

- Mishkin, Frederic S., and Adam S. Posen. 1997. Inflation targeting: Lessons from four countries. *Federal Reserve Bank of New York Economic Policy Review* 3, no. 3 (August): 9–110.
- Taylor, John B. 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39:195–214.

Comment James H. Stock

Glenn Rudebusch and Lars Svensson have provided a clear and interesting treatment of a large number of policy rules within a bivariate vector autoregression (VAR). They model the interest rate (the federal funds rate) as an exogenous variable under the perfect control of the Fed. Changes in the interest rate affect the deviation of real output from potential, which in turn affects inflation through an output-based Phillips curve. Control rules are evaluated in terms of their expected loss, which is a function of the variances of inflation, potential GDP, and the interest rate.

Their paper is clearly and precisely written and the results are well presented. Their discussion of loss functions and targets is lucid and compelling. The modeling decisions they made are sensible and permit the evaluation of a large number of rules. In future work along these lines, it would be of interest to consider a larger VAR that includes an additional interest rate (so that the Fed is not implicitly given control over the entire term structure in the simulations). Similarly, most methods for constructing potential GDP are questionable, and theirs is no exception. The pitfalls of estimating potential GDP could be sidestepped by specifying the Phillips curve in terms of unemployment rather than potential GDP. It would be useful to see whether their findings, particularly the importance of large coefficients in Taylor-type rules, hold up under these extensions. These comments are relatively minor, however, and in general their paper constitutes an excellent contribution to the literature on monetary policy rules.

Because Rudebusch and Svensson's paper is so clean and self-contained, in the remainder of these comments I will turn to the broader question that is one of the motivations for this conference, the construction and evaluation of control rules in the presence of model uncertainty. A policy rule that performs well under reasonable perturbations of a model, or under different plausible models, is robust to that model uncertainty. Although policy robustness is an underlying theme of this conference, it is important to emphasize two limitations of the robustness results reported in this volume.

James H. Stock is professor of political economy at the Kennedy School of Government, Harvard University, and a research associate of the National Bureau of Economic Research.

The author is grateful to Glenn Rudebusch, Tom Sargent, and Lars Svensson for helpful discussions. This research was supported in part by National Science Foundation grant SBR-9409629.

First, the “conference rules” have been evaluated by various authors using their estimated models, but each of the estimated models contains considerable model uncertainty arising from the estimation of the model parameters. It is possible that a rule is robust across point estimates of models, which might be similar in important dynamic respects (after all, the models are estimated using the same data), but that the rule is not robust to 1 standard error changes in the parameters of the models. Robustness to sampling uncertainty needs to be investigated more carefully before any conclusions can be drawn about the robustness of the policy rules considered in this volume (I return to this point below).

Second, a theme of several papers is that inflation-forecast-targeting rules (in which the monetary authority adjusts the interest rate to move an inflation forecast toward a target) perform well in many of the models considered here. However, this conclusion is drawn by evaluating the performance of the inflation-targeting rule using the same model that is used to compute the inflation forecast. In contrast, the essence of policy robustness is whether a specific quantitative rule performs well under a model *other than that used to develop the policy*. Rudebusch and Svensson find that inflation-forecast-targeting rules, based on conditional inflation forecasts produced by their models, work well when evaluated using their model. The proper check of robustness, however, is whether inflation-forecast-targeting rules based on, say, the Rudebusch-Svensson model forecasts work well when the true model is something else.

To illustrate this point, suppose the Fed hires Rudebusch and Svensson to make their conditional forecasts: the Fed provides them a trial value of the interest rate, Rudebusch and Svensson compute inflation forecasts, and they iterate until the inflation forecasts from their model satisfy the Fed policymakers. Now suppose, however, that Rudebusch and Svensson’s research assistant mistakenly feeds the conditional U.S. interest rate into a model of the Swedish economy rather than their U.S. model, so that Rudebusch and Svensson report back Swedish rather than U.S. inflation. One would expect this inflation-forecast-targeting rule, thus implemented, to produce outcomes for the U.S. that are badly wrong: the model used to generate the inflation forecasts differs sharply from the true model. While one would hope that such a gross mistake would not happen in practice, the essential point is that evaluating the robustness of inflation-targeting rules requires the evaluation of the model’s conditional forecasts when that model is false. I know of no research on monetary policy rules that undertakes that evaluation.

