Acknowledging Misspecification in Macroeconomic Theory

Lars Peter Hansen and Thomas J. Sargent

We explore methods for confronting model misspecification in macroeconomics. We construct dynamic equilibria in which private agents and policy makers recognize that models are approximations. We explore two generalizations of rational expectations equilibria. In one of these equilibria, decision makers use dynamic evolution equations that are imperfect statistical approximations, and in the other misspecification is impossible to detect even from infinite samples of time-series data. In the first of these equilibria, decision rules are tailored to be robust to the allowable statistical discrepancies. Using frequency domain methods, we show that robust decision makers treat model misspecification like time-series econometricians.

Key words: Robustness; Model misspecification; Monetary policy; Rational expectations

Lars Peter Hansen: University of Chicago (E-mail: l-hansen@uchicago.edu)
Thomas J. Sargent: Stanford University (E-mail: sargent@leland.stanford.edu)

This paper was prepared for the Society of Economic Dynamics Conference in Costa Rica in June 2000 and for the Ninth International Conference sponsored by the Institute for Monetary and Economic Studies, Bank of Japan, in Tokyo on July 3–4, 2000. We thank Lars Svensson and Grace Tsiang for helpful comments.

I. Rational Expectations versus Misspecification

Subgame perfect and rational expectations equilibrium models do not permit a self-contained analysis of model misspecification. But sometimes model builders suspect misspecification, and so might the agents in their model.[1] To study that, we must modify rational expectations. But in doing so, we want to respect and extend the inspiration underlying rational expectations, which was to deny that a model builder knows more about the data generating mechanism than do the agents inside his model.

This paper describes possible reactions of model builders and agents to two different types of model misspecification. The first type is difficult to detect in time-series samples of the moderate sizes typically at our disposable. A second type of model misspecification is impossible to detect even in infinite samples drawn from an equilibrium.

A. Rational Expectations Models

A model is a probability distribution over a sequence. A rational expectations equilibrium is a fixed point of a mapping from agents' personal models of an economy to the actual model. Much of the empirical power of rational expectations models comes from identifying agents' models with the data generating mechanism. Leading examples are the cross-equation restrictions coming from agents' using conditional expectations to forecast and the moment conditions emanating from Euler equations. A persuasive argument for imposing rational expectations is that agents have incentives to revise their personal models to remove readily detectable gaps between them and the empirical distributions. The rational expectations equilibrium concept is often defended as the limit point of some more or less explicitly specified learning process in which all personal probabilities eventually merge with a model's population probability.

B. Recognitions of Misspecification

A rational expectations equilibrium (indexed by a vector of parameters) *is* a likelihood function. Many authors of rational expectations models express or reveal concerns about model misspecification by declining to use the model (i.e., the likelihood function) for empirical work. One example is the widespread practice of using seasonally adjusted and/or low-frequency adjusted data. Those adjustments have been justified formally by stressing the model's inadequacy at particular frequencies and by appealing to some frequency domain version of an approximation criterion like that of Sims (1972), which is minimized by least squares estimates of a misspecified model. Sims (1993) and Hansen and Sargent (1993) describe explicit justifications that distinguish the model from the unknown true data generating process. They posit that the true generating process has behavior at the seasonal frequencies that cannot be explained by the model except at parameter values that

1. By studying how agents who fear misspecification can promote cautious behavior and boost market prices of risk, we do not intend to deny that economists have made tremendous progress by using equilibrium concepts that ignore model misspecification.

cause bad fits at the non-seasonal frequencies. Maximum likelihood estimates make the model best fit the frequencies contributing the most variance to the data set. When the model is most poorly specified at the seasonal frequencies, then using seasonally adjusted data can trick the maximum likelihood method to emphasize frequencies where the model is better specified. That can give better estimates of parameters describing tastes and technologies.[2]

Less formal reasons for divorcing the data analysis from the model also appeal to model misspecification. For example, calibrators say that their models are approximations aimed at explaining only "stylized" facts or particular features of the time series.

Notice that in both the formal defenses of data filtering and the informal practice of calibration, the economist's model typically remains a rational expectations model inhabited by agents who do not doubt the model. Thus, such analyses do not let the agents inside the economist's model share his doubts about model specification.

II. Agents Who Share Economists' Doubts

But the intent of rational expectations is to put the economist and the agents inside his model on the same footing. Letting the agents contemplate model misspecification reopens fundamental issues that divided Knight, Fellner, and Ellsberg from Savage, and that were set aside when, by adopting rational expectations, macroeconomists erased all model ambiguity from their agents' minds.

A. Savage versus Knight

Knight (1921) distinguished risky events, which could be described by a probability distribution, from a worse type of ignorance that he called uncertainty and that could not be described by a probability distribution. He thought that profits compensated entrepreneurs for bearing uncertainty. Especially in some urn examples that prefigured Ellsberg (1961), we see Knight thinking about decision making in the face of possible model misspecifications.[3] Savage contradicted Knight. Savage (1954) proposed a set of axioms about behavior that undermined Knight's distinction between risk and uncertainty. A person behaving according to Savage's axioms has a well-defined personal probability distribution that rationalizes his behavior as an expected utility maximizer. Savage's system undermined Knight by removing the agent's possible model misspecification as a concern of the model builder.

B. Muth versus Savage

For Savage, it was not an aspect of rationality that personal probabilities be "correct." But for followers of Muth (1961), it was. By equating personal probabilities with objective ones, rational expectations assumes away possible model misspecifications

2. The remarkable feature of these results is that better estimates of taste and technology parameters are acquired by imposing *false* cross-equation restrictions and by accepting *worse* estimates of the parameters governing information and agents' forecasts. Two-sided seasonal adjustment distorts the temporal and information properties processes that agents are trying to forecast.
3. As Ellsberg (1961) points out, Knight's introspection about urns did not produce the paradox that Ellsberg is famous for.

and disposes of diversity of personal probabilities. Rational expectations substantially weakens the appeal of Savage's "solution" of the model specification problems that concerned Knight because it so severely restricts personal probabilities.

C. Ellsberg versus Savage

On the basis of experimental evidence, Ellsberg (1961) and Fellner (1961) challenged Savage's theory. Fellner (1965) proposed a semiprobabilistic framework in which agents used context-specific "slanted probabilities" to make decisions in ways that violate the Savage axioms. The Ellsberg paradox motivated Gilboa and Schmeidler (1989) to formulate a new set of axioms that accommodate model ambiguity. Gilboa and Schmeidler's axioms give agents not a unique personal probability distribution but a *set* of distributions. They posit that agents make decisions as the min-max outcomes of a two-person game in which the agent chooses a utility maximizing decision and a malevolent nature chooses a minimizing probability distribution from within the set chosen by the agents. They show that such behavior can explain the Ellsberg paradox.

Convinced by the Ellsberg paradox and inspired by Gilboa and Schmeidler's formulation, Epstein and Wang (1994), Epstein and Melino (1995), and Chen and Epstein (1998) have constructed dynamic models in which agents are adverse to model ambiguity. Some of this work represents model ambiguity by a class of probability distributions generated by the epsilon-contaminations used in the robust statistics literature.

D. Brunner and Meltzer

Brunner and Meltzer (1967) discussed the role model misspecification in the design of monetary policy. They challenge the existence of a "fully identified, highly confirmed theory of macroeconomic processes." They write:

> we acknowledge that many of the questions raised here can be answered more fully if (and only if) more useful knowledge about the structure of the economy is assumed or obtained. Put otherwise, the theorist may choose to ignore this problem by assuming the possession of reliable information currently outside the scope of quantitative economics. The policy maker is not as fortunate.[4]

Brunner and Meltzer went on to suggest a min-max strategy for selecting among endogenous indicators of monetary policy. They are led to this rule by acknowledging and hence confronting the ambiguity that policy makers have about models of the macroeconomic economy.

