Volume Title: NBER Macroeconomics Annual 2002, Volume 17

Volume Author/Editor: Mark Gertler and Kenneth Rogoff, editors

Volume Publisher: MIT Press

Volume ISBN: 0-262-07246-7

Volume URL: http://www.nber.org/books/gert03-1

Conference Date: April 5-6, 2002

Publication Date: January 2003

Title: "Rules vs. Discretion" After Twenty-Five Years

Author: Nancy L. Stokey

URL: http://www.nber.org/chapters/c11072

Nancy L. Stokey
UNIVERSITY OF CHICAGO

"Rules vs. Discretion" after Twenty-five Years

1. *Introduction*

From time immemorial citizens have complained about their governments. When the government is a greedy despot or the society is composed of private agents with conflicting goals, it is easy to see why complaints arise. Twenty-five years ago Kydland and Prescott (1977) showed something more surprising: even in a society with identical households (with identical tastes and opportunities, and the same choices to make) and a perfectly benevolent government (one that wants to maximize the utility of this representative household), in some circumstances bad outcomes may occur. These situations seem to involve no conflict of interest, either among different groups of households or between the private sector and the government, and the outcomes are "bad" in the sense that better alternatives are obviously available and seem to be—almost—within reach. Settings where this paradox arises include patent protection, capital levies, default on debt, disaster relief, and monetary policy.

Two elements are needed to create such a situation. First, anticipations about future government policy must be important in shaping current decisions in the private sector. Second, there must be a public-good aspect—an *external effect*—of the private-sector choices that are influenced by anticipated policy.

In a setting with these two features, even a benevolent government typically has an incentive to mislead the private sector about the policies that will be implemented in the future, in order to manipulate their cur-

I am grateful to Fernando Alvarez, V. V. Chari, Patrick Kehoe, Robert Lucas, Roger Myerson, and Christopher Phelan for useful discussions, and to the Research Division of the Federal Reserve Bank of Minneapolis for hosting me during a visit where much of this work was done.

rent decisions and enhance the external effect. After the private-sector choices have been made, the government's incentives put weight on only the direct (contemporaneous) effect of the policy. Thus, it has an incentive to implement a different policy from the one announced. If private-sector agents are rational, however, they foresee that the government's incentives will change and refuse to be misled in the first place. The resulting outcome seems bad if the enhanced external effect, which must be forgone, is large. All agents would all be better off if all could be fooled, but rational behavior precludes this possibility: in equilibrium the private sector must anticipate correctly the policymaker's incentives and choices. Thus, the time-consistency problem offers an explanation for what seem to be paradoxically bad policy outcomes.[1]

A key issue in settings where the time-consistency problem arises is the ability or inability of the government to make binding commitments about future policy: *Rules* imply commitment, while *discretion* implies its absence. Commitment is important if anticipations about future government actions influence the current choices of the private agents in the economy. With the ability to commit, the government can tie the hands of its successors in a way that may improve outcomes. Without that ability the private sector fears—with good reason—that today's government will make promises that its successors will refuse to honor.

If commitment is lacking, a framework that incorporates game-theoretic elements is needed to model the policymaker's incentives.[2] And as Barro and Gorden (1983) showed early on, such a formulation also points the way to a resolution of the problem: within a game-theoretic framework it is easy to show that if the game is repeated and agents are not too impatient, there are reputation equilibria in which the "good" outcome prevails along the equilibrium path. That is, a policymaker can be disciplined by reputation considerations even if he has discretion.

The time-consistency issue has been intensively studied over the past twenty-five years, and many of the main theoretical issues have been resolved. Interesting substantive applications are, of course, still being developed. But rather than review the theoretical literature again[3] or attempt

1. If the government is not benevolent—if it has objectives different from those of the private sector—the same incentive to mislead can arise. But with conflicting objectives that is unsurprising. What is astonishing about Kydland and Prescott's examples is that all parties seem to share the same objectives. This appearance is somewhat illusory, in the sense that the payoffs of the private sector agents are symmetric, not identical: the private sector is not a "team." See Chari, Kehoe, and Prescott (1989) for a more detailed discussion.
2. Chari and Kehoe (1990) and Stokey (1991) offer two slightly different general frameworks.
3. See Rogoff (1987, 1989) for an excellent survey of the early literature, Persson and Tabellini (1994) for a discussion of many applied questions, and McCallum (1999) for a discussion of issues related to monetary policy.

to survey the applications (which are too numerous even to list), we will look at two issues that remain. Both deal with choices about the policy regime to be used.

The first issue is reputation building. A policy instrument that can be monitored more closely implies less frequent breakdowns in a reputation equilibrium. Thus, ease of observability is one criterion involved in the choice among discretionary instruments. Here we will look at a central bank deciding whether to peg an exchange rate or to set a rate of money growth. The model focuses on the trade-off between *observability*, the accuracy with which the private sector can monitor the central bank's actions, and *tightness*, how closely the instrument is linked to the object of ultimate interest, the inflation rate. We will show that the ease of monitoring an exchange-rate policy may outweigh other costs it imposes relative to a money growth policy.

The second issue is the robustness of a policy mechanism against mismanagement. One reason to prefer rules over discretion is that governments are not always as intelligent, benevolent, and farsighted as the Ramsey government found in theoretical discussions of policy. Policymakers who are misguided, greedy, or myopic sometimes hold office. Rules that are hard to change may offer protection against these less than ideal types of government officials. The robustness argument is one of the motivations in Friedman's (1948) recommendations on aggregative policy, and it is one that seems worth reviving. Many of the biggest policy blunders seems to arise from incompetence or special-interest-group pressures, rather than the classic time-consistency issue.[4]

Here we will look at robustness using a model in which the type of government in power, Ramsey or myopic, changes randomly from period to period. In this setting the (farsighted) Ramsey government faces an especially difficult task, since the possibility of the myopic type adversely affects private-sector behavior. Hence when the Ramsey government is in power it must distort its own policy in a way that offsets the policy of the myopic type. If the probability of the myopic type is high enough, a simple policy rule can be advantageous. A well-designed rule places an important restriction on the policy of the myopic type, while leading to only a mild change in the policy of the Ramsey type.

4. There are many examples of policy that was arguably well intentioned but surely misguided. Cole and Ohanian (2001) argue that the National Labor Relations Act was important in prolonging the Great Depression by keeping wage rates too high. The inflationary episodes experienced in many countries during the 1970s may offer another example. See Ireland (1999) and Clarida, Gali, and Gertler (1999) for a further discussion. Phelan (2001) offers an interesting model in which a government that is greedy, but also intelligent and patient, may for long episodes behave like one that is benevolent.

These two issues are examined in next two sections. The concluding section discusses some of the results.

2. *Reputation Building*

The ability of a government policymaker to establish and maintain a reputation for reliable conduct depends on how well the public can observe his actions. A policy instrument that is more easily monitored—one that allows the private sector to detect deviations from announced policy rules more easily—has an obvious advantage in allowing the policymaker to build and keep a reputation. Hence observability is often a key issue. But a policy instrument that is more observable may be less tightly connected to the ultimate target, and consequently there is a tension between observability and tightness.[5]

In a recent paper Atkeson and Kehoe (2001) look at the problem of a central bank choosing between two instruments for conducting monetary policy. The bank's options are to peg an exchange rate or to target the rate of money growth. If it pegs an exchange rate, realized inflation is equal to the rate of depreciation in the exchange rate plus an exogenous shock term that represents the foreign rate of inflation. If the bank targets money growth, realized inflation is equal to the rate of money growth plus an exogenous shock term that represents a domestic velocity shock. The central bank chooses its instrument period by period and may switch instruments at any time. If any reversions occur along the equilibrium path, the most severe punishment is implemented.

Notice that with either instrument the object of interest—the inflation rate—is imperfectly related to the bank's action. The exogenous shock terms—the foreign rate of inflation in the first case and the domestic velocity shock in the second—are beyond the bank's control and are unknown when the bank is making its policy decision. In general the two shocks will have different variances, and those variances are important inputs into the banks decision.

The two instruments also differ along a second dimension. The public is assumed to observe the exchange rate directly, so any deviation is immediately detected. That is, the exchange rate is assumed to be a perfectly observable instrument. Consequently, with the exchange rate as the in-

5. Observability is different from what most authors call transparency. In discussions of monetary policy the latter is typically used to refer to the clarity with which the private sector can observe the central bank's *objectives*. The term observability will be used here to refer to the clarity with which the private sector can observe the bank's *actions*.

strument there exist equilibria in which the *threat* of reversion disciplines central-bank behavior, but no reversions actually occur along the equilibrium path.

The money growth rate, on the other hand, is not directly observed. Thus, if the central bank uses the money growth rate for its instrument, the private sector can only infer something about its behavior by looking at the realized rate of inflation. Hence, under a money-growth regime, (accidental) reversions cannot be avoided. As in Green and Porter's (1984) cartel model, the imperfect monitoring technology is the source of these reversions.