The remainder of these comments take up this problem in the form of parametric model uncertainty, by which I mean uncertainty that can be summarized as uncertainty about the value of a finite-dimensional parameter. This complements Sargent’s comments on Ball’s paper (chap. 3 of this volume), in which Sargent considers the case of uncertainty that is nonparametric in the sense that the uncertainty can be formalized as over elements of an infinite-dimensional space. In particular, I will consider two approaches to parametric

uncertainty. The first is a Bayesian approach that grows out of Brainard's (1967) early work on parameter uncertainty. I will argue that while this approach is appealing from the perspective of decision theory, and while it can yield intuitive results, in practice it places informational demands on policy-makers that are wholly unrealistic and therefore fails to provide a useful framework for constructing practical policies. In its place, I propose using minimax methods to construct optimal robust policies and implement these methods quantitatively in the Rudebusch-Svensson model.

Bayesian Approaches to Model Uncertainty

The Bayesian decision analytic approach to control under parametric uncertainty posits a loss function that is a function of future macroeconomic variables. The decision maker is assumed to have priors over all parameters in the model. Optimal policy is then solved by finding the policy that minimizes the expected loss, integrating over the parameters with respect to the prior density. This is conventionally done in the context of a single model. However, in this volume several distinct models are presented, so it is of interest to consider the result of this procedure when there is uncertainty over the class of models as well. In particular, consider two stylized single-equation models of inflation:

$$(1) \quad \pi_t = \beta x_{t-1} + \varepsilon_t,$$

$$(2) \quad \pi_t = \alpha \pi_{t-1} + \gamma x_{t-1} + \eta_t,$$

where π_t is inflation and x_t is the control variable. Evidently, the two models differ only in whether lagged inflation has an effect on future inflation. Suppose that the decision maker has Gaussian priors over β in the first model, so that $\beta \sim N(\bar{\beta}, \sigma_\beta^2)$. For the second model, the decision maker has the priors $\gamma \sim N(\bar{\gamma}, \sigma_\gamma^2)$.

Suppose the decision maker has quadratic loss, $(\pi_t - \pi^*)^2$, where π^* is the target rate of inflation. If the decision maker were sure that model (1) is correct, then the optimal policy would be

$$(3) \quad x_{t-1}^{*,1} = [\bar{\beta}/(\bar{\beta}^2 + \sigma_\beta^2)]\pi^*.$$

On the other hand, if the decision maker were sure that model (2) is correct, the optimal policy would be

$$(4) \quad x_{t-1}^{*,2} = [\bar{\gamma}/(\bar{\gamma}^2 + \sigma_\gamma^2)](\pi^* - \alpha \pi_{t-1}).$$

Now suppose that the decision maker does not know which model is correct but is sure that one of them is; he or she assigns prior probability λ to the event that model (1) is the true model. In this case, the optimal policy is

$$(5) \quad x_{t-1}^* = \lambda x_{t-1}^{*,1} + (1 - \lambda)x_{t-1}^{*,2}.$$

The noteworthy feature of this result here is that when there is uncertainty over classes of models rather than just (smooth) uncertainty over the parameters in a model, the optimal policy is a linear combination of the two optimal policies in the individual models. At least in this simple example, then, one could imagine giving a board of policymakers the optimal policies resulting from the individual models and letting each policymaker compute his or her individual weighted average of these model-based policies, based on each individual's views of how likely a particular model is to be correct.