We share the concern of Brunner and Meltzer. We will operationalize this concern by introducing formally perturbations or potential model misspecifications into a benchmark dynamic model. These perturbations could be viewed as indexing large families of dynamic models as in dynamic extensions of the Gilboa-Schmeidler multiple prior formulation. We prefer to think about these as errors in a convenient,

4. See Brunner and Meltzer (1967), p. 188.

but misspecified dynamic macroeconomic model. Our choice of perturbations comes from a source that served macroeconomists well before, especially at the dawn of rational expectations.

III. Control Theory

The mathematical literatures on control and estimation theory were the main sources of tools for building rational expectations models. That was natural, because before 1975 or so, control theory was about how to design optimal policies under the assumption that a decision maker's model, typically a controlled Markov process, is correctly specified. This is exactly the problem that rational expectations modelers want the agents inside their models to solve.

But just when macroeconomists were importing the ordinary correct-model version of control theory to refine rational expectations models, control theorists were finding that policies they had designed under the assumption that a model is correct sometimes performed much worse than they should. Such practical problems caused control theorists to soften their assumption about knowing the correct model and to think about ways to design policies for models that were more or less good approximations. Starting in the mid-1970s, they created tools for designing policies that would work well for a set of possible models that were in a sense close to a basic approximating model. In the process, they devised manageable ways of formulating a set of models surrounding an approximating model. To design policies in light of that set of models, they used a min-max approach like that of Gilboa and Schmeidler.

IV. Robust Control Theory

We want to import, adapt, and extend some of the robust control methods to build models of economic agents who experience model ambiguity. We briefly sketch major components of our analysis.

A. The Approximating Model

There is a single approximating model defining transition probabilities. For example, consider the following linear-quadratic state space system:

$$x_{t+1} = Ax_t + Bu_t + Cw_{t+1} \tag{1}$$

$$z_t = Hx_t + Ju_t \tag{2}$$

where x_t is a state vector, u_t is a control vector, and z_t is a target vector, all at date t. In addition, suppose that $\{w_{t+1}\}$ is a vector of independent and identically normally distributed *shocks* with mean zero and covariance matrix given by I. The target vector is used to define preferences via

$$-\frac{1}{2}\sum_{t=0}^{\infty}\beta^t E\,|z_t|^2 \tag{3}$$

where $0 < \beta < 1$ is a discount factor and E is the mathematical expectation operator. The aim of the decision maker is to maximize this objective function by choice of control law $u_t = -Fx_t$.

B. Distortions

A particular way of distorting those probabilities defines a set of models that express an agent's model ambiguity. First, we need a concise way to describe a class of alternative specifications. For a given policy $u = -Fx$, the state equation (1) defines a Markov transition kernel $\pi(x'|x)$. For all x, let $v(x'|x)$ be positive for all x. We can use $v(x'|x)$ to define a distorted model via $\pi^v(x'|x) = v(x'|x)\pi(x'|x)/Ev(x'|x)$, where the division by $Ev(x'|x)$ lets the quotient be a probability. Keeping $v(x'|x)$ strictly positive means that the two models are mutually absolutely continuous, which makes the models difficult to distinguish in small data sets. *Conditional relative entropy* is a measure of the discrepancy between the approximating model and the distorted model. It is defined as

$$\text{ent} = \int \log \frac{\pi^v(x'|x)}{\pi(x'|x)}\,\pi^v(dx'|x). \tag{4}$$

Conditional entropy is thus the conditional expectation of the log likelihood ratio of the distorted with respect to the approximating model, evaluated with respect to the distorting model.

The distortion just described preserves the first-order Markov nature of the model. This occurs because $v(x'|x)$ because only x shows up in the conditioning. More general distortions allow lags of x to show up in the conditioning information set. To take a specific example, we represent perturbations to model (1) by distorting the conditional mean of the shock process away from zero:

$$x_{t+1} = Ax_t + Bu_t + C(w_{t+1} + v_t) \tag{5}$$

$$v_t = f_t(x_t, \ldots, x_{t-n}). \tag{6}$$

Here equation (6) is a set of distortions to the conditional means of the innovations of the state equation. In equation (6), these are permitted to feed back on lagged values of the state and thereby represent misspecified dynamics. For the particular model of the discrepancy (6), it can be established that

$$\text{ent} = \frac{1}{2}v_t'v_t.$$

This distortion substitutes a higher-order nonlinear Markov model for the first-order linear approximating model. More generally, the finite-order Markov structure can be relaxed by supposing that v_t depends on the infinite past of x_t. The essential requirement is the mutual absolute continuity between the approximating model and its perturbed counterpart.

C. Conservative Valuation

We use a min-max operation to define a conservative way of evaluating continuation utility. Let $V(x_{t+1})$ be a continuation value function. Fix a control law $u_t = -Fx_t$ so that the transition law under a distorted model becomes

$$x_{t+1} = A_0 x_t + C(w_{t+1} + v_t), \tag{7}$$

where $A_0 = A - BF$. We define an operator $R_\theta(V(x_{t+1}))$ for distorting the continuation values as the indirect utility function for the following minimization problem:

$$R_\theta(V(x_{t+1})) = \min_{v_t} \{\theta v_t' v_t + E(V(x_{t+1}))\}, \tag{8}$$

where the minimization is subject to constraint (7). Here $\theta \leq +\infty$ is a parameter that penalizes the entropy between the distorted and approximating model and thereby describes the size of the set of alternative models for which the decision maker wants a robust rule. This parameter is context-specific and depends on the confidence in his model that the historical data used to build it inspire. Below, we illustrate how detection probabilities for discriminating between models can be used to discipline the choice of θ.

Let $\hat{v}_t = G_\theta x_t$ attain the minimum on the right side. The following useful robustness bound follows from the minimization on the right side of equation (8):

$$EV(A_0 + C(w_{t+1} + v_t)) \geq R_\theta(V(x_{t+1})) - \theta v_t' v_t. \tag{9}$$

The left side is the expected continuation value under a distorted model. The right side is a lower bound on that continuation value. The first term is a conservative value of the continuation value under the approximating ($v_t = 0$) model. The second term on the right gives a bound on the rate at which performance deteriorates with the entropy of the misspecification. Decreasing the penalty parameter θ lowers the $R_\theta V(x_{t+1})$, thereby increasing the conservative nature of $R_\theta V(x_{t+1})$ as an estimate under the approximating model and also decreasing the rate at which the bound on performance deteriorates with entropy. Thus, the indirect utility function induces a conservative way of evaluating continuation utility that can be regarded as a context-specific distortion of the usual conditional expectation operator. There is a class of operators indexed by a single parameter that summarizes the size of the set of alternative models.

It can be shown that the operator R_θ can be represented as

$$\begin{aligned} R_\theta(V(x_{t+1})) &\approx h^{-1}E_t(h(V(x_{t+1})) \\ &\approx \theta \hat{v}_t' \hat{v}_t + E_t(V(x_{t+1})) \end{aligned}$$

where $h(V) = \exp\left(\frac{-V}{\theta}\right)$ and $\hat{v}_t$ is the minimizing choice of v_t.[5] Note that h is a convex function. This is an operator used by Epstein and Zin (1989), Weil (1993), and others, but with no connection to robustness.

The operator R_θ can be used to rework the theory of asset pricing or to design robust decision rules. The operator also has been used by Hansen and Sargent (1995) to define a discounted version of risk-sensitive preference specification of Jacobson (1973) and Whittle (1990). They parameterize risk-sensitivity by a parameter $\sigma \equiv -\theta^{-1}$, where the $\sigma < 0$ imposes an additional adjustment for risk that acts like a preference for robustness.