Atkeson and Kehoe show that if the central bank can commit to a policy, then it chooses the instrument with the smaller variance for its shock. That is, with commitment only tightness is valued. They also show that if the central bank cannot commit, then it prefers to use the perfectly observable instrument, the exchange rate, even if the variance of the foreign inflation shock is somewhat larger than the variance of the domestic velocity shock.[6]

In this section we will look at a slightly modified version of Atkeson and Kehoe's model that highlights the main conclusions. First, we will require the government to make a one-time decision about which instrument to use, instead of choosing the instrument period by period. Second, we will use reversions to the one-shot Nash equilibrium instead of the most-severe-punishment path. Third, we will formulate the model in the classic Ramsey tradition, as one in which the government's objective is to maximize the utility of the representative household. Finally, we will allow an alternative version of the inflation process under a money-growth rule.

2.1 THE ECONOMY

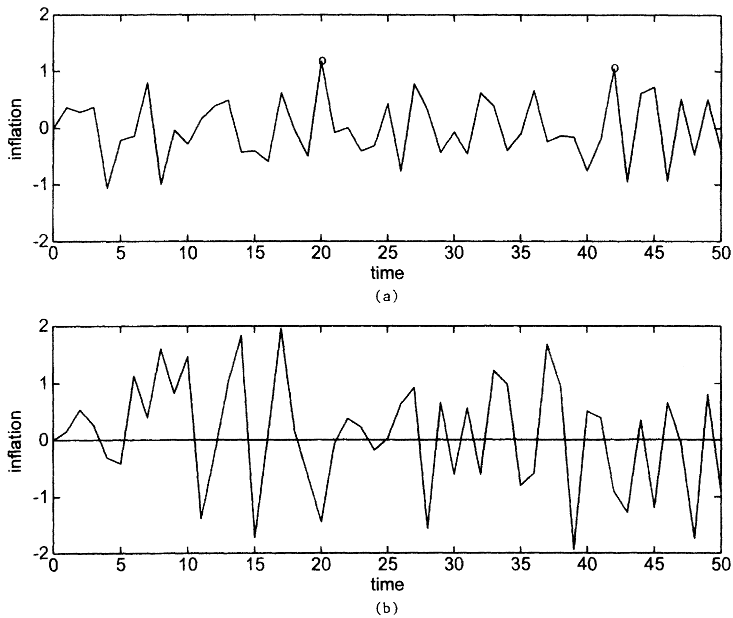
Consider a central bank choosing between money growth and an exchange rate as the instrument for conducting monetary policy. Suppose

$$\mu + v = \pi = e + \zeta,$$

where μ is the money growth rate, v is a velocity shock, π is the inflation rate, e is the rate of depreciation in the exchange rate, and ζ is the foreign rate of inflation, all in logs. Assume that the shocks v, ζ are i.i.d. and

6. Of course, many other issues affect this choice as well. For example, in an early contribution Poole (1970) focuses on the sources of shocks.

Figure 1 REALIZED INFLATION RATE: (a) MONEY-GROWTH REGIME, (b) EXCHANGE-RATE REGIME



independent of each other, with means of zero and variances $\sigma_v^2, \sigma_\zeta^2 > 0$. Assume that π and e are observed.[7]

Under a money-growth policy the instrument is μ, the (noisy) signal is π, and the velocity shock v affects the realized inflation rate. Although $e = \pi - \zeta$ is observed, it is not useful in assessing the central bank's performance: e is a noisy signal about π, and π is observed directly. Under an exchange-rate policy the instrument is e, the (noiseless) signal is e, and the foreign inflation rate ζ affects the realized inflation rate.

Figure 1 illustrates the trade-off. Figure 1a displays the realized rate of

7. To incorporate serial correlation in the shocks, define μ and e to include expected changes in velocity and foreign inflation, respectively. Then v and ζ, interpreted respectively as innovations in velocity and exchange-rate depreciation, are serially uncorrelated and have means of zero. With serial correlation we must also ask whether it remains reasonable to assume that e is perfectly observable. If the foreign rate of inflation is serially correlated, then e is observable only if $E[\pi^f]$, the central bank's forecast of the foreign inflation rate, is observable. If the central bank announces its estimate $E[\pi^f]$ each period, and if the private sector can verify this forecast independently, then the model goes through without change. As Goodfriend (1986) notes, a central bank's main forecasting advantage derives from its earlier access to data. But presumably the domestic central bank has little advantage in acquiring the data relevant for the foreign inflation rate. Hence the assumption that the private sector can verify the bank's announcement seems reasonable.

inflation under a money-growth rule. Since the actual rate of money growth (which is always zero in equilibrium) is not observed by the private sector, the reputation equilibrium involves reversions when the realized inflation rate exceeds some (optimally chosen) threshold. The small circles depict situations where a reversion is triggered.

Figure 1b displays the situation under an exchange-rate rule. The horizontal line is the actual rate of depreciation in the exchange rate, and the fluctuations around it depict the realized rate of inflation. The variance of realized inflation is larger than under a money-growth rule, but since the exchange rate is observed directly, no reversions occur. Thus, the optimal choice trades off the higher ongoing cost of larger fluctuations under the exchange-rate regime against the cost of occasional reversions under a money-growth regime.

A slightly more complicated model of money growth incorporates output growth. Suppose

$$\pi = (\hat{\mu} - g) + u,$$

where $\hat{\mu}$ is money growth, g is real GDP growth over the period, and u is a velocity shock. Let

$$g^e = g + \varepsilon$$

be the central bank's (imperfect) forecast of real growth, where ε is the forecast error. Assume that the shocks ε, u, ζ are i.i.d. and independent of each other, with means of zero and variances σ_ε^2, σ_u^2, $\sigma_\zeta^2 > 0$. Under a money-growth policy the bank can be viewed as choosing

$$\mu = \hat{\mu} - g^e,$$

the excess of money growth over expected real growth. For simplicity we will continue to call μ the rate of money growth. Assume that the private sector cannot observe g^e, but does observe $\hat{\mu}$ and g. Then

$$\hat{\mu} - g = \mu + \varepsilon$$

is its signal about the bank's action, and

$$\pi = \mu + \varepsilon + u$$

is the realized inflation rate.[8]

8. Our second model of money growth is similar in spirit to Canzoneri's (1985) model. There the central bank was assumed to have private information about a velocity shock; here it has private information about real output growth.

Thus, in both models of money growth the signal about the central bank's action is noisy, so reputation equilibria involve reversions ("punishments") along equilibrium outcome paths. In the first model the realized inflation rate is itself the signal, while in the second the inflation rate is the signal plus additional noise. To capture both models of money growth, the framework analyzed here allows two shocks. For the first interpretation one shock is set identically to zero.

In the next two sections we will characterize a certain class of reputation equilibria and calculate expected payoffs along the equilibrium outcome paths. These equilibria are then compared with those for the exchange-rate model. Since the signal is noiseless under an exchange-rate regime, no reversions occur along the equilibrium outcome path.

2.2 HOUSEHOLD BEHAVIOR

Under a money-growth rule the timing of events within each period is:

1. the government sets the money growth rate μ;
2. each household chooses w, interpreted as a rate of wage growth, in anticipation of the current inflation rate;
3. the signal

$$s = \mu + \varepsilon$$

and the inflation rate

$$\pi = \mu + \varepsilon + u$$

are observed, where ε and u are the exogenous shocks. In the simple model $\varepsilon \equiv v$ and $u \equiv 0$.

Let $\bar{w}$ denote the average rate of wage growth in the economy. The one-period loss for a household that sets the wage w is

$$L(w, \bar{w}, \pi) = \frac{a}{2}\pi^2 + \frac{b}{2}(\bar{w} - \pi)^2 + \frac{d}{2}(w - \alpha - \pi)^2,$$

where $\alpha > 0$, and where $a, b, d > 0$ with $(a + b + d)/2 = 1$ are relative weights.

The household's loss function has a "new Keynesian" interpretation.[9] Suppose each household is the monopolistic supplier of a differentiated commodity produced with labor as the only input. Since households set

9. See Ireland (1997) for a more detailed justification of a similar payoff function.

wages before the current inflation rate is known, wages are sticky for one period.

Suppose each household's target wage is $W = (1 + \hat{\alpha})\, P$, where P is the average price level in the economy and $\hat{\alpha}$ is the desired markup. It is convenient to renormalize units each period so $P_{-1} = 1$, and let $w = \ln W$, $\pi = \ln P$, and $\alpha = \ln(1 + \hat{\alpha})$.

The first term in the loss function represents the "shoe leather" cost of inflation. It depends only on the actual rate of inflation P/P_{-1}, and with the chosen normalization it is proportional to $[\ln(P/P_{-1})]^2 = \pi^2$.