Although this result has intuitive appeal, there are reasons to doubt that its simple lessons can be made general enough to be useful for practical policy making. First, on a technical level, dynamic models with learning imply very different rules, in which there can be experimentation to learn about the parameters of the model (cf. Wieland 1996, 1998). It is not clear how this would generalize to the multimodel setting.

Second and more fundamentally, the calculations here require an unrealistic amount of information. Key to these calculations are the existence of prior distributions, which for nonlinear models need to be joint priors over all the parameters. While it is plausible that policymakers might have opinions about the value of the NAIRU or the slope of the Phillips curve, it is not plausible that they would have opinions about, say, the covariance between $\alpha_{\pi 3}$ and β_{y2} in equations (1) and (2) in Rudebusch and Svensson's paper. Indeed, there has been great debate about how to construct priors for large autoregressive roots in univariate autoregressive models (see, e.g., the special issue of *Econometric Theory*, August/October 1994); I believe that a fair summary of this debate is that various Bayesians have agreed to disagree over how to construct their priors. If experts cannot construct priors for univariate autoregressions, it is entirely unrealistic for noneconometrician policymakers to construct priors for multiequation nonlinear dynamic models. Unfortunately, such priors are a necessity for the foregoing calculations, so the conventional decision analytic approach does not seem to be a promising direction for developing practical policy rules that address model uncertainty. It is therefore useful to explore an alternative approach based on minimax approaches to model uncertainty.

Minimax Approaches to Model Uncertainty

An alternative approach is to evaluate policies by their worst-case performance across the various models under consideration. The best policy from this perspective is the minimax policy that has the lowest maximum risk. Because Rudebusch and Svensson do not consider parameter uncertainty, as an illustration I will consider the effect of parameter uncertainty on policy choice using the Rudebusch-Svensson model.

Specifically, I consider their model (1)–(2), with their point estimates, and focus on the effects of uncertainty in two of the parameters, α_y and β_r . These are the two most interesting parameters of their model from an economic perspective: α_y is the slope of the (potential GDP) Phillips curve, and β_r is the impact effect on the GDP gap of a change in the interest rate.