D. Robust Decision Rule

A robust decision rule is produced by the Markov-perfect equilibrium of a two-person zero-sum game in which maximizing agent chooses a policy and a minimizing agent chooses a model. To compute a robust control rule, we use the Markov-perfect equilibrium of the two-agent zero-sum dynamic game:[6]

$$V(x) = \max_u \min_v \left\{-\frac{1}{2}z'z + \frac{\beta\theta}{2}v'v + \beta E_t V(x^*)\right\} \tag{10}$$

subject to

$$x^* = Ax + Bu + C(w + v).$$

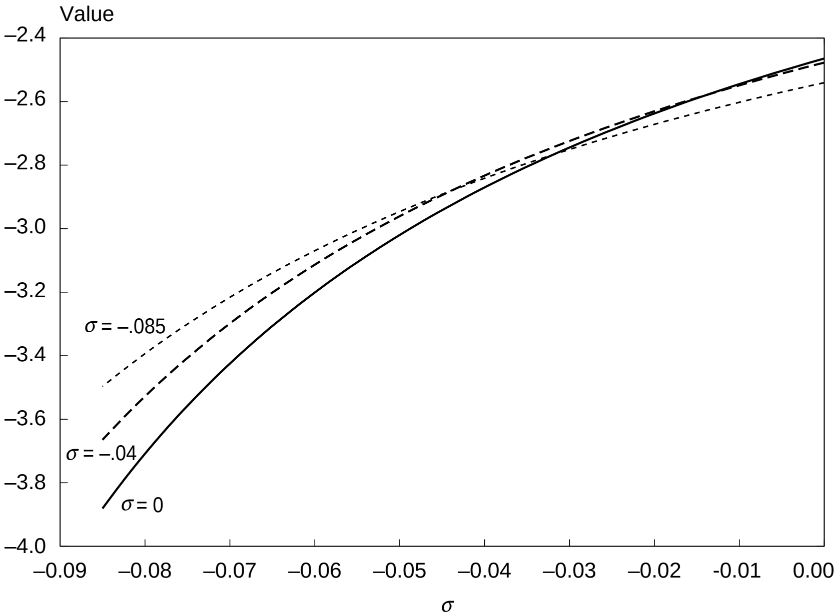
Here, $\theta > 0$ is a penalty parameter that constrains the minimizing agent; θ governs the degree of robustness achieved by the associated decision rule. When the robustness parameter θ takes the value $+\infty$, we have ordinary control theory because it is too costly for the minimizing agent to set a nonzero distortion v_t. Lower values of θ achieve some robustness (again see the role of θ in the robustness bound (9)).

To illustrate robustness, we present some figures from Hansen and Sargent (2000c) based on a monetary policy model of Ball (1999). Ball's model is particularly simple because the private sector is "backward-looking." This reduces the monetary policy problem to a simple control problem. While this simplifies our depiction of robustness, it is of more substantive interest to investigate models in which private sector decision-makers are "forward-looking." Hansen and Sargent (2000b) study models in which both the Federal Reserve and private agents are forward-looking and concerned about model misspecification.

Under various worst-case models indexed by the value of $\sigma = -\theta^{-1}$ on the horizontal axis, Figure 1 shows the values of $-Ez_t'z_t$ for a monetary policy model of Ball (1999) for three rules that we have labeled with values of σ = 0, –.04, –.085. Later, we shall describe how these different settings of $\sigma = -\theta^{-1}$ correspond to different sizes of the set of alternative models for which the decision maker seeks robustness.

5. We use the notation ≈ because there is a difference in the constant term that becomes small when we take a continuous-time diffusion limit.
6. There is only one value function because it is a zero-sum game.

Figure 1 Value of $-Ez_t'z_t$ for Three Decision Rules



Note: When the data are generated by the worst-case model associated with the value of σ on the horizontal axis: $\sigma = 0$ rule (solid line), $\sigma = -.04$ rule (dashed-dotted line), $\sigma = -.085$ (dashed) line. The robustness parameter is given by $\theta = -1/\sigma$.

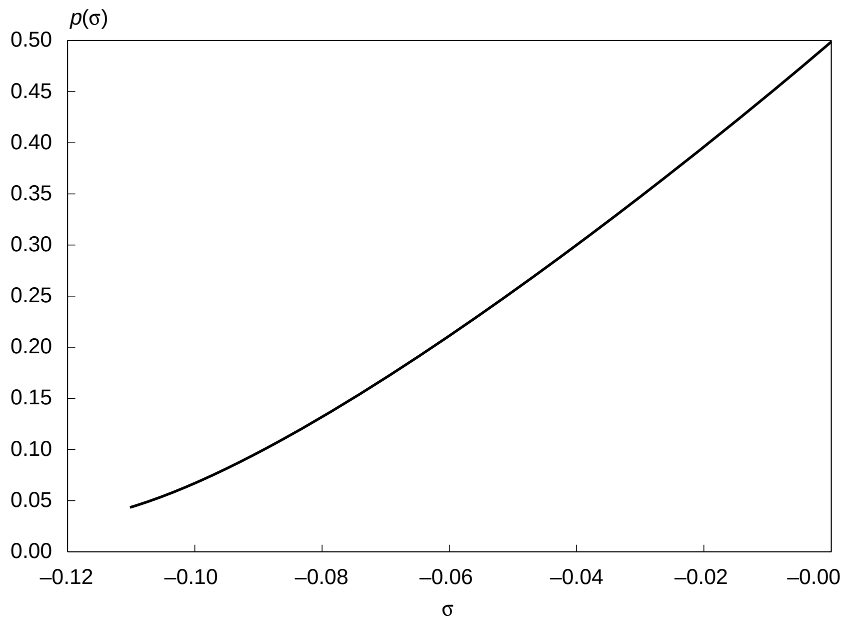
For Ball, $-Ez_t'z_t$ is the sum of variances of inflation and a variable measuring unemployment. The three fixed rules solve equation (10) for the indicated value of σ. The value of $-Ez_t'z_t$ is plotted for each of the fixed rules evaluated for the law of motion $x_{t+1} = (A - BF)x_t + C(w_{t+1} + G(\tilde{\sigma})x_t)$ where $G(\tilde{\sigma})x_t$ denotes the minimizing rule for v_t associated with the value $\sigma = \tilde{\sigma}$ on the horizontal axis. The way the curves cross indicates how the $\sigma = -.04$ and $\sigma = -.085$ rules are not optimal if the model is specified correctly (if $\sigma = 0$ on the horizontal axis), but do better than the optimal rule against the model misspecifications associated with the distortions associated with movements along the horizontal axis. Notice how the σ used to design the rule affects the slope of the payoff line. We now briefly turn to describe how σ might be chosen.

E. Detection Probabilities

The Bellman equation (10) specifies and penalizes model distortions in terms of the conditional relative entropy of a distorted model with respect to an approximating model. Conditional relative entropy is a measure of the discrepancy between two models that appears in the statistical theory of discriminating between two models. We can link detection error probabilities to conditional relative entropy, as in Anderson, Hansen, and Sargent (1999) and Hansen, Sargent, and Wang (2000). This allows us to discipline our choice of the single free parameter θ that our approach brings relative to rational expectations.

For a sample of 147 observations, Figure 2 displays a set of Bayesian detection error probabilities for comparing Ball's model with the worst-case model from equation (10) that is associated with the value of $\sigma = -\theta^{-1}$ on the axis. The detection error probability is .5 for $\sigma = 0$ (Ball's model and the $\sigma = 0$ worst-case model are identical, and therefore detection errors occur half the time). As we lower σ, the worst-case model from equation (10) diverges more and more from Ball's model (because $v_t'v_t$ rises) and the detection error probability falls. For $\sigma = -.04$, the detection error probability is still .25: this high fraction of wrong judgments from a model comparison test tells us that it is difficult to distinguish Ball's model from the worst-case $\sigma = -.04$ model with 147 observations. Therefore, we think it is reasonable for the monetary authority in Ball's model to want to be robust against misspecifications parameterized by such values of σ. In this way, we propose to use a table of detection error probabilities like that encoded in Figure 2 to discipline our selection of σ. See Anderson, Hansen, and Sargent (1999) and Hansen, Sargent, and Wang (2000) for applications to asset pricing.