The second term represents the household's interests as a consumer. Its surplus is maximized if other producers set wages at $\bar{W} = P$, and its relative loss is proportional to $[\ln(\bar{W}/P)]^2 = (\bar{w} - \pi)^2$.

The last term represents the household's interests as a producer. Its surplus is maximized if its wage equals the target value, and its relative loss is proportional to

$$\left[\ln\left(\frac{W}{(1 + \hat{\alpha})P}\right)\right]^2 = (w - \alpha - \pi)^2.$$

Notice that $\mu = E[\pi|\mu]$ is the expected rate of inflation, conditional on the value μ for money growth. Let μ^a denote the rate of money growth anticipated by households. Then μ^a is also the inflation rate expected by households, where the word "expected" encompasses uncertainty about the central bank's action as well as uncertainty about the shock.

Consider the expected value of the current period loss if μ^a is anticipated and μ is carried out. Households set wages at $w = \mu^a + \alpha$, so the expected loss is

$$\begin{aligned}\Lambda(\mu^a, \mu) &= E[L(\mu^a + \alpha, \mu^a + \alpha, \mu + \varepsilon + u)] \\ &= \frac{a}{2}\mu^2 - b\alpha(\mu - \mu^a) + \frac{b + d}{2}(\mu - \mu^a)^2 + M, \qquad (1)\end{aligned}$$

where

$$M \equiv \sigma_\varepsilon^2 + \sigma_u^2 + \frac{b}{2}\alpha^2$$

is an unavoidable part of the expected loss. The first two terms in Λ, which are exactly as in Barro and Gordon (1983), are important for the incentive constraints for the central bank. The third term and its derivative vanish when households correctly anticipate the action of the central bank, $\mu^a = \mu$, as they do in equilibrium. The last term, M, is important for cost comparisons across instruments.

The second term in Λ can be interpreted as a Phillips-curve coefficient. If households anticipate an average rate of inflation μ^a, then the central bank can reduce this part of the expected loss by setting the money growth rate a little higher, $\mu > \mu^a$. Of course, a higher value for μ increases the first and third terms in Λ, putting a bound on the net gain from unanticipated inflation.

In equilibrium households correctly anticipate the action of the central bank, $\mu^a = \mu$, so the expected loss is

$$\Lambda(\mu, \mu) = \frac{a}{2}\mu^2 + M.$$

Consequently, if the central bank could precommit, it would set $\mu = 0$ to minimize this loss. Call $\mu = 0$ the *Ramsey rate* of money growth. For the reasons noted above, if $\mu^a = 0$ is anticipated, short-run considerations tempt the central bank to set $\mu > 0$.

Define μ^N to be the unique rate of money growth with the property that if μ^N is anticipated by households, so they set wages at $w = \mu^N + \alpha$, then the central bank has no short-run temptation to deviate. The latter requires $\Lambda_2(\mu^N, \mu^N) = 0$, so

$$\mu^N = \frac{b\alpha}{a}.$$

Call μ^N the *Nash rate* of money growth.

Let

$$\delta \equiv \Lambda(\mu^N, \mu^N) - \Lambda(0, 0) = \frac{\alpha}{2}(\mu^N)^2 = \frac{(b\alpha)^2}{2a}$$

denote the difference between the expected losses (over one period) under the Nash and Ramsey money growth rates.

2.3 MARKOV EQUILIBRIA

The game described above is infinitely repeated, and future losses are discounted by the constant factor $\beta \in (0, 1)$ per period. If β is close to one, as we will assume here, the repeated game has many subgame-perfect equilibria. We will focus on a particular subset: Markov equilibria in which there are two states, *good* and *bad,* that also satisfy some other restrictions. In the rest of this section we will briefly describe this set of equilibria and sketch the argument for characterizing the subset that minimize expected discounted losses. A more detailed discussion is provided in Appendix A.

Each equilibrium in the class we are considering is characterized by rates of money growth (μ^g, μ^b) for the central bank and rates of wage growth (w^g, w^b) for the representative household for each state, and rules for updating the state at the end of each period. These must satisfy the usual equilibrium conditions. The additional restrictions are twofold.

First, we will focus on equilibria in which the central bank chooses the Ramsey rate of money growth in the good state, $\mu^g = 0$, and the Nash rate in the bad state, $\mu^b = \mu^N$. It then follows immediately that the rates of wage growth chosen by households are $w^g = 0 + \alpha$ and $w^b = \mu^N + \alpha$.

Second, we will restrict the class of rules for updating the state. We will assume that only the current signal s is used and that it is used in a particular way in each state. Specifically, if the economy is currently in the good state, households compare the signal with a one-sided threshold S^g, and the state remains good in the next period if and only if $s \leq S^g$. If the economy is currently in the bad state, households check whether the signal lies in a symmetric interval around μ^N, and the state reverts to good in the next period if and only if $s \in [\mu^N - \varepsilon^b, \mu^N + \varepsilon^b]$. The simple structure of these equilibria makes them appealing candidates for attention.

The pair of thresholds (S^g, ε^b) must also satisfy incentive compatibility (IC) constraints for the central bank in each state. These constraints ensure that any deviation from the equilibrium rate of money growth, 0 or μ^N, is unattractive to the bank.

DEFINITION *Simple two-state Markov equilibria* are characterized by money growth rates $\mu^g = 0$ and $\mu^b = \mu^N$, rates of wage growth $w^g = 0 + \alpha$ and $w^b = \mu^N + \alpha$, and updating rules that use only the current signal. Depending on the current state, the state next period is good if and only if $s \leq S^g$ or $s \in [\mu^N - \varepsilon^b, \mu^N + \varepsilon^b]$, where the critical values $S^g, \varepsilon^b \geq 0$ satisfy the IC constraints for the central bank.

The symmetric form of the test in the bad state ensures that the IC constraint holds in that state. The IC constraint in the good state imposes an additional restriction on the pair (S^g, ε^b). We turn next to a brief discussion of that constraint.

Instead of using S^g and ε^b, it is convenient to analyze the model in terms of the corresponding probabilities p of a reversion from the good state to the bad and q of a return in the other direction. It is also useful to place a mild restriction on the distribution of the shock ε.

ASSUMPTION 1 ε has a continuous, symmetric, unimodal density $f(\varepsilon)$ with mean zero, whose support is all of **R**.

Under Assumption 1 the reversion probability p can be adjusted continuously from 0 to ½ by adjusting S^g from 0 to $+\infty$; and the return probability q can be adjusted continuously from 0 to 1 by adjusting ε^b from 0 to $+\infty$. Normal distributions with mean zero satisfy this assumption and will be used in the examples.[10]

It is useful to define the function

$$\gamma(p) \equiv f(F^{-1}(1 - p)), \qquad p \in (0, 1),$$

where F is the c.d.f. for f. Then $\varepsilon = F^{-1}(1 - p)$ is the value for the shock that leaves probability p in the upper tail, and $f(\varepsilon)$ is the height of the density function at this point. Thus, $\gamma(\cdot)$ maps probabilities in the upper tail into levels for the density function. We will also use the hazard function, $h(p) = \gamma(p)/p$.

Fix β and define the function

$$\psi(p, q; \beta) = \frac{1}{1 - \beta(1 - p - q)}.$$

Recall that δ is the incremental expected loss from being in the bad state rather than the good in the current period. If the switching probabilities are (p, q), then $\delta\psi(p, q)$ is the expected discounted value of the (current and future) incremental losses from being (currently) in the bad state. That is, $\psi(p, q)$ takes account of all future switches back and forth between states, discounting and weighting them appropriately. Note that ψ is decreasing in p and q: higher switching probabilities reduce the difference between the states.

Fix the parameters (a, b, α) and the density f; let $\mu^N = b\alpha/a$ be the Nash inflation rate, and let γ, h be the functions defined above. The set of probabilities $(p, q) \in (0, ½] \times [0, 1]$ that satisfy the central bank's IC constraint in the good state are those for which

$$\gamma(p)\beta\delta\psi(p, q) \geq b\alpha.$$

10. If there are equilibria with $p > ½$, then there are also equilibra with $p \leq ½$, and the latter have lower costs. Hence we focus on them.

The interpretation is as follows: increasing the money growth rate above $\mu^g = 0$ leads to a marginal gain of $b\alpha$ in the current period and a marginal increase of $\gamma(p)$ in the probability of reversion to the bad state. The latter is multiplied by $\beta\delta\psi(p, q)$, the expected discounted loss if a reversion occurs. Using h instead of γ and rearranging terms, we can rewrite this constraint as

$$\beta\delta p\psi(p, q) \geq \frac{b\alpha}{h(p)}. \tag{2}$$

Suppose the pair (p, q) satisfies (2). If the economy is currently in the good state, the expected discounted cost of future reversions is $\beta\delta p\psi(p, q)$. Hence the equilibria that minimize expected discounted losses are those that solve

$$\min_{p,q \in (0,\, 1/2] \times [0,\, 1]} \quad \beta\delta p\psi(p, q) \qquad \text{s.t. (2).} \tag{3}$$

Proposition 1 characterizes the set of equilibria that minimize expected losses among all simple two-state Markov equilibria.