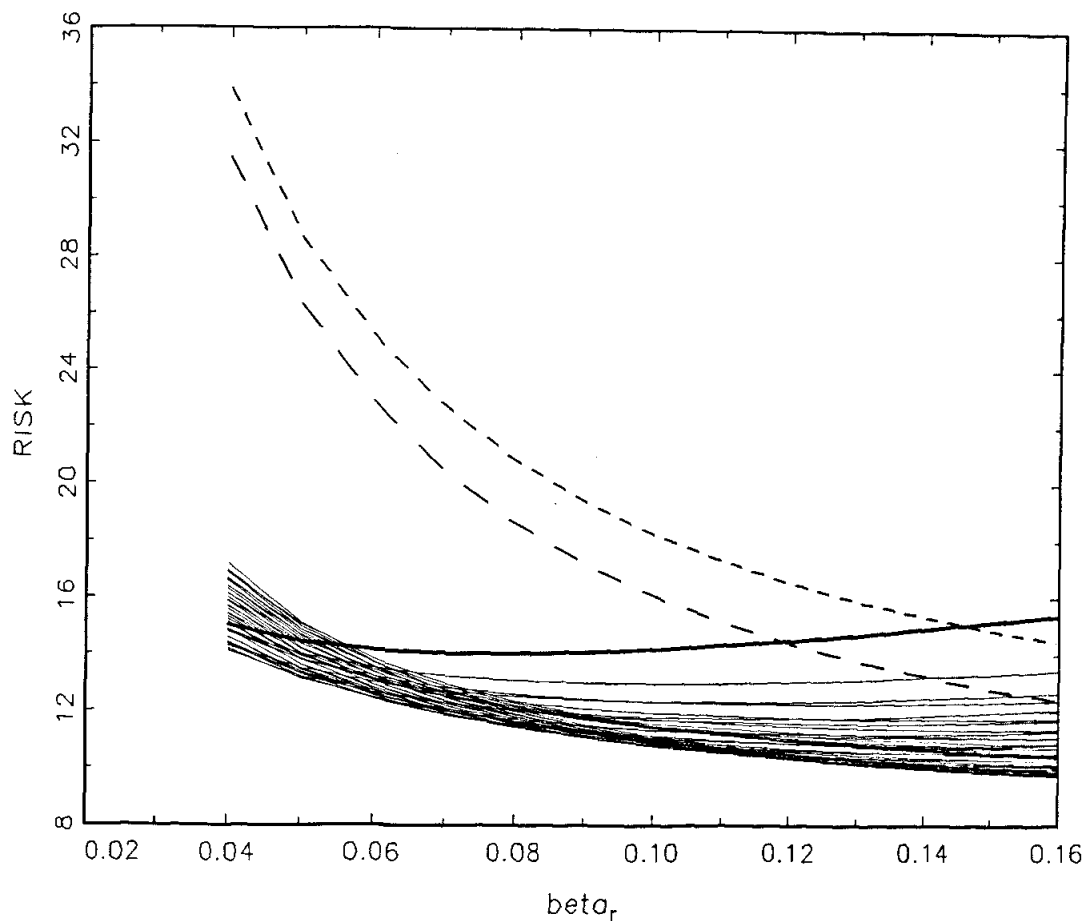


Fig. 5C.1 Risk functions for various policy rules, evaluated in the Rudebusch-Svensson model for $0.04 \leq \beta_r \leq 0.16$ and $\alpha_y = 0.20$

The loss function considered here is the one of the Rudebusch-Svensson loss functions,

$$(6) \quad \text{Loss} = \text{var}(\pi_t) + \text{var}(y_t) + \frac{1}{2} \text{var}(\Delta i_t)$$

in their notation. To capture model uncertainty, values of parameters α_y and β_r within 2 standard errors of their point estimates were considered; that is, the parameters were varied in the ranges $0.08 \leq \alpha_y \leq 0.20$ and $0.04 \leq \beta_r \leq 0.16$.

The policy rules considered here are two-parameter modified Taylor rules of the type considered by Rudebusch and Svensson, specifically,

$$(7) \quad i_t = g_\pi \bar{\pi}_t + g_y y_t.$$

Three types of policy rules were considered: the Taylor rule ($g_\pi = 1.5$, $g_y = 0.5$) and a modified Taylor rule with somewhat more response to output fluctuations ($g_\pi = 1.5$, $g_y = 1.0$); model-specific optimal rules of the type (7), in which the parameters are optimal for particular values of α_y and β_r ; and the minimax rule that minimizes expected loss over all parameter values.

Slices of the risk function surface are presented in figure 5C.1 for these various policy rules; the slices present risk as a function of β_r for $\alpha_y = 0.20$. The upper lines are the risks of the two conference rules, the Taylor rule (*short*