Figure 2 Detection Error Probability



Note: (Coordinate axis) as a function of $\sigma = -1/\theta$ for Ball's model.

F. Precaution

A preference for robustness induces context-specific precaution. In asset pricing models, this boosts market prices of risk and pushes a model's predictions in the direction of the data with respect to the equity premium puzzle. See Hansen, Sargent, and Tallarini (1999) and Hansen, Sargent, and Wang (2000). In permanent income models, it induces precautionary savings. In sticky-price models of monetary policy, it can induce a policy authority to be more aggressive in response to shocks than one who knows the model. Such precaution has an interpretation in terms of a frequency domain representation of the criterion function under various model perturbations.

For the complex scalar ζ, let $G(\zeta)$ be the transfer function from the shocks w_t to the targets z_t. Let $'$ denote matrix transposition and complex conjugation, $\Gamma = \{\zeta : |\zeta| = \sqrt{\beta}\}$, and $d\lambda(\zeta) = 1/[2\pi i\sqrt{\beta\zeta}]d\zeta.$ Then the criterion (3) can be represented as

$$H_2 = -\int_{\Gamma} \text{trace}[G(\zeta)'G(\zeta)]d\lambda(\zeta).$$

Where $-\delta_j(\zeta)$ is the j-th eigenvalue of $G(\zeta)'G(\zeta)$, we have

$$H_2 = \sum_j \int_{\Gamma} -\delta_j(\zeta)d\lambda(\zeta). \tag{11}$$

Hansen and Sargent (2000a) show that a robust rule is induced by using the criterion

$$\text{ent}(\theta) = \int_{\Gamma} \log\det[\theta I - G(\zeta)'G(\zeta)]d\lambda(\zeta)$$

or

$$\text{ent}(\theta) = \sum_j \int_{\Gamma} \log[\theta - \delta_j(\lambda)]d\lambda(\zeta). \tag{12}$$

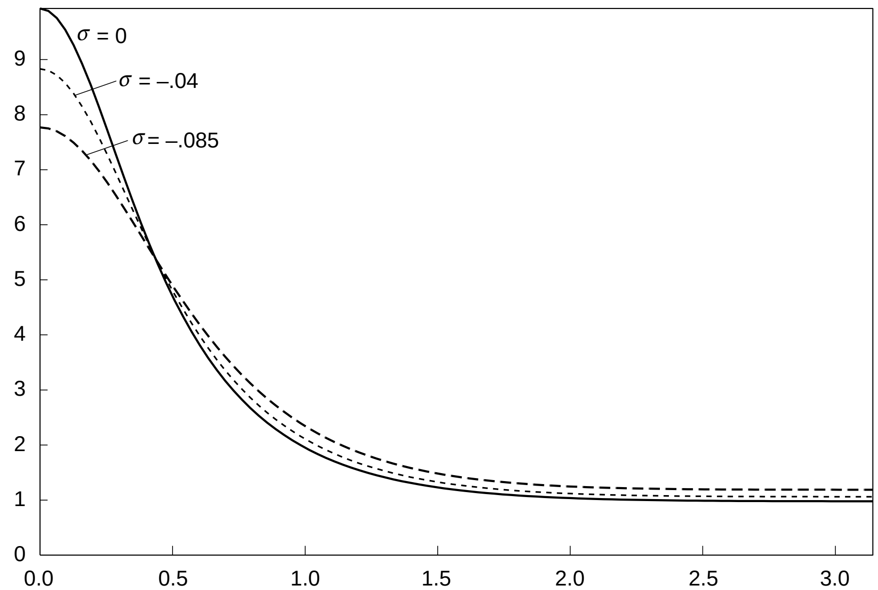
Because $\log(\theta - \delta)$ is a concave function of $-\delta$, criterion (12) is obtained from criterion (11) by putting a concave transformation inside the integration. Aversion to model misspecification is thus represented as additional "risk aversion" across frequencies instead of across states of nature. Under criterion (12), the decision maker prefers decision rules that render trace $G(\zeta)'G(\zeta)$ flat across frequencies.

For example, the curve for $\sigma = 0$ in Figure 3 depicts a frequency domain decomposition of $Ez_t'z_t$ for the optimal rule under Ball's model. Notice that it is biggest at low frequencies. This prompts the minimizing agent in problem (10) to make what Ball's model are supposed to be i.i.d. shocks instead be serially correlated (by making v_t feed back appropriately on x_t). The maximizing agent in problem (10) responds by changing the decision rule so that it is less vulnerable to low-frequency misspecifications. In Ball's model, this can be done by having the monetary authority adjust interest rates more aggressively in its "Taylor rule." Notice how the frequency decompositions under the $\sigma = -.04$ and $\sigma = -.085$ rules are flatter.[7] They thereby achieve robustness by rendering themselves less vulnerable to low-frequency misspecifications.

A permanent income model is also most vulnerable to misspecifications of the income process at the lowest frequencies, since it is designed to do a good job at smoothing high-frequency movements. For a permanent income model, a preference for robustness with respect to the specification of the income process then induces precautionary saving of a type that does not depend on the third derivative of the value function.

7. These are frequency decompositions of the Ball's criterion function operating under the robust rules when the approximating model governs the data.

Figure 3 Frequency Decompositions of $Ez_t'z_t$ for Objective Function of Ball's Model under Three Decision Rules



Note: (σ = 0, –.04, –.085). The robustness parameter satisfies: $\theta = -1/\sigma$.

G. Multi-Agent Settings

We have discussed only single-agent decision theory, despite the presence of the minimizing second agent. The minimizing agent in the Bellman equation (10) is fictitious, a computational device for the maximizing agent to attain a robust decision rule. Macroeconomists' routines use the idea of a representative agent to study aggregate phenomena using single-agent decision theory, for example by studying planning problems and their decentralizations. We can use a version of equation (10) in conjunction with such a representative agent device. A representative agent formulation would attribute a common approximating model and a common set of admissible model perturbations to the representative agent and the government. Robust Ramsey problems can be based on versions of problem (10) augmented with implementability constraints. See Hansen and Sargent (2000a, 2000b) for some examples.

V. Self-Confirming Equilibria

We have focused on misspecifications that are difficult to detect with moderate-sized data sets, but that can be distinguished with infinite ones. We now turn to a more subtle kind of misspecification, one that is beyond the capacity of detection error probabilities to unearth even in infinite samples. It underlies the concept of self-confirming equilibrium, a type of rational expectations that seems natural for macroeconomics. A self-confirming equilibrium attributes possibly distinct personal probabilities (models) to each agent in the model. Those personal probabilities (1) are permitted to differ on events that occur with zero probability in equilibrium, but (2) must agree on events that occur with positive probability in equilibrium. Requirement (2) means that the differences among personal probabilities cannot be detected even from infinite samples from the equilibrium. Fudenberg and Kreps (1995a, 1995b), Fudenberg and Levine (1993), and Sargent (1999) advocate the concept of self-confirming equilibrium partly because it is the natural limit point of a set of adaptive learning schemes (see Fudenberg and Levine [1998]). The argument that agents will eventually adapt to eliminate discrepancies between their model and empirical probabilities strengthens the appeal of a self-confirming equilibrium but does nothing to promote subgame perfection.[8] A self-confirming equilibrium is a type of rational expectations equilibrium. However, feature (1) permits agents to have misspecified models that fit the equilibrium data as well as any other model but that miss the "causal structure." The beliefs of a large player about what will occur off the equilibrium path can influence his choices and therefore outcomes along the equilibrium path.