PROPOSITION *Let $f(\varepsilon)$ satisfy Assumption 1. Then*

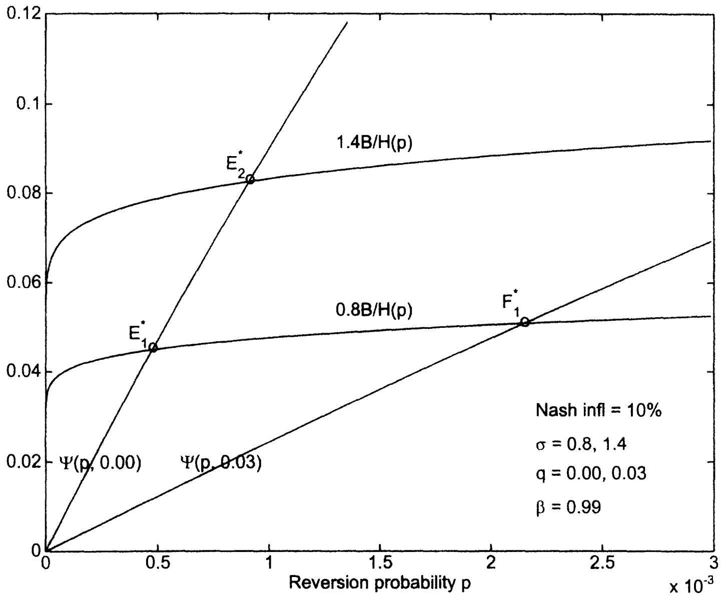
(i) *any pair $(p, q) \in (0, 1/2] \times [0, 1]$ satisfying (2) characterizes a simple two-state Markov equilibrium;*
(ii) *the set of such equilibria is nonempty if and only if (2) holds for $q = 0$, for some $p \in (0, 1/2]$;*
(iii) *a pair (p^*, q^*) attains the minimum expected loss if and only if it solves (3), and a solution exists if the set of equilibria is nonempty;*
(iv) *if (p^*, q^*) is a loss-minimizing pair and $q^* < 1$, then the expected loss per period, conditional on starting in the good state, is*

$$C = \Lambda(0, 0) + \frac{b\alpha}{h(p^*)}; \tag{4}$$

(v) *if $f(\varepsilon)$ is a normal density, then the solution (p^*, q^*) is unique (if one exists) and $q^* = 0$.*

The first and third claims summarize the discussion of (2) and (3). The second claim follows from the fact that ψ is decreasing in q. The fourth follows from the fact that if (p^*, q^*) is loss-minimizing and $q^* < 1$, then (2) holds with equality. If it did not, q^* could be increased, shortening

Figure 2 EQUILIBRIUM REVERSION PROBABILITIES



reversions and further reducing expected losses. To illustrate the last claim we turn to an example.

If $f_i(\cdot)$ is a normal $(0, \sigma_i^2)$ density, the associated hazard function is $h_i(p) = H(p)/\sigma_i$, where H is the hazard function for a normal (0, 1). The function H is decreasing for $p \leq 1/2$. (Recall that p is the probability in the *upper* tail.)

Figure 2 displays the function $\Psi(p, q; \beta) = \beta p \psi(p, q; \beta)$, for $\beta = 0.99$ and $q = 0.0, 0.03$; and the function $b\alpha/\delta h_i(p) = 2/\mu^N h_i(p)$, for $\mu^N = 10\%$ and $\sigma_i = 0.8, 1.4$. Suppose $\sigma = \sigma_1 = 0.8$. The points E_1^* and F_1^* occur where the $2/\mu^N h_1(p)$ curve crosses the $\Psi(p, q)$ curves, for $q = 0$ and 0.03. Call the x-coordinates of these points $p_1^{min}(q)$. For each q, the IC constraint (2) holds to the right of this point, so there are equilibria for $p \geq p_1^{min}(q)$. Reducing q extends the feasible range for p downward, reflecting the fact that the central bank's IC constraint involves a trade-off: longer punishments (lower q) permit less frequent punishments (lower p).

For each fixed q, the pair $(p_1^{min}(q), q)$ minimizes expected discounted costs. And since a pair of this form satisfies the IC constraint with equality, the expected discounted cost of future reversions is proportional to the quantity on the vertical axis. Hence the minimum expected loss overall is attained at $E_1^* = (p_1^*, q_1^*) = (p_1^{min}(0), 0)$.

The figure is qualitatively the same for any parameter values, provided the shock ε has a normal distribution, establishing claim (v). For $\sigma = \sigma_2 = 1.4$ the (unique) minimum-cost equilibrium occurs at the crossing point E_2^*, again with $q^* = 0$. Notice that increasing σ raises the minimum expected cost: a less informative signal requires a higher reversion probability p^*.

Figure 3a displays the optimal reversion probability p^* as a function of the standard deviation σ, for Nash inflation rates of 3%, 5%, 10%, and 20%. Looking along each curve, we see that increasing the standard deviation of the shock—reducing the accuracy of the signal—leads to more frequent reversions. Looking across curves, we see that increasing the Nash inflation rate—raising the cost of reversions—reduces the frequency of reversions. Figure 3b displays the corresponding thresholds for the inflation rate.

The conclusion that $q^* = 0$ is a direct consequence of the fact that the hazard function $h(p)$ for a normal density is a decreasing function. It holds for other distributions with that property, but not in general. For example, suppose $f(\varepsilon)$ has an exponential distribution in the relevant range,

$$f(\varepsilon) = \frac{1}{2}\eta e^{-\eta\varepsilon}, \qquad \varepsilon \geq 0.$$

Then the hazard rate is constant in the region of interest: $h(p) = \eta$, $p \leq \frac{1}{2}$. For this distribution the curve $2/\mu^N h(p)$ in Figure 2a is a horizontal line. Hence if there are any equilibria at all, there are many that attain the minimum expected cost, each with the form $(p^{\min}(q), q)$, $q \in Q^*$. These equilibria have switching probabilities that rise and fall together.

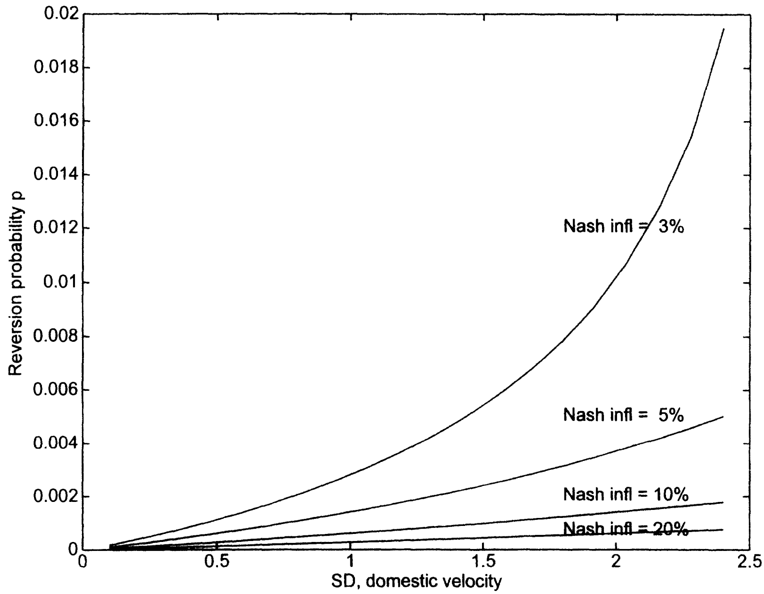
Alternatively, if f has an increasing hazard rate in the relevant range, then the cost-minimizing equilibrium is again unique and has $q = 1$.