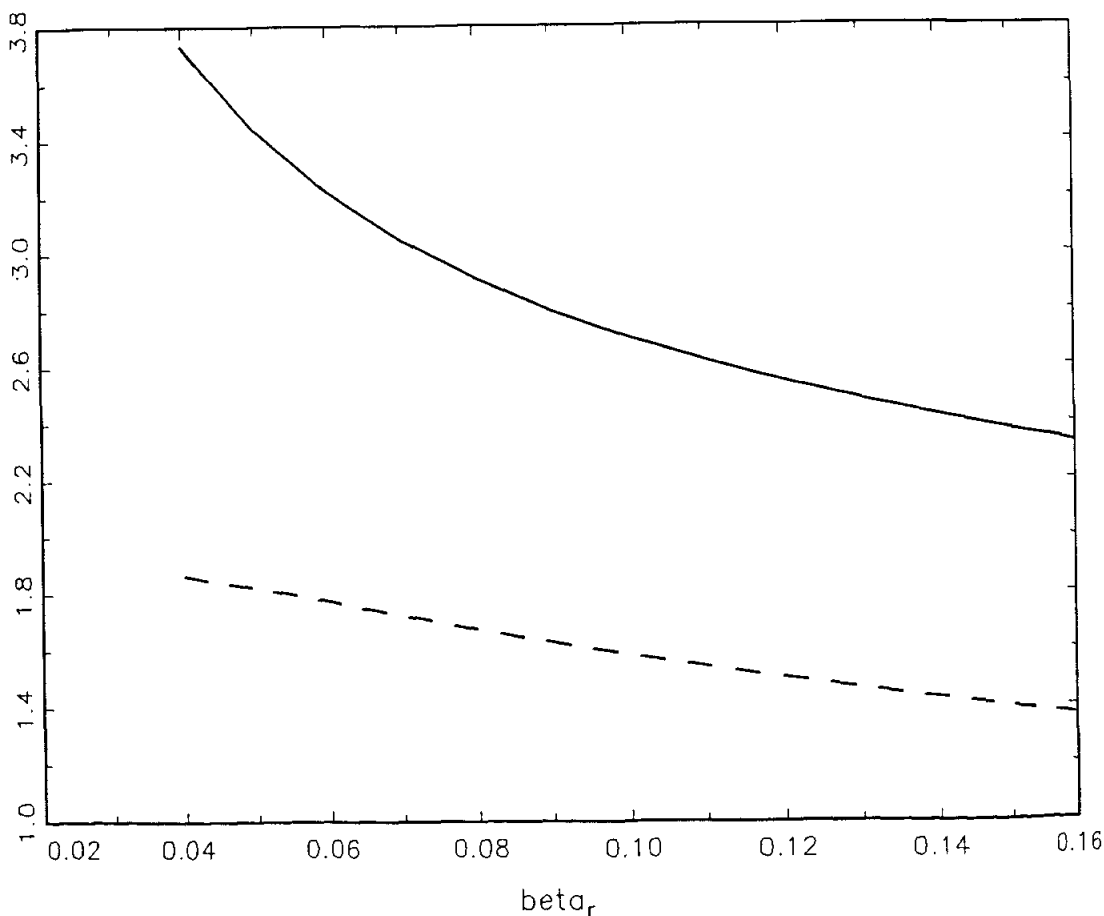


Fig. 5C.2 Optimal parameter values for Taylor-type rules in the Rudebusch-Svensson model for $0.04 \leq \beta_r \leq 0.16$ and $\alpha_y = 0.14$

dashes) and the rule with $g_\pi = 1.5$ and $g_y = 1.0$ (*long dashes*). Each of the light solid lines is the risk function for a policy that is optimal for a particular value of (α_y, β_r) ; the lower envelope of these dotted lines constitutes an envelope of the lowest possible risk, across these parameter values. The heavy solid line is the risk of the policy that is minimax over $0.08 \leq \alpha_y \leq 0.20$ and $0.04 \leq \beta_r \leq 0.16$. (The model-optimal and minimax policies were computed by a simulated annealing algorithm with 1,000 random trials.)

Several observations are apparent. First, the Taylor rule has very large maximum risk. The risk is greatest when β_r is lowest. When monetary policy has little effect (β_r is small), the Taylor rule produces movements in interest rates that are too small to stabilize output and inflation as quantified by the loss function (6). It turns out that the minimax rule has a risk function that is tangent to the risk envelope, with the point of tangency corresponding to the model in which monetary policy has the smallest direct impact on the GDP gap and the Phillips curve is flattest ($\beta_r = 0.04$ and $\alpha_y = 0.08$). In the Rudebusch-Svensson model, this corresponds to the case in which monetary policy is least effective. Here the minimax policy is obtained by producing the optimal rule in the least favorable case for monetary control of inflation and output.