That self-confirming equilibria permit large players—in particular, governments—to have wrong views provides ways to resolve what we at Minnesota used to call "Wallace's conundrum" in the mid-1970s.[9] Wallace had in mind a subgame perfect equilibrium. He noted that there is no room for policy advice in a rational expectations equilibrium where private agents know the conditional probabilities of future choices of the government. In such a (subgame perfect) equilibrium, there are no free variables for government agents to choose: their behavior rules are already known and responded to by private agents. For example, if a researcher believes that the historical data obey a Ramsey equilibrium for some dynamic optimal tax problem, he has no advice to give other than maybe to change the environment.

Self-confirming equilibria contain some room for advice based on improved specifications of the government's model.[10] But such advice is likely to be resisted because the government's model fits the historical data as well as the critic's model. Therefore, those who criticize the government's model must do so on purely theoretical grounds, or else wait for unprecedented events to expose an inferior fit of the government's model.

8. Bray and Kreps (1987) present a Bayesian analysis of equilibrium learning without misspecification. They express regret that they precluded an analysis of learning *about* a model by having set things up so that their agents can be Bayesians (which means that they know the model from the beginning and learn only *within* the model). Agent's model misspecification, which disappears only eventually, is a key part of how Fudenberg and Levine (1998) and others analyze learning about an equilibrium.
9. See Sargent and Wallace (1976).
10. Much of macroeconomists' advice has been of that form: think of the arguments about the natural unemployment rate theory, which were about using new cross-equation restrictions to reinterpret existing econometric relations.

References

Anderson, E., L. P. Hansen, and T. Sargent, "Robustness, Detection and the Price of Risk," mimeo, 1999.

Ball, L., "Policy Rules for Open Economies," in J. Taylor, ed. *Monetary Policy Rules*, University of Chicago Press, 1999, pp. 127–144.

Bray, M. M., and D. M. Kreps, "Rational Learning and Rational Expectations," in G. R. Feiwel, ed. *Arrow and the Ascent of Modern Economic Theory*, New York University Press, 1987, pp. 597–625.

Brunner, K., and A. H. Meltzer, "The Meaning of Monetary Indicators," in G. Horwich, ed. *Monetary Process and Policy: A Symposium*, Homewood, Illinois: Richard D. Irwin, Inc., 1967.

Chen, Z., and L. G. Epstein, "Ambiguity, Risk, and Asset Returns in Continuous Time," mimeo, 1998.

Ellsberg, D., "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75, 1961, pp. 643–669.

Epstein, L. G., and A. Melino, "A Revealed Preference Analysis of Asset Pricing under Recursive Utility," *Review of Economic Studies*, 62, 1995, pp. 597–618.

———, and T. Wang, "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62 (3), 1994, pp. 283–322.

———, and S. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Asset Returns: A Theoretical Framework," *Econometrica*, 57, 1989, pp. 937–969.

Fellner, W., "Distortion of Subjective Probabilities as a Reaction to Uncertainty," *Quarterly Journal of Economics*, 75, 1961, pp. 670–689.

———, *Probability and Profit: A Study of Economic Behavior along Bayesian Lines*, Homewood, Illinois: Richard D. Irwin, Inc., 1965.

Fudenberg, D., and D. M. Kreps, "Learning in Extensive Games, I: Self-Confirming and Nash Equilibrium," *Games and Economic Behavior*, 8, 1995a, pp. 20–55.

———, and ———, "Learning in Extensive Games, II: Experimentation and Nash Equilibrium," mimeo, Harvard University, 1995b.

———, and D. K. Levine, "Self-Confirming Equilibrium," *Econometrica*, 61, 1993, pp. 523–545.

———, and ———, *The Theory of Learning in Games*, MIT Press, 1998.

Gilboa, I., and D. Schmeidler, "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18 (2), 1989, pp. 141–153.

Hansen, L. P., and T. Sargent, "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 1993, pp. 21–55.

———, and ———, "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control*, 40, 1995, pp. 968–971.

———, and ———, "Elements of Robust Filtering and Control for Macroeconomics," mimeo, 2000a.

———, and ———, "Robust Control and Filtering in Forward-Looking Models," unpublished manuscript, 2000b.

———, and ———, "Wanting Robustness in Macroeconomics," unpublished manuscript, 2000c.

———, ———, and T. Tallarini, "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 1999, pp. 873–907.

———, ———, and N. E. Wang, "Robust Permanent Income and Pricing with Filtering," mimeo, 2000.

Jacobson, D. H., "Optimal Linear Systems with Exponential Performance Criteria and Their Relation to Differential Games," *IEEE Transactions on Automatic Control*, 18, 1973, pp. 124–131.

Knight, F. H., *Risk, Uncertainty and Profit*, Houghton Mifflin Company, 1921.

Muth, J. F., "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29 (3), 1961, pp. 315–335.

Sargent, T. J., *Conquest of American Inflation*, Princeton University Press, 1999.

———, and N. Wallace, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics*, 2 (2), 1976, pp. 169–183.

Savage, L. J., *The Foundations of Statistics*, John Wiley and Sons, 1954.

Sims, C. A., "Approximate Prior Restrictions in Distributed Lag Estimation," *Journal of the American Statistical Association*, 67, 1972, pp. 169–175.

———, "Rational Expectations Modeling with Seasonally Adjusted Data," *Journal of Econometrics*, 55, 1993, pp. 9–19.

Weil, P., "Precautionary Savings and the Permanent Income Hypothesis," *Review of Economic Studies*, 60, 1993, pp. 367–383.

Whittle, P., *Risk-Sensitive Optimal Control*, John Wiley and Sons, 1990.

Comment

FUMIO HAYASHI
University of Tokyo

Serving as a discussant for a Lars Hansen paper is not only an honor but also a challenge. Throughout my career as an academic, I have read his writings and learned a lot from them. I have found his style to be optimal for presenting ideas in the inherently difficult field that he covers. However, this particular paper I agreed to discuss was hard to read. I therefore devote more than half of my discussion to summarize the paper. This is followed by a list of my comments on the paper.

To describe the robust control problem considered in the paper, let us start from the relatively familiar ground, which is the standard linear control theory with quadratic objectives described at the beginning of the paper's Chapter IV. There is a state vector x_t and a control vector u_t, which follow a system of linear difference equations given in equations (1) and (2) of the paper. I will refer to this system as the *model*. Here, $\{w_t\}$ is a sequence of shocks with unit variance. A confusing feature of the paper is that in some parts of the paper this shock sequence is a random sequence with mean zero, while in other parts it is either a deterministic sequence or (if I am not mistaken) a random sequence with nonzero mean (as in the part of the paper that discusses Ball's macro model). To encompass all these cases, I will allow the mean of w_t, denoted μ_t, to differ from zero. So μ_t is a sequence of time-dependent intercepts of the model.

The objective is to maximize the negative of the expected discounted value of a loss function given in equation (3) of the paper, that is, to minimize the expected discounted loss:

$$\sum_{t=0}^{\infty} \beta^t E\,|z_t|^2. \qquad (1)$$

The controller's objective is to choose a feedback rule $u_t = Fx_t$ that minimizes this cost. In the context of monetary policy, the model is a macro model involving such variables as the inflation rate and output as the state variables and the interest rate as the control variable.

The model's parameter vector α consists of the coefficients of the differential equations (A, B, C, H, J). Write the expected loss shown above as

$$R(F, \alpha, \{\mu_t\}). \qquad (2)$$

Here, F is the feedback rule to be chosen by the controller and $\{\mu_t\}$ is the sequence of the mean of the shocks.

With this notation, the standard control problem can be written as

$$\min_F R(F, \alpha, \{0\}), \tag{3}$$

where {0} is a sequence of zeros. As the notation hopefully makes clear, the controller here is assumed to know the true model parameter α and the shock sequence is known to have mean zero (i.e., $\mu_t = 0$ for all t).