2.4. OBSERVABILITY AND TIGHTNESS

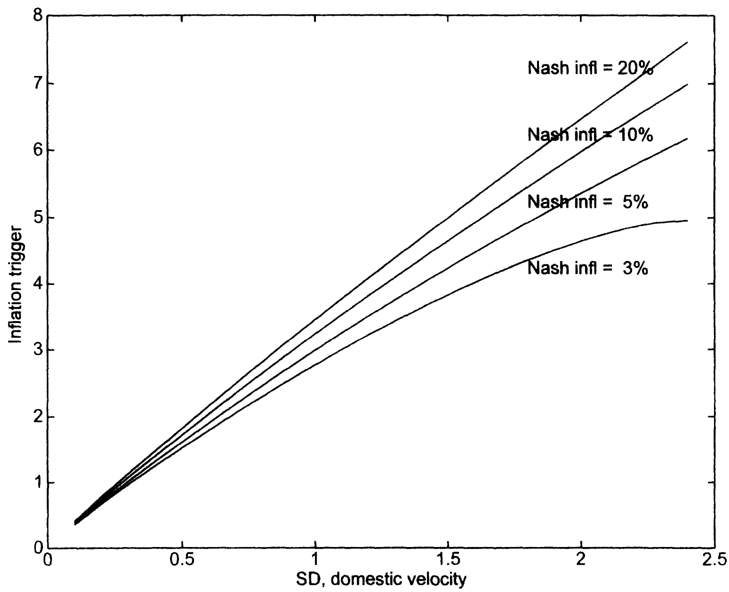
With the characterization of the least-cost equilibria in hand, we can return to the central bank's problem of choosing between the two potential policy instruments, money growth and the exchange rate. Recall from (1) that $\Lambda(0, 0) = M = \sigma_\varepsilon^2 + \sigma_u^2 + b\alpha^2/2$ is the expected loss per period, ignoring reversions. Since no reversions occur under the exchange-rate regime and there is only one shock, the expected loss per period is simply

$$C^{ex} = \sigma_\xi^2 + \frac{1}{2}b\alpha^2.$$

Figure 3 (a) OPTIMAL REVERSION PROBABILITIES; (b) OPTIMAL INFLATION THRESHOLDS



(a)



(b)

Under a money-growth regime the expected cost of reversions must also be included. Consider first the simple model of money growth. The velocity shock intervenes between the money growth rate and the signal, $s = \pi = \mu + v$, and it is the only shock. Hence $\varepsilon = v$ and $u = 0$, and Assumption 1 must hold for the velocity shock. Let h_v^* denote the hazard rate in a cost-minimizing equilibrium. Then the expected cost per period is

$$C^{mg} = \sigma_v^2 + \frac{1}{2}b\alpha^2 + \frac{b\alpha}{h_v^*}.$$

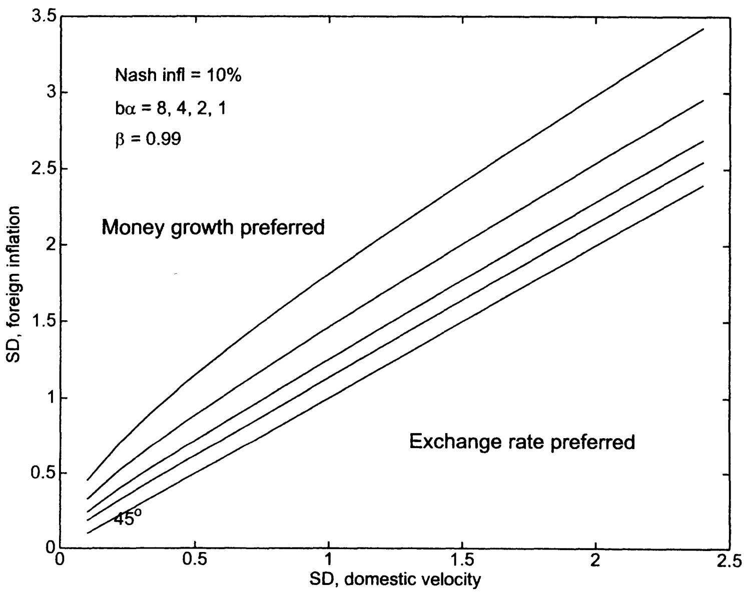
Comparing the two costs, we find that the exchange rate is preferred to money growth as an instrument if and only if

$$\sigma_\xi^2 \leq \sigma_v^2 + \frac{b\alpha}{h_v^*}.$$

If $\sigma_\xi^2 \leq \sigma_v^2$, then the exchange rate is obviously preferred: it is both tighter and more observable. If $\sigma_\xi^2 > \sigma_v^2$, then the exchange rate is the preferred instrument if and only if the higher cost from its looser relationship with the target (the higher variance of its shock) is more than offset by the expected cost of reversions under a money-growth policy.

Figure 4 displays the tradeoff for $\beta = 0.99$, a Nash inflation rate of $\mu^N =$

Figure 4 INSTRUMENT CHOICE



$b\alpha/a = 10\%$, and the four values $b\alpha = 1, 2, 4, 8$ for the Phillips-curve coefficient. [The corresponding weights on the shoe-leather cost of inflation, the first term in (1), are $a = 0.1, 0.2, 0.4, 0.8$.] The exchange rate is the preferred instrument along and below the 45° line, where the standard deviation of the foreign inflation shock is no greater than that of the domestic velocity shock. In addition it is preferred if the former is somewhat larger than the latter, with the exact position of the separating curve depending on the parameters.

A higher value for the Phillips-curve coefficient $b\alpha$ increases the central bank's incentive to deviate under a money-growth regime, increasing the size of the region where the exchange rate is preferred instrument. That coefficient measures the gain from surprise inflation, which in the model here is interpreted as arising because of monopolistic (rather than perfect) competition among producers (households). But it can have other interpretations as well. For example, it might represent the value of additional seignorage revenue, or the benefit from devaluing outstanding (nominal) debt.

For the complex model of money growth the signal is $s = \mu + \varepsilon$, where ε is the error in the bank's forecast of GDP growth, and Assumption 1 must hold for ε. Repeating the argument above and letting h_ε^* denote the optimal hazard rate, we find that

$$C^{mg} = \sigma_\varepsilon^2 + \sigma_u^2 = \frac{1}{2}b\alpha^2 + \frac{b\alpha}{h_\varepsilon^*},$$

so the exchange rate is the preferred instrument if and only if

$$\sigma_\zeta^2 - \sigma_u^2 \leq \sigma_\varepsilon^2 + \frac{b\alpha}{h_\varepsilon^*}.$$

If $\sigma_\zeta^2 \leq \sigma_\varepsilon^2 + \sigma_u^2$, then the exchange rate is obviously preferred. Otherwise there is, as before, a trade-off between tightness and observability. Figure 4 still applies, with the axes relabeled: on the horizontal axis is σ_ε, and on the vertical is $\sqrt{\sigma_\zeta^2 - \sigma_u^2}$.

With a normal distribution for ε, the optimal punishment length is infinite: $q^* = 0$. Such an outcome strains the imagination: presumably a new central banker or a new institution altogether would be put in place in finite time. It is very easy to modify the model here to deliver that result, by adding a strictly positive lower bound, $q^0 > 0$, on the return probability. The argument above proceeds exactly as before (cf. Figure 2), and the (unique) equilibrium has $q^* = q^0$. The reversion length is random, and it is straightforward to calculate its expected value as a function of q^0. Since the additional restriction operates on the money growth instrument, the result is to enlarge the region of parameter space where the exchange rate is the preferred instrument.

3. *Robustness*

Not all governments are as benevolent and clever as a Ramsey government. The possibility that the government is "bad," which may mean greedy, incompetent, or myopic, creates difficulties for a "good" (Ramsey) government. Some of the difficulties are unavoidable: a legacy of large outstanding debt, bad legislation, etc. can be difficult to undo. In addition, the behavior of the private sector will be predicated on a certain apprehension about the nature of the administration currently in power. In this section we will show that if a "good" government cannot easily distinguish itself from a "bad" one, this mistrust by the private sector makes its task more onerous. In such an environment a simple policy rule can be very useful, even if it cannot respond to shocks in the environment. In the model here, the fact that a rule reduces or eliminates the potential damage done by a "bad" government has a very useful effect on private behavior. This effect far outweighs the small additional gain that a "good" government could attain with discretion. As will be shown, even a moderate probability that the government is "bad" makes the rule worthwhile.

Suppose that there are two types of governments, *Ramsey* and "bad." Reputation equilibria are delicate, and there are countless ways for the other type of government to deviate from the Ramsey policy. Here we will assume that the "bad" type is myopic, setting current tax rates to maximize current-period utility. The Ramsey government behaves in the usual fashion, raising revenue in a way that maximizes the expected discounted utility of the representative household. For simplicity, we will assume that the government's type is i.i.d.

The environment is adopted from Fischer's (1980) paper. Each household receives an endowment of goods that can be invested or consumed directly. Invested goods earn a return but are also subject to taxation. The household can also use labor to produce goods. The government must finance an exogenous expenditure sequence. The tension is between the government's short-run temptation to use a (nondistorting) capital levy to finance current expenditures, and the adverse effect such a policy has on the incentive to invest. The expenditure sequence is stochastic, and for simplicity is taken to be i.i.d.

First we study a setting where policy is discretionary. If the government is known to be the Ramsey type and the discount factor is sufficiently close to one, then the standard reputation argument applies. In the setting here, the Ramsey government uses a carefully calculated capital tax to finance part of spending and to provide insurance against the high expenditure shock. The capital tax varies with the expenditure level, but its expected value is low enough so that investment is worthwhile.