The model-specific optimal parameter values are plotted in figure 5C.2 for

$\alpha_y = 0.14$. Evidently, when the impact effect of monetary policy is small, the optimal response of monetary policy to inflation (*solid line*) and the output gap (*dashed line*) is large. This is the case for the minimax policy, in which $g_\pi = 3.86$ and $g_y = 1.48$. The minimax risk across all models for this policy is 15.61. For the Rudebusch-Svensson model with this parametric uncertainty, the minimax-optimal Taylor-type rule exhibits very strong reactions to inflation and the output gap to guard against the possibility that the true response of the economy to monetary policy is weak.

These results are only illustrative, but they indicate that quite different conclusions can be reached once we admit that there is parameter uncertainty in our models. In the Rudebusch-Svensson model, recognizing parameter uncertainty leads to “conservative” policies that exhibit more aggressive responses than are optimal for the point estimates of the model. It would be interesting to see this sort of analysis undertaken for some of the other models presented in this volume.

References

- Brainard, W. 1967. Uncertainty and the effectiveness of policy. *American Economic Review* 57:411–25.
- Wieland, V. 1996. Monetary policy, parameter uncertainty and optimal learning. Washington, D.C.: Board of Governors of the Federal Reserve System. Manuscript.
- . 1998. Monetary policy and uncertainty about the natural unemployment rate. Finance and Economics Discussion Paper no. 98-22. Washington, D.C.: Board of Governors of the Federal Reserve System.

Discussion Summary

Arturo Estrella asked whether the good performance of smoothing rules in the paper is related to the fact that the IS curve depends on the difference between the short-term nominal interest rate and recent inflation. Changes in the nominal rate reflected in the smoothing rules could be proxying for the difference between the nominal rate and recent inflation. *Svensson* replied that the reason for the bad performance of difference rules was not clear. There is a tendency to get an eigenvalue equal to one or above because the coefficients sum to one in the Phillips curve.

Andrew Haldane noted that most inflation-targeting countries seem to be small open economies. It would therefore be interesting to see how the results of the paper change in an open economy setting. *Svensson's* work and the *Batini* and *Haldane* paper presented at the conference suggest that the change in results would be substantial. Consider the example of simple rules. The two simple rules that perform well in the paper are the Taylor rule and a constant-interest-rate inflation forecast rule. In a model with only two equations, aggregate demand and aggregate supply, these rules, which condition on just two

variables, come not surprisingly close to being fully optimal. In a setting with an important role for exchange rates, Svensson's work on inflation targeting in small open economies indicates that the Taylor rule might not do very well. The second rule holds interest rates constant, which is not admissible with a no-arbitrage condition in a forward-looking open economy setting. Regarding the latter point, *Rudebusch* replied that one of the reasons for the paper to look at constant-interest-rate inflation forecast rules is that inflation-targeting central banks, such as the Bank of England, produce these forecasts in their inflation reports. Therefore, these forecasts seem to be of interest for policy. *Svensson* agreed that simple rules work well because the model is simple enough for inflation and output to be sufficient statistics. With more variables, for example, fiscal policy, simple rules would work less well.

Volker Wieland noted that in the presence of uncertainty about multiplicative parameters, such as the effectiveness of monetary policy, in a linear model, the optimal rule exhibits a more cautious policy response. However, additive uncertainty, such as uncertainty about the natural rate, does not matter in a linear-quadratic framework. In a nonlinear model, additive uncertainty begins to matter. Nonlinearities could, for example, be in the preferences or in the constraints, such as zero-bound constraints on nominal interest rates or nonlinear Phillips curves. *John Williams* mentioned that in his own work on parameter uncertainty using the U.S. model (*Williams* 1997), the value of the loss function and the implicit optimal rule vary greatly with the parameter governing the slope of the Phillips curve. While this parameter is thus a key parameter for monetary policy, it is unfortunately also the least precisely estimated parameter of the model.

