no constants or linear terms in the objective function and all stochastic terms in the constraints. Models with stochastic terms in the objective function, or cross-product and linear terms there, can be transformed into the canonical form. Thus even if the government could deduce the form of some or all of the constraints on the canonical form from the VAR (which it can't in this example), it would not necessarily have found the economy's "technology".

If Z_t were an observable variable, there would be an exact linear relation connecting Z_t and past data to current G_t which could be quickly deduced from the data. Our example depends on thinking of policy makers as agents who can collect unpublished, informal data and use it to make a guess at W_t, but can't characterize their data as a number and, say, plot it against their previous choices of G_t. This is of course not very realistic, but neither is the idea of a unitary, optimizing macro policy maker. If policy makers were truly rational agents, of course, they could fit and solve econometric models in their heads and would have no need of econometricians.

The example is rigged so that identification does not even require instrumental variables estimation -- it uses a Wold causal chain style of identification. It could have been rigged so that the identification problem was a little more interesting.

The Lucas critique raises no problems for the example in the preceding section, not because expectation-formation does not enter the model (it is an important part of the model's dynamics), but because the model is one in which policy is already optimal and persists in being so. Thus the process of policy choice does not change the expectations formation behavior implicit in the model's structure. The point of the example is that it is one in which the Lucas statement, quoted above, about the uselessness of models in the general form of (1), is completely incorrect if taken at face value.

The only coherent interpretation of the Lucas critique is that it states that if one uses a model which incorrectly describes the reaction of expectations formation to policy choice, it will produce incorrect evaluations of policy. The implication is not that econometric evaluation of policy using models fitted to history is impossible, but that it requires correct specification of the reaction of the economy to policy. And the notion strongly associated with the rational expectations literature, that the only correct way to model such reactions is to ignore the probabilistic character of policy choice, is spurious baggage.

Some readers might be convinced by the example, but suppose that it only shows a minor qualification to the Lucas critique -- that when there are no changes in the policy rule, policy choice using econometric conditional projections is at least possibly logically coherent. When we contemplate truly important changes in policy, or changes which may approach permanence, then we are in the realm of changes in policy rule, and it may be thought that Lucas is still right that econometric conditional projections cannot be used for analyzing changes in rule.

But this brings us back to the argument in the first part of the paper. Changes in rule are changes in policy variables, just as is any other kind of change in policy. Very persistent or very large changes in policy are likely to generate nonlinear effects in the reaction of the economy to policy. These nonlinearities

-- e.g. that the reaction of the economy to 6 years of below average inflation cannot be correctly determined by adding up 12 copies of the reaction of the economy to 2 quarters of below average inflation -- must be modeled accurately if conditional projections with the model are to be accurate. But there is no logical distinction between using an accurate nonlinear model to project the effects of persistent or large policy changes and using an accurate linear model to project the effects of smaller, short run policy changes. It is true that the nonlinear models appropriate to projecting the effects of large or persistent policy changes are likely to be complicated, and that the uncertainty about how the data should be interpreted is likely therefore to be larger in such cases. There may be plausible alternative identifying assumptions available which yield different conditional projections without conflicting with the data. Putting the same thing another way, the role of untested subjective judgement (i.e., "a priori theory") may be large. But this is to some extent true in any econometric policy analysis and raises no new issues of principle.

There may be some policy issues where the simple rational expectations policy analysis paradigm -- treating policy as given by a rule with deterministic parameters, which are to be changed once and for all, with no one knowing beforehand that the change may occur and no one doubting afterward that the change is permanent -- is a useful approximate simplifying assumption. To the extent that the rational expectations literature has led us to suppose that all "real" policy change must fit into this internally inconsistent mold, it is has led us onto sterile ground.

5. Conclusions

It should be clear that, though the model in the previous section is specially rigged, what is crucial to the rigging is the information structure. Whenever the policy authority makes imperfectly predictable choices because it observes a noisy information variable, while the public finds the government's information variable redundant, we will reproduce the identifying assumptions justifying the form of econometric policy evaluation which is valid in this model.

In more complicated models, identification of the structural relation connecting policy actions to the private economy's reaction is likely to be less easy than it is here. Within the context of models which like that in this paper yield a stationary linear structure while policy varies, different assumptions about the source of the variation in policy could make reaction unidentifiable or could lead to a need to identify it by use of techniques which could be close to those of standard simultaneous equations theory. Examples of identification techniques which work directly with a VAR reduced form appear in Blanchard and Watson [198] and Litterman [198].

The generic version of the point this paper makes is as follows: we can ordinarily expect that accurate conditional projections of the effects of policy choices will be useful to policy makers; we can expect that there will be autonomous variation in policy in the data, so that statistical estimation will be valuable in preparing such conditional projections; and the identifying assumptions required to bring the data to bear in forming the projections will be much less than a complete behavioral interpretation of each equation in the probability model, in fact not even requiring separation of expectational dynamics from other dynamics.

REFERENCES

Blanchard, Olivier, and Mark Watson [1984], "Are Business Cycles All Alike", NBER discussion paper, presented at March 1984 conference.

Cooley, Thomas F., Stephen F. LeRoy, and Neil Raymon [1984], "Modeling Policy Interventions," discussion paper, University of California, Santa Barbara, and University of Missouri, Columbia.

Kydland, F. and E. Prescott [1977], "Rules Rather than Discretion: the Inconsistency of Optimal Plans," Journal of Political Economy, v.85.

Litterman, R.L. [1984], "The Costs of Intermediate Targeting," Federal Reserve Bank of Minneapolis Discussion Paper.

Lucas, Robert E., Jr. [1976], in The Phillips Curve and Labor Markets, Karl Brunner and Alan Meltzer, ed.'s. Amsterdam: North Holland.

Roberds, Will [1985], "Policy Analysis under Stochastic Replanning," Federal Reserve Bank of Minneapolis Research Department Working Paper number 278.

Sargent, T.J. [1984], "Vector Autoregressions, Expectations, and Advice," American Economic Review, May proceedings issue.

Sims, C.A. [1975], "Exogeneity and Causal Orderings," in New Methods in Business Cycle Research, Federal Reserve Bank of Minneapolis.

_______ [1982], "Policy Analysis with Econometric Models," Brookings Papers on Economic Activity.

_______ [1984], "Solving Stochastic Equilibrium Models "Backwards", Univerisity of Minnesota Center for Economic Research Discussion Paper number 206.