If the type of government is uncertain, but the probability of the myopic type is not too high, a reputation equilibrium still exists. The policy of the Ramsey government is qualitatively similar to the previous case. The main difference is that the Ramsey government must offer a high enough expected return on capital during the periods when it is in office to compensate the household for the fact that capital earns a negative expected return when the myopic government is in power. The policy adopted by the Ramsey government becomes rather odd, and expected utility declines as the probability of a myopic government increases. The Ramsey government is willing to continue participating in this equilibrium because abandoning it means that households stop investing, which entails a substantial cost. (For sufficiently high probabilities reputation equilibria cease to exist, but here we will focus on probabilities that are below that threshold.)

We then consider what happens if, instead of allowing the government discretion in setting fiscal policy each period, the society adopts a policy rule placing an upper bound on the capital tax. If the probability of the Ramsey type is sufficiently close to one, this rule reduces welfare, since the insurance feature of a variable capital tax is lost. But if the probability of the myopic type is high enough, the rule is welfare-enhancing.

3.1 THE ENVIRONMENT

Each period the household receives an endowment of goods, ω, and an endowment of time. It can invest all or part of its goods endowment in a productive activity, and it can hide the rest. Let $\theta \in [0, 1]$ denote the fraction of the goods endowment that is invested. Investments earn a rate of return $r > 0$, but they can also be taxed. Hidden goods earn no return but cannot be taxed. Time spent working produces goods according to the linear technology $q = w\ell$, where $w > 0$ is an implicit wage rate and ℓ is labor supply.

Households value private consumption goods c and time worked ℓ according to a utility function that is additively separable and linear in labor supply:

$$U = \mathrm{E}\left\{\sum_{t=0}^{\infty} \beta^t [u(c(t)) - \ell(t)]\right\}.$$

Assume u is strictly increasing, strictly concave, and twice differentiable, and $0 < \beta < 1$. Assume $u'((1 + r)\omega) \gg w$, so that the household chooses to work even if it is consuming its entire endowment, with interest, and faces a positive tax on labor income.

Government expenditure is exogenous and stochastic. For simplicity assume it takes only two values, $g_1 = 0$ and $g_2 = g > 0$, and that the realizations are i.i.d. Let $\pi_1 = \pi$ and $\pi_2 = 1 - \pi$ denote the probabilities.

In each period the government levies flat-rate taxes $\tau_k \in [0, 1 + r]$, $\tau_\ell \in [0, 1]$ on capital and on labor income. The government cannot issue debt, so its budget must be balanced each period. Assume

$$r\omega < (1 - \pi)g, \tag{5}$$

so that the required revenue cannot be raised with a capital tax that leaves the household with a positive expected return on investment. For simplicity assume in addition that $g < (1 + r)\omega$, so that the required revenue *can* be raised with a confiscatory capital tax. Finally, assume that g is small enough so that it can be financed entirely with a labor tax when the household hides its good endowment.

3.2 RAMSEY GOVERNMENT

First consider an economy in which it is known for sure that the government is the Ramsey type. We are interested in settings where there is a reputation equilibrium of the usual form. The tax policy in that equilibrium is the one that the Ramsey government would employ if it could commit ex ante to fixed, state-contingent tax rates. For discount factors β that are sufficiently close to unity there an equilibrium of this form, supported by the threat of a reversion to the one-shot Nash equilibrium.

Suppose the government could precommit, and consider the problem of choosing the optimal tax policy subject to the constraints imposed by household behavior. In this stationary environment with i.i.d. expenditure shocks and no state variables, the solution is a stationary tax policy $\{(\tau_{\ell i}, \tau_{ki}), i = 1, 2\}$ that maximizes the household's expected utility per period, where subscripts $i = 1, 2$ denote the values of the tax rates, consumption, etc. in the two states.

Suppose that the household has invested all of its endowment, and consider its problem after the state i has been realized and the current tax rates $(\tau_{ki}, \tau_{\ell i})$ are known. Its problem is

$$\max_{c_i, \ell_i}[u(c_i) - \ell_i] \tag{6}$$

$$\text{s.t.} \quad c_i = (1 + r - \tau_{ki})\omega + (1 - \tau_{\ell i})w\ell_i, \quad i = 1, 2. \tag{7}$$

The equilibrium allocation must also satisfy the market-clearing condition for goods:

$$c_i + g_i = w\ell_i + (1 + r)\omega, \qquad i = 1, 2; \tag{8}$$

and the government's budget constraint (redundant, by Walras' law) must hold:

$$g_i = \tau_{ki}\omega + \tau_{\ell i}w\ell_i, \qquad i = 1, 2.$$

Finally, notice that the household's net income gain from investment in state i is $(r - \tau_{ki})\omega$. The household is willing to invest its endowment if and only if the associated change in expected utility is positive. Hence investment occurs if and only if the capital tax satisfies the rate of return constraint

$$\sum_i \pi_i u'(c_i)\omega(r - \tau_{ki}) \geq 0. \tag{9}$$

The Ramsey government's problem is

$$\max \sum_i \pi_i[u(c_i) - \ell_i]$$

subject to (8), (9), and the constraints imposed by household optimization. As shown in Appendix A, the solution $\{\theta^R, (c_i^R, \ell_i^R, \tau_{ki}^R, \tau_{\ell i}^R), i = 1, 2\}$, with $\theta^R = 1$, has the following features:

(i) consumption is the same in the two states, $c_1^R = c_2^R$;
(ii) the labor tax is the same in the two states, $\tau_{\ell 1}^R = \tau_{\ell 2}^R$;
(iii) labor supply is higher by g in the second state, $\ell_2^R = \ell_1^R + g$;
(iv) the expected capital tax is equal to the rate of return, $\sum_i \pi_i \tau_{ki}^R = r$;
(v) capital is subsidized when spending is low and taxed when it is high, $\tau_{k_1}^R < 0 < \tau_{k_2}^R$.

Features (i)–(iii) follow from the assumption that utility is linear in labor supply. Given (i), result (iv) is an immediate consequence of the rate-of-return restriction in (9). Result (v) is an instance of the principle developed in Zhu (1992) and in Chari, Christiano, and Kehoe (1994): the capital tax in a stochastic setting can act as a perfect substitute for state-contingent debt of the type discussed in Lucas and Stokey (1983).

As in the monetary model of the previous section, the Ramsey policy can be sustained as the outcome in a reputation equilibrium in which the behavior of the government is disciplined by the threat of reversion to the repeated one-shot Nash equilibrium. In the latter equilibrium households have no incentive to invest. They hoard their goods endowment and all spending is financed with contemporaneous labor taxes. If spending is low the labor tax rate is zero, $\tau_{\ell 1}^{N} = 0$. If spending is high the labor tax $\tau_{\ell 2}^{N} > 0$ is set at the minimum level needed to raise the required revenue g. Any capital tax policy that violates (9) can be used, but no revenue is collected from it.

Notice that there are temptations to deviate in both states of the world. In the low-spending state there is a one-time gain from setting both tax rates to zero, and in the high-spending state there is a one-time gain from using a large capital levy. But for β sufficiently close to one the Ramsey government resists both temptations.

3.3 MIXED TYPES ($\lambda > 0$)

The equilibrium described above is valid for an economy in which it is known with certainty that the government in office is a Ramsey government. Suppose instead that the government's type is i.i.d., and let λ be the probability of the myopic type. Let *mx* and *Rx* denote values under myopic and Ramsey governments respectively in this mixed environment. We will look at equilibria in which households still have an incentive to invest in the mixed economy, so $\theta^{x} = 1$. If λ is not too large and β is sufficiently close to one, such equilibria exist.

The behavior of the myopic government is straightforward. In the low-spending state it sets both tax rates to zero, $\tau_{k_1}^{mx} = \tau_{\ell 1}^{mx} = 0$; and in the high-spending state it raises all of the required revenue from a capital tax, setting the labor tax to zero: $\tau_{k2}^{mx} = g/\omega$ and $\tau_{\ell 2}^{mx} = 0$. The household's problem is as in (6)–(7). Since the labor tax is the same in both states, it follows immediately that consumption is the same in both states: $c_1^{mx} = c_2^{mx} = c^{mx}$. The labor supplies in the two states are then determined by (8).

In a world with a positive (but small enough) probability of a myopic government, the Ramsey government must alter its strategy, since otherwise households will not be willing to invest. Conditional on the myopic type holding office, the capital tax is $\tau_{k2}^{mx} = g/\omega$ with probability $1 - \pi$ and zero otherwise. Hence a household faces an expected utility loss of

$$L \equiv u'(c^{mx})\left\{\left[(1 - \pi)\frac{g}{\omega} - r\right]\omega\right\} > 0$$

if it invests its entire endowment. The assumption in (5) implies that the term in braces is positive. The Ramsey type must offset this loss by offering an expected gain when it is in office.