Frederic Mishkin made two points justifying rules based on constant-interest-rate inflation forecasts in the context of a closed economy. First, these rules help central banks communicate with the public. Second, these rules help guide discussions about monetary policy in central bank meetings. *Svensson* illustrated these points by noting that in the case of a strict inflation-forecast-targeting rule, the warranted change in interest rates can be expressed as the difference between the unchanged-interest-rate inflation forecast and the inflation target, divided by the policy multiplier, which is easy to communicate. In practice, inflation reports show such constant-interest-rate inflation forecasts.

William Poole stressed that to understand rules for the federal funds rate, it is essential to have two interest rates in the model because of the following reasoning. One of the attractive features of money growth rules is that the economy has a built-in stabilizing mechanism: with constant money growth, shocks to aggregate demand change interest rates, thus keeping the economy from "running off." Something similar happened in recent years with the Federal Reserve's federal funds rate targeting: long interest rates have changed in response to anticipated future federal funds rate moves, even when the Federal Reserve did not change the federal funds rate much. So the fact that bond markets are forward looking is a built-in stabilizing mechanism.

Ben McCallum approved of the emphasis on terminology in the paper. The distinction between targeting a variable and responding to a variable warrants consideration. However, the notion of inflation targeting is odd in the context of high λ -values, that is, when the weight on output variability is much higher than the weight on inflation. *Svensson* agreed with *McCallum* that to use the term “inflation targeting,” the weight on inflation should be significant.

Robert King remarked that the term “interest rate smoothing” is usually used to denote inertia in the level of interest rates, represented by a large response coefficient on the lagged interest rate and small coefficients on contemporaneous output and inflation. From both the Rotemberg and Woodford and the Batini and Haldane papers it seems as if forward-looking models could rationalize that pattern of response. How does such a rule perform in the Rudebusch-Svensson model? *Rudebusch* replied that these rules are not desirable in their model. Moreover, *Rudebusch* disagreed with *King*’s characterization of interest rate smoothing. Whether a rule smooths the interest rate depends on how persistent the arguments of the rule are. John Taylor’s original rule has small response coefficients with no lagged term, and yet, it produces a path for the funds rate that is as smooth as the historical series. In a model with persistence in the output gap and inflation, it is not clear whether a large coefficient on the lagged interest rate is needed to smooth the funds rate.

Ben Friedman noted that the point about Brainard-type uncertainty rules driving the policymaker toward more conservatism depends not only on the model but also on the policy rule and the policy instrument. In the context of this discussion, the policy instrument is the interest rate, and therefore conservatism presumably means less variation in interest rates. In a world with money demand shocks, conservative policy thus leads to higher variability in monetary quantities. However, with a policy rule based on the monetary base, conservatism means that the money base grows more closely along a fixed growth path, which, for the same reasons but now played in reverse, means more interest rate volatility. *Friedman* asked whether this tension is handled conceptually in the approach presented in *Stock*’s comment. *Stock* replied that in the example used in his discussion, the policy rule was based on interest rates. In a comparison of different instruments, it is not obvious that the optimal combination rule is going to be spanned by the submodels. *Stock* also remarked that conservatism does not necessarily mean gradualism. In his simulations, the Taylor rule was the most conservative rule in the sense that the response coefficients were smallest. However, the Taylor rule generated large losses and was far away from a minimax or optimal solution. *Bob Hall* remarked that the same question arises in prison sentences because of the unknown deterrent effect. Is it conservative to give felons short sentences?

Tom Sargent questioned the conclusion drawn in *Stock*’s comment regarding the averaging of rules. If the analysis suggested by *Stock* is pursued with the model at hand, a dynamic model, the posterior over models becomes part of the state of the control problem, such as in *Volker Wieland*’s thesis, implying

that decision rules in this problem become functions of this distribution. Furthermore, the control problem is going to unleash an experimentation motive. If a decision maker is confronted with more than one possible model and a prior over those models, he wants to manipulate the data to learn more. The minimax caution characterization is a static problem, which will not survive the dynamics.

Reference

Williams, John C. 1997. Simple rules for monetary policy. Washington, D.C.: Board of Governors of the Federal Reserve System. Manuscript.