In the emerging economic literature on robust control (where Lars Hansen and Thomas Sargent are two of the main contributors), the controller is not assumed to know the model parameter; the coefficients α can differ from the true parameter value. From what I know about the literature, there are two distinct formulations. Both formulations employ the "minimax" criterion. Let Δ be the unknown non-random deviation of the true parameter value from α, and let D be the set of possible deviations. The minimax problem considered by the first formulation can be written as

$$\text{Robust Control I: } \min_F \max_{\Delta \in \mathsf{D}} R(F, \alpha + \Delta, \{0\}). \tag{4}$$

Therefore, it is assumed, as above, that the shock sequence is known to have zero mean. This formulation has been applied to a macro model similar in spirit to Ball's model by Stock (1999) for a particular specification of the set D.

The second formulation, which is the one taken by Hansen and Sargent (and also, I gather, in the engineering literature on robust control), differs from the first in that the controller knows the true coefficient parameters to be α but the true model has nonzero shock means. Let M be the set of possible mean sequences. The second formulation can be written as

$$\text{Robust Control IIa: } \min_F \max_{\{\mu_t\} \in M} R(F, \alpha, \{\mu_t\}). \tag{5}$$

A natural idea, which arises in any constrained optimization problem, is to represent the constraint as a penalty in the objective function. A variant of this second formulation can be written as

$$\text{Robust Control IIb: } \min_F \max_{\{\mu_t\}} \left[R(F, \alpha, \{\mu_t\}) - \theta \sum_{t=0}^{\infty} \beta^t \mu_t' \mu_t \right]. \tag{6}$$

Now, there is no constraint on $\{\mu_t\}$ (or, the set M is replaced by a very large set), but there is a penalty term $\theta \sum_{t=0}^{\infty} \beta^t \mu_t' \mu_t$ in the objective function. If $\theta = \infty$, then the $\{\mu_t\}$ that maximizes the objective function is such that $\mu_t = 0$ for all t. Thus, the case with $\theta = \infty$ is the standard control problem discussed above. This formulation of indexing the problem by θ is attractive for several reasons. First, it provides a convenient way of measuring the degree of ignorance on the part of the controller. Second, as shown in Section IV.E of the paper, some intelligent discussion is possible about how θ might be chosen. Third, as extensively discussed in Hansen and Sargent (2000), the whole apparatus of linear-quadratic control can be employed to compute the solution. Sargent (1999) applied this formulation IIb to compute the robust feedback rule for Ball's macro model for different choices of θ. Some of the results of this computation are also reported in the paper.

Having described what I think is the crux of the paper, I now present several comments in ascending order of substantiveness.

- Stock (1999), which adopted Robust Control I described above, finds that the robust rule entails *policy aggressiveness*, at least for the particular macro model and for a particular range of deviations D. That is, in his words, the robust monetary policy "exhibits very strong reaction to inflation and the output gap to guard against the possibility that the true response of the economy to monetary policy is weak." Sargent (1999), who applied Robust Control IIb to Ball's macro model, also reports policy aggressiveness. Is policy aggressiveness a generic feature of robust control? If not, what are the features of the model that entails policy aggressiveness?
- I would have liked to see a discussion of how the three formulations—Robust Control I, IIa, and IIb—are related. In particular, is there some sort of equivalence theorem stating that for some choice of D and M the feedback rule given by Robust Control I is the same as that of Robust Control IIa?
- My understanding is that Gilboa and Schmeidler (1989) provided a set of axioms under which the minimax rule is a rational behavior when the decision maker does not know the model. It seems to me that Robust Control I and IIa can be justified on those grounds. I would like to know if there is any similar theoretical basis for Robust Control IIb.
- There is a question of which formulation of robust control is more useful to the central bank. Although there seems to be a broad agreement as to what the loss function should be, economists and policy makers disagree about what the true macro model is. It might be plausibly argued that the model uncertainty is about the slope coefficients, rather than about the intercept. If so, Robust Control I should be preferred.
- I am not sure exactly how this paper relates to the theme of the conference, which is about monetary policy in the age of near-zero interest rate. It is probably true that the effect of monetary policy is more uncertain when the interest rate is zero than when positive, but this increased uncertainty is due to a nonlinearity brought about by the non-negativity constraint on the nominal interest rate. It seems to me that the first-order business for researchers is to examine the implication of this nonlinearity.
- The Bank of Japan's monetary policy is determined by committee. My impression from reading the minutes of the Policy Board meeting is that the board members subscribe to different models and that each member thinks his or her own model is the true one. Here, robust control theory may be useful as a *positive* theory of monetary policy. A rational way of making policy would be to take the range of models held by the board members as the D in Robust Control I and adopt the solution to the minimax problem.

References

Gilboa, I., and D. Schmeidler, "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18 (2), 1989, pp. 141–153.

Hansen, L., and T. Sargent, "Elements of Robust Control and Filtering in Macroeconomics," manuscript, 2000.

Sargent, T., "Comment on 'Policy Rules for Open Economies,'" in J. B. Taylor, ed. *Monetary Policy Rules*, University of Chicago Press, 1999, pp. 144–154.

Stock, J., "Comment on 'Policy Rules for Open Economies,'" in J. B. Taylor, ed. *Monetary Policy Rules*, University of Chicago Press, 1999, pp. 253–259.

Comment

TIFF MACKLEM[11]

Bank of Canada

A great deal has changed in a decade. The average rate of inflation in the Organisation for Economic Co-operation and Development (OECD) countries in 1990 was 6.2 percent. By the end of the decade, it had fallen to 2.5 percent. With this success in unwinding inflation, the focus of monetary policy has shifted away from reducing inflation to how best to maintain low and stable inflation so as to reap the benefits of price stability in terms of improved economic performance. With this shift in the focus of policy has come a welcome rebirth of research on monetary control. This latest contribution by Lars Hansen and Thomas Sargent is a leading example of how economic theory and control theory can be combined to confront very real issues faced by central bankers who must make decisions in an uncertain world. Central bankers like low inflation. We are also very pleased that the decision problems we face are being addressed by some of the most influential thinkers in the profession.

My comments proceed in three parts. First, by making reference to some of the inputs into monetary policy decisions at the Bank of Canada, I want to convince you that the issues Hansen and Sargent address really are of concern to central banks. Second, I compare the outcomes of robust control to observed policy, and speculate on the relevance of robust control to practical central banking. Finally, I contrast and compare traditional optimal control in models with multiplicative parameter uncertainty to robust control and suggest that maybe they are not as different as they appear.

I. Uncertainty and Robustness in Policy Making

Hansen and Sargent's starting point can be summarized by two assertions: (1) economic models are approximations; and (2) economic policy should take this into account. It is hard to imagine anyone at a central bank disagreeing with either of these assertions.

11. These comments benefited from discussions with David Longworth, John Murray and James Yetman. The views in this comment are my own. No responsibility for them should be attributed to the Bank of Canada.

Anyone who has worked at a central bank has seen economic thinking and the models change much more quickly than the economy could possibly be changing. Hopefully, this evolution in economic thinking reflects progress toward the "true" model. However, even if—with enough time—we could build a nearly exact model of the economy, the economy is not a constant. Thus, the reality is that the best we can hope for is better approximations. This raises the question: What are central banks doing to take into account the fact that their models are approximations?

The short answer is they diversify. Policy makers do not rely on a single set of assumptions with a single model. Rather, they consider a variety of assumptions and a variety of models, as well as a wide range of evidence that does not fit neatly into their models (precisely because these models are approximations). To make this more concrete, I will briefly describe the inputs into policy decisions at the Bank of Canada—matters I know something about. Most central banks do something roughly similar. Then I want to compare what central banks do to Hansen and Sargent's recommendation—robust monetary control.