In particular, the Ramsey type must raise the subsidy on capital in the low-spending state and/or cut the capital tax in the high-spending state so that, averaging over both types of government, the household faces a nonnegative expected rate of return. Thus, in the mixed economy with probability λ of a myopic government, the rate-of-return constraint for the Ramsey government is

$$\sum_i \pi_i u'(c_i^R)\omega(r - \tau_{ki}^R) \geq \frac{\lambda}{1-\lambda} L. \tag{10}$$

For $\lambda = 0$ this inequality reduces to the one in (9), but for $\lambda > 0$ the right side is positive and increasing in λ.

The problem of the Ramsey government in the mixed economy is as before, with (10) in place of (9). As shown in Appendix B, for any fixed $\lambda > 0$ the solution $\{\theta^{Rx}, (c_i^{Rx}, \ell_i^{Rx}, \tau_{ki}^{Rx}, \tau_{\ell i}^{Rx}), i = 1, 2\}$ retains many of the qualitative features of the solution for $\lambda = 0$. Properties (i)–(iii) are unchanged: consumption and the labor tax are the same across the two states, and labor supply is higher by g in the second state. The analogue of property (iv) says that (10) holds with equality. Property (v) continues to hold if λ is not too large. In principle, however, the Ramsey type might subsidize capital in both states if λ is large enough.

Changes in the probability of a myopic administration affect the allocation under the Ramsey government as one would expect: consumption c^{Rx} is decreasing in λ; the labor tax τ_ℓ^{Rx} is increasing in λ; and both capital taxes τ_{ki}^{Rx} are decreasing in λ. That is, the subsidy on capital in the low-spending state is larger, and the tax on capital in the high-spending state is smaller. Expected utility, conditional on a Ramsey government being in office, is decreasing in λ.

As λ rises, the Ramsey government must increase the distorting labor tax to subsidize capital more heavily when spending is low and to finance a greater share of expenditure when spending is high. These costs are endured because there is a substantial gain to maintaining the incentives to invest.

3.4 A POLICY RULE

Alternatively, society could adopt a simple policy rule mandating a cap on the capital tax that is low enough to ensure that households have an incentive to invest. In our simple model the optimal cap is $\bar{\tau}_k = r/(1 - \pi)$. Both the Ramsey and myopic types use the same policy under the rule. In the low-spending state tax rates are zero, $\tau^u_{k1} = \tau^u_{\ell 1} = 0$. In the high-spending state the capital tax is set at the mandated maximum, $\tau^u_{k2} = \bar{\tau}_k$, and the labor tax $\tau^u_{\ell 2} > 0$ at the lowest rate consistent with budget balance. Expected utility under this policy rule is not as high as under the Ramsey policy, but the rule is robust against the blunders of the myopic government.

3.5 AN EXAMPLE

In this section we will look at a simple numerical example that illustrates an important point: the difference in expected utility under the reputation equilibrium compared with the policy rule is quite modest, even if the government is certain to be the Ramsey type. In addition, expected utility in the reputation equilibrium declines as the probability of the myopic type rises, and eventually the policy rule dominates.By contrast, the expected utility gain from using the rule rather than enduring the one-shot Nash outcome is very substantial. This result reflects the fact that the rule was deliberately constructed to exploit a large potential gain, ignoring small ones.

Utility is logarithmic, $u(c) = a \ln c$, and the parameter values are

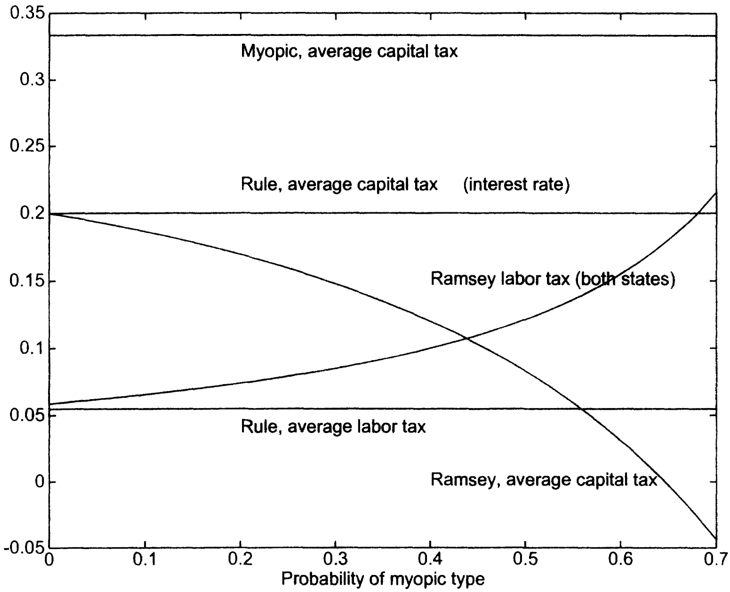
$$a = 10, \qquad w = 1, \qquad \omega = 3, \qquad r = 0.2, \qquad g = 2, \qquad \pi = 1/2.$$

The discount factor β is assumed to be sufficiently close to one so that the reputation equilibrium exists.

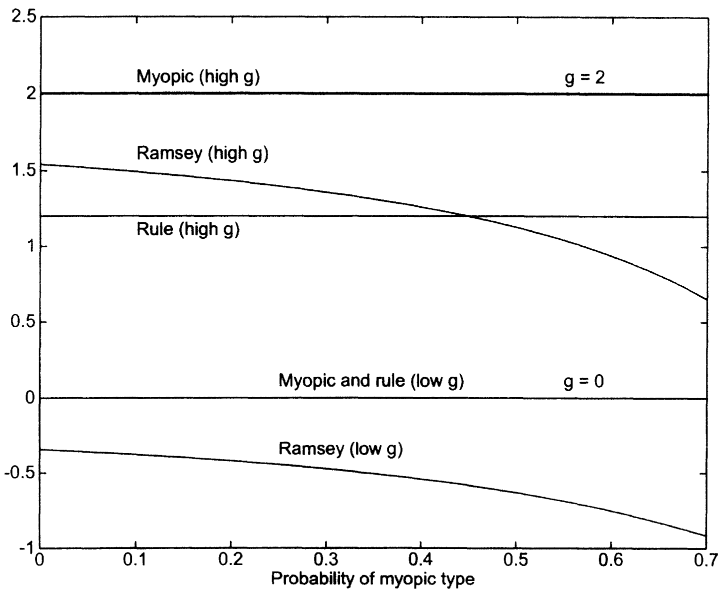
Figure 5a–d displays the equilibrium outcomes as the probability of the myopic type increases from 0% to 70%. Obviously, nothing happens to the policies or outcomes under the myopic type or under the policy rule. What do change are the policies adopted by the Ramsey type and the weighted averages in the economy with mixed types.

Figure 5a displays the tax rates. Under the myopic type the average capital tax rate is 33% (an average of 67% and 0%), well above the 20% rate of return on capital. The labor tax is zero. Under the Ramsey type the average capital tax is 20% (an average of 51% and −11%) with $\lambda = 0$ and declines monotonically as λ rises. The labor tax is positive and increases with λ, offsetting the declining revenues from the capital tax. The reason for this pattern is clear: the Ramsey type adjusts its policy to maintain the incentive for households to invest. Under the policy rule the aver-

Figure 5 (a) TAX RATES; (b) CAPITAL-TAX REVENUE; (c) CONSUMPTION; (d) EXPECTED UTILITY

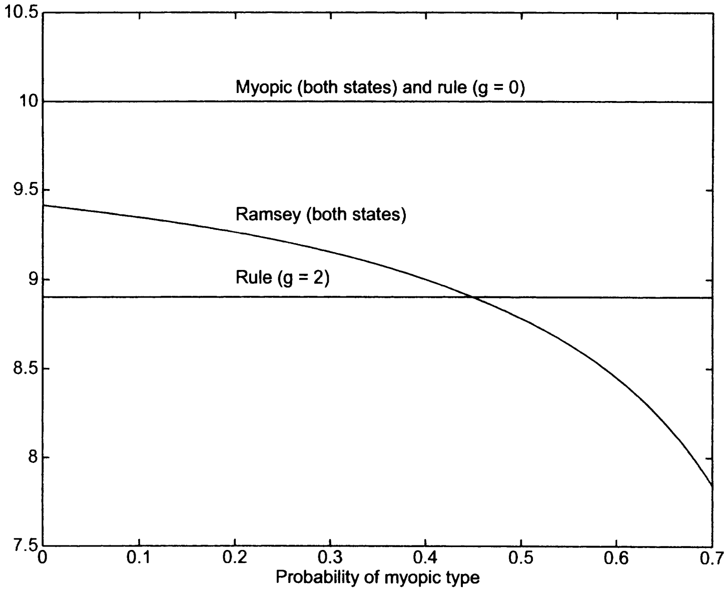


(a)

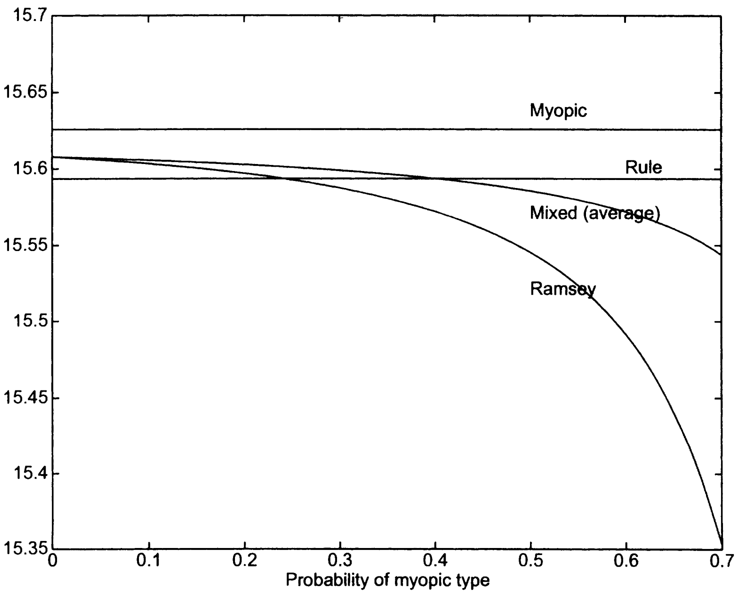


(b)

Figure 5 CONTINUED



(c)



(d)

age capital tax is 20% (an average of 40% and 0%), and the average labor tax is a little over 5%.

Figure 5b displays revenue from the capital tax. Recall that government expenditure is 2 or 0. Under the myopic type revenue from the capital tax exactly covers spending: it is 2 or 0, depending on the state, and the labor tax is not used. Under the Ramsey type, if $\lambda = 0$ revenue from the capital tax is 1.55 or -0.35, depending on the state. As λ increases, both figures decline (the subsidy in the zero-spending state gets larger). Under the policy rule the revenue from the capital tax is $\bar{\tau}_k \omega = r\omega/(1 - \pi) =$ 1.2 or 0, and the labor tax is used when spending is high.

Figure 5c displays consumption. Consumption is the same in both states under the myopic or Ramsey types, since each type sets the same labor tax in both states. Consumption falls rather sharply under the Ramsey government as λ rises. This change is a direct consequence of the rising labor tax. Under the policy rule consumption differs in the two states, since the labor tax varies.

Figure 5d displays expected utility under the myopic and Ramsey governments, as well as the weighted average, and under the policy rule. The rule delivers higher expected utility if $\lambda > 40\%$.

The figures for the one-shot Nash equilibrium are not displayed, since they are—literally—off the charts. Households do not invest, so there is no interest income and all revenue must be raised from the labor tax. When $g = 0$, labor supply is 7 and consumption is 10. When $g = 2$, the labor tax is 40%, labor supply is 5, and consumption is 10. The expected utility is 14.5. This dismal outcome deters the Ramsey type from abandoning the reputation equilibrium for reasonable β-values.

This simple model illustrates several points. The first is quantitative. A policy rule that is simple but well designed can capture much of the benefit available from commitment. Here the first-order effect comes from maintaining the incentive to invest, as can be seen by comparing expected utility under the Nash regime and under the policy rule. The simple rule cannot capture the further gains available from implicit insurance, but these are much smaller. Indeed, they vanish altogether if λ is large enough.

In addition, the behavior of the Ramsey government in this simple model suggests that the political-economy issues surrounding the reputational equilibrium cannot be neglected. Running for election on the Ramsey platform in this economy would be a difficult task indeed!

Finally, note that the damage a "bad" government can inflict is much larger if capital is long-lived. To keep the model here simple, capital was assumed to last for only one period. If capital is durable and expensive, "bad" government behavior may have much worse consequences.

4. Conclusion

To conclude it is useful to touch on some issues that the two formal models do not address. We begin with issues related to the monetary model.

As noted above, and as many authors have emphasized, models like the one analyzed here have a vast multiplicity of equilibria. We compared monetary instruments by looking at the best equilibrium within a certain class. But why should we suppose that the best equilibrium is likely to arise? In addressing this very practical question, it is useful to keep in mind that many of the equilibria in these games have similar outcome paths. In particular, there are many equilibria in which the bank plays the Ramsey strategy as long as its reputation is intact.

These equilibria differ in their description of how the bank loses its reputation, resulting in a reversion to a bad outcome, and in the precise description of the nature of the reversion. Here we assumed that one-shot Nash behavior prevailed during reversions and that the end of a reversion episode was linked to an observation of the signal, but neither feature is critical. For example, the most severe punishment could be used instead of one-shot Nash. And even if one-shot Nash is used, the reversions could be of fixed length, of completely random length, allow returns to the good state as a complicated function of current and past signals, etc. Indeed, the return probability could be interpreted as the (random) length of time required to reorganize the central bank or to install a new head of the bank in office. As an empirical matter it would be very difficult to distinguish sharply among these equilibria. They differ only in their descriptions of reversion behavior, and reversions are (necessarily) rare.

More importantly, the model's description of behavior during a reversion episode seems better taken with several grains of salt. During good times the central bank's behavior is stable and predictable. This is also roughly true in practice, and the model captures this behavior quite well. Reversions are not so precisely scripted, in reality or in the model. Choosing among reputation equilibria (on a theoretical level) and distinguishing among them (empirically) are equally difficult tasks. But they are also unimportant tasks, in the sense that the important aspect of behavior in the model is the robust feature shared by all the reputation equilibria, behavior during good times.

Having argued that choosing among reputation equilibria is not terribly important, there remains the issue of how reputations are established. Formal models that permit reputation equilibria always have a large multiplicity of other equilibria as well. Indeed, simply repeating the bad outcome is one possibility. As an empirical matter, countries with stable governments often manage to build and maintain reputations for good behavior in areas where it matters most: the public debt is honored; capital taxes are stable and not too high; intellectual-property rights are protected by patents, copyrights, and trademarks; and monetary policy is fairly stable. Reputations are central in explaining good outcomes in settings like these, where the policymaker has substantial discretion, but on a theoretical level little is understood about how reputations are built, how credibility is established.[11]

Since theory provides little or no guidance here, it may be more fruitful to view this as an empirical issue. Perhaps this is the role of the central banker (an individual) as opposed to the central bank (an institution). A successful central banker is one who can steer the economy toward a good equilibrium. Success requires that the central bank take the appropriate actions, but that is not enough. The central banker must convince the private sector that the bank will behave that way. Indeed, he (or she) must persuade the private sector that there is a commonly held belief that the bank will behave appropriately. Perhaps "leadership" is the name we give to the elusive qualities that enable some individuals to succeed at this task.

Initially establishing a reputation for good behavior is a critical task for a central banker. Adopting a more observable instrument for conducting policy, pegging an exchange rate rather than using the money growth rate, may ease the banker's task during the critical initial phase when he is attempting to establish a reputation for good behavior. Establishing a currency board is another way to accomplish the same task, in the sense that it acts as an easy-to-monitor instrument for conducting monetary policy.[12]

Of course, in the long run monetary and fiscal policy are linked through the government's budget constraint. Good monetary policy is simply infeasible without a conservative (balanced budget) fiscal policy. A government that runs substantial deficits, with no prospect of surpluses to retire the accumulating debt, will eventually fail in its efforts to float new bond issues. The problem is exacerbated if, as is typically the case, old debt must be rolled over as well. At some point the only feasible options are outright default, a large devaluation, or both. A government facing that situation typically finds the seignorage revenue from a large devaluation too attractive to resist, and monetary policy becomes the fiscal policy of last resort.[13]

11. See Faust and Svensson (2001) for an interesting exception.
12. Rogoff (1985) suggests an intriguing solution: simply appointing a central banker who places more weight on price stability.
13. See Zarazaga (1995) for a very interesting model in which episodic bouts of high inflation occur when decentralized fiscal policy is combined with centralized monetary policy.

The model analyzed here focuses on an issue that is critical for some central banks: those for which establishing and maintaining a reputation is a first priority. After that task has been largely accomplished, as it has been in the United States, in Japan, and elsewhere, other issues take center stage: which targets should be used, which monetary aggregates should be given the greatest attention, etc. These are important issues, but only after a reputation for good conduct has been fairly well established. If a bridge is in danger of collapsing, there is little point in repairing potholes. Only after the structural problems have been addressed is it useful to think about the quality of the road surface.

The model of fiscal policy analyzed above illustrates one danger from allowing too much discretion. The myopic government in the model can be thought of as representing administrations subject to a variety of short-run political pressures, arising from many possible sources. The model highlights the fact that a well-designed policy rule is one that pays attention to first-order effects (here, the incentive to invest), although it may neglect more subtle issues (here, the insurance available from a more subtle capital tax).

The model here has a representative household, but the same issue arises when there is heterogeneity. Differences among households create some divergence in views about fiscal policy, but if those differences are modest, there may still be a fair amount of common ground. There may be a set of policy rules