Policy decisions at the Bank of Canada are based on a portfolio of analysis, some model-based, some more judgmental. The starting point is the staff economic projection. The projection is based on a structural model of the Canadian economy (the Quarterly Projection Model). The projection is essentially a deterministic simulation of the model using the base case monetary rule (basically a forward-looking Taylor rule of the type estimated by Clarida, Galí, and Gertler [1997] for several countries). Around this base case, we also routinely simulate a variety of alternative scenarios. A first set of alternatives considers the implications of using different monetary rules. These include simulations with the nominal interest rate held constant for six to eight quarters (then reverting to the base case rule), simulations using the markets expectations for interest rates and the exchange rate over the next year or so (again reverting to the base case rule beyond this horizon), and simulations with a more standard Taylor rule that uses the contemporaneous deviation between inflation and its target (rather than the forecast deviation). A second set of alternatives considers the implications of varying key assumptions regarding variables that are exogenous to the model (such as global demand, world commodity prices, and domestic productivity growth), or the dynamics of key endogenous relationships in the model (such as expectations). These alternatives are chosen on the basis of a judgmental assessment of the major risks facing the Canadian economy. Some of these are "bad" outcomes, such as a hard-landing scenario for the U.S. economy or an exogenous upward shift in inflation expectations. Others are more attractive outcomes, such as an acceleration in productivity growth, which nonetheless require the appropriate monetary policy response. So, in the context of the model, substantial effort is expended to consider the robustness of the monetary policy advice coming out of the model.

Moreover, the projection using the structural macro model is only one input into policy. Formal forecasts are also conducted using money-based vector-error-correction models that seek to exploit information in the monetary aggregates. As with the structural-model forecast, several scenarios are considered for these money-based forecasts. In addition, a short-run national forecast is built up from judgmental regional outlooks. Finally, augmenting these forecasts are information from financial markets, from industry liaison activities, and from our ongoing assessment of a broad range of economic indicators. All this is to say that a lot of time and energy is devoted to looking at the issues from a variety of perspectives. This diversity is motivated by uncertainty about the current state of the economy, uncertainty about how the economy will respond to shocks, and uncertainty about the appropriate policy response.

Is this robust monetary control? At a formal level, clearly it is not. But, just as households maximize utility without any knowledge of utility functions or the Kuhn-Tucker theorem, perhaps central bank decision makers use robust rules without formal knowledge of robust control theory or the axiomatic treatment of Gilboa and Schmeidler. To begin to address this issue, it is useful to consider the outcomes implied by robust control theory.

II. Outcomes of Robust Control and Observed Policy Compared

At the heart of robust control is the injection of pessimism into monetary decisions. Formally, pessimism is injected by an inner minimization in which the policy maker performs a worst-case analysis. Against this worst-case scenario, the policy maker then maximizes by choosing the monetary response that leads to the best outcome in this worst-case scenario. The implication of the injection of pessimism is that policy is adjusted (relative to standard control theory) in such a way as to provide insurance against the worst-case outcome.

Hansen and Sargent's latest paper does not discuss in any detail what this implies for the behavior of inflation, output, and interest rates, relative to the standard-control case. And I will confess up front that I do not know what the implications are in general. But at least in the context of simple, linear-quadratic, backward-looking models, previous work by Sargent (1999) has shown that robust rules are more aggressive than the rules produced by "standard" or "modern" control theory. The intuition for this result is straightforward. Taking out insurance against bad outcomes means moving aggressively so as to maximize the chance of heading off these bad outcomes or at least making them less bad.

Does this type of aggressive action aimed at taking out insurance against very bad outcomes look like observed central bank behavior? The short answer is no and yes.

A. No

In simple backward-looking models in which the central bank minimizes the squared deviations of inflation from target and output from potential output, Svensson (1997) and Ball (1997) have shown that standard control theory yields an optimal rule with the Taylor form. However, empirical versions of the Taylor rule that are either estimated or calibrated to fit actual central bank behavior generally yield coefficients on inflation and output that are well below those suggested as optimal in the Ball-Svensson models. This result has spawned a growing literature on why central banks appear to smooth interest rates relative to what is implied by standard control theory.[12] Robust rules, at least in the context of these models, are even more aggressive than standard control rules. While several plausible explanations have been offered for why observed central bank behavior is not as aggressive as standard control rules suggest (e.g., Sack [1998], Woodford [1998]), robust rules raise the bar further. Perhaps even more ingenious arguments can be used to square theory with reality. Nevertheless, robust rules would appear to be moving theory and observed behavior further apart, not closer together. So by this standard, observed policy outcomes do not look much like those implied by robust control.

B. Yes

The discussion so far has focused on the role of the central bank in maintaining price stability. Formally, this is represented by a quadratic objective function in inflation and output. But central banks have another role—financial stability. And here, I think, observed policy looks much more like robust control. In particular, there are numerous examples of situations in which central banks changed interest rates aggressively in what looks to be an attempt to take out insurance against financial instability and the possibility that financial instability may endanger price stability. A few examples:

- The stock market crash of 1987. Following the crash, the Federal Reserve dropped the Federal Funds rate by 125 basis points against a background of robust growth and potential inflationary pressures.
- In September 1998, following the difficulties at Long-Term Capital Management and concern about financial fragility, the Federal Reserve lowered the Federal Funds rate by 75 basis points, again against a background of robust economic growth with the potential for inflationary pressures.
- Following the 1997 financial crisis in Asia, most countries responded by tightening monetary and fiscal policies despite the likely negative impact of the financial turmoil on economic activity and the absence of obvious macro problems—e.g., output growth had been strong, inflation was in hand, and fiscal situations were generally healthy.
- In August 1998, the Bank of Canada raised the bank rate 100 basis points to restore confidence in the external value of the Canadian dollar, despite the fact that the domestic macro situation was being adversely affected by the Asian crisis.

These examples all describe situations in which the central bank took actions that appear puzzling relative to a "narrow macro" view that focuses on inflation and output, but can be rationalized as taking out insurance against a very bad outcome. In this sense, they look very much like a maxi-min strategy in which the central bank action is designed to maximize welfare in the worst-case scenario.

Why does central bank behavior look like robust control in some situations but not in others? The occasions in which behavior looks like robust control share two common features. First, they are typically situations that depend on things that we have only very rudimentary models of—things like the psychology of perceptions and financial market disintermediation. Second, they are situations in which things can change very quickly. These two ingredients appear to fit neatly into Hansen and

12. See Goodhart (1999) for an eloquent exploration of the issue.

Sargent's framework. The absence of good models means there is a considerable degree of Knightian uncertainty in the sense that policy makers do not know the distribution of outcomes. And the reality that the situation can deteriorate very rapidly means the worst-case scenario could develop before the policy makers have another opportunity to reconsider the stance of policy.

In contrast, in the more narrowly defined macro view typically studied in the literature on monetary rules, we do have reasonable models. While crude, these models are based on estimated relationships that do hold more or less over time. Policy makers are therefore better equipped to analyze risks and make probabilistic statements about outcomes. Second inflation and output, unlike perceptions, do not change overnight. Policy makers can therefore afford to take a more gradualist approach, trading off some preemptiveness for the benefits of the more complete information that comes with waiting.

To summarize, robust monetary control appears to have a lot to do with how monetary policy reacts when facing a crisis or a potential crisis, but less to do with monetary policy in more "normal" situations. Looking at the inputs into central bank decisions (at least at the Bank of Canada) the diversified approach suggests that in "normal" times policy makers do not put special emphasis on the worst-case outcomes. Rather they look at all the outcomes and try to chart a "middle" course based on an assessment of the relative probabilities of these outcomes. This produces behavior that is much more gradual than suggested by robust control. In moving forward, I encourage Hansen and Sargent (and others) to apply robust control to the growing literatures on financial stability and dealing with crises.

III. Robust Control and Standard Control Compared

Following Brainard's seminal 1967 paper, the conventional wisdom has been that parameter (or multiplicative) uncertainty results in more muted monetary responses relative to the certainty equivalence case. This conventional wisdom has been reinforced by empirical applications, such as Sack (1998), that show that incorporating multiplicative uncertainty goes at least partway to explaining why observed policy actions are less aggressive than those suggested by standard control theory in models with only additive uncertainty. As shown by Hansen and Sargent, robust control theory, like multiplicative uncertainty, results in a breakdown of certainty equivalence. But while the conventional wisdom is that multiplicative uncertainty results in more muted policy responses, robust control calls for more aggressive actions. Apparently, Brainard and robust control are at odds.

In closing, I want to challenge this interpretation and suggest that the outcomes of robust control and standard control with multiplicative uncertainty may have more in common than has been suggested to date in the literature. To do this, I draw on a recent study by one of my colleagues—Gabriel Srour.

Srour (1999) considers the simple backward-looking Ball-Svensson model in which the policy maker seeks to minimize the expected weighted sum of deviations of inflation from the target and deviations of output from potential:

$$\alpha E(y - y^*)^2 + (1 - \alpha)E(\pi - \pi^*)^2, \tag{1}$$

where π is the inflation rate, π^* is the inflation target, and y and y^* are actual and potential output (in logs). The economy is described by a Phillips curve and an aggregate demand equation:

$$\begin{aligned} \pi_{t+1} - \pi^* &= a_t(\pi_t - \pi^*) + d_t(y_t - y^*) + \varepsilon_{t+1} \\ y_{t+1} - y^* &= b_t(y_t - y^*) - c_t(r_t - r^*) + \eta_{t+1}, \end{aligned} \tag{2}$$

where r_t is the short-term real interest rate (the instrument of monetary policy), r^* is the equilibrium real interest rate, and ε_{t+1} and η_{t+1} are white-noise shocks. The twist that Srour adds relative to Ball and Svensson is that a, b, c, and d are mutually uncorrelated and i.i.d. random variables with means 1, $\bar{b}$, $\bar{c}$, and $\bar{d}$, respectively.

Srour shows that the optimal rule with multiplicative uncertainty continues to have the Taylor form (as in the Ball-Svensson case with only additive uncertainty):

$$r_t - r^* = A(y_t - y^*) + B(\pi_t - \pi^*). \tag{3}$$

The impact of uncertainty is to alter the optimal coefficients A and B relative to what they would be in the certainty equivalence case with no parameter uncertainty (call these values $\bar{A}$ and $\bar{B}$). Srour shows that, in general, the relationship between A and $\bar{A}$ and between B and $\bar{B}$ is ambiguous, and depends on the relative uncertainties of all the parameters in the model. In other words, multiplicative uncertainty does not necessarily lead to more muted monetary responses.

How does this arise? To answer this question, Srour looks at the implications of uncertainty in the parameters a, b, c, and d one at a time. What he shows is that uncertainty in c and d—the coefficients that translate interest rate changes to changes in output and inflation—leads to less aggressive monetary responses (i.e., $A < \bar{A}$ and $B < \bar{B}$). Thus, for example, uncertainty about the interest elasticity of demand leads the policy maker to move interest rates less in response to shocks. This is simply the classic result obtained by Brainard (1967). In contrast, uncertainty about a—how surprises in the Phillips curve translate into future inflation—results in a more aggressive monetary response (i.e., $A > \bar{A}$ and $B > \bar{B}$). The intuition for this last result is that uncertainty about a raises the expected variance of future inflation. To mitigate this, the monetary authority has the incentive to bring inflation back to target more quickly so as to lower the volatility of inflation in future periods.

What I take away from this is that standard control theory with multiplicative uncertainty can produce outcomes that look like the more aggressive outcomes predicted by robust control theory. Perhaps this example is just a special case of robust control theory. Ostensibly, however, this is not the case in the sense that there is no inner minimization in Srour's example. In other words, the aggressive response does not arise because of the injection of pessimism. It arises simply because a more aggressive response will minimize the expected variance.

My final point is a question for Hansen and Sargent: Does robust control always (or even almost always) lead to more aggressive responses than in the certainty equivalence case? I do not pretend to know the answer to this question. But one feature of the examples considered to date with backward-looking, linear-quadratic models raises a question. In the example considered in Sargent (1999), the Phillips curve is of the accelerationist variety—i.e., a = 1. This is the case in which inflation surprises lead to permanently higher inflation (other things equal), and I wonder if it is this assumption that is interacting with the pessimistic element to produce the more aggressive monetary responses. To put it another way, suppose a = 0.5, would robust control still produce a more aggressive response than in the certainty equivalence case? Would robust control rules look more like observed behavior?

I raise this point because with credible inflation control, Sargent (1971) taught us almost three decades ago that a should be less than one. And, indeed consistent with Sargent's classic result, there is now evidence for a number of countries that a has fallen in the 1990s as inflation control has improved and the credibility of price stability has increased. Does this mean that robust control might not predict such aggressive responses after all?

References

Ball, Laurence, "Efficient Rules for Monetary Policy," NBER Working Paper No. 5952, National Bureau of Economic Research, 1997.

Brainard, William, "Uncertainty and the Effectiveness of Policy," *American Economic Review*, 57, Papers and Proceedings, 1967, pp. 411–425.

Clarida, Richard, Jordi Galí, and Mark Gertler, "Monetary Policy Rules in Practice: Some International Evidence," NBER Working Paper No. 6254, National Bureau of Economic Research, 1997.

Goodhart, Charles, "Central Bankers and Uncertainty," Annual Keynes Lecture at the British Academy, published in *Bank of England Quarterly Bulletin*, February 1999, pp. 102–114.

Sack, Brian, "Does the Fed Act Gradually? A VAR Analysis," Finance and Economics Discussion Paper Series No. 1998-17, Board of Governors of the Federal Reserve System, 1998.

Sargent, Thomas, "A Note on the Accelerationist Controversy," *Journal of Money, Credit and Banking*, 8, 1971.

———, "Comment on 'Policy Rules for Open Economies,'" in John B. Taylor, ed. *Monetary Policy Rules*, University of Chicago Press, 1999, pp. 144–154.

Srour, Gabriel, "Inflation Targeting under Uncertainty," Bank of Canada, Technical Report No. 85, 1999.

Svensson, Lars E. O., "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets," *European Economic Review*, 41, 1997, pp. 1111–1146.

Woodford, Michael, "Optimal Monetary Policy Inertia," mimeo, Princeton University, 1998.

General Discussion

In reply to Tiff Macklem's comment on how monetary policy should respond to crisis situations, Lars P. Hansen explained that while it is not explicitly stated, his paper did envision two types of exogenous shocks: a "jump diffusion process" with rapid changes, and a “standard Brownian process.” He said that it would be possible to describe the former as “crisis conditions” and the latter as “ordinary conditions.”

Marvin Goodfriend noted that the model used in Hansen’s paper was an “engineering model” with a high degree of universality, but to make it applicable to the practical design of economic policy, the idiosyncrasy of economic conditions should be carefully examined. Vítor Gaspar pointed out differences in the frameworks appropriate to analyze individual decision-making and collective decision-making. In this context, he emphasized the importance of analyzing the decision-making process in policy committees. Spencer Dale commented on the difficulty of comparative evaluations among various models, which the central bank must confront when seeking robustness for its economic model.

Hiroshi Fujiki agreed with Hansen on the usefulness of the application of asset prices for the central bank, but asked two questions: (1) what kind of asset pricing model was assumed to underlie the argument that changes in the subjective discount factor and robustness could be identified; and (2) whether robustness or the permanent income hypothesis could better explain the weakness of consumption experienced by Japan in recent years. On the latter question, Hansen responded that the central bank should actively utilize the information obtained from asset prices, particularly by estimating equity price models to clarify the role played by robustness.

Shin-ichi Fukuda hinted to the possibility of the existence of multiple equilibria for asset prices in Hansen’s model. Lars E. O. Svensson pointed out that Brunner and Meltzer had discussed robust optimal control more than 30 years ago, and felt that this ought to be referred to in Hansen’s paper.

As a practical implication of the exercise on the monetary policy, Guy M. Meredith suggested that the success of the flexible U.S. monetary policy during the last several years could have many indications for current Japanese monetary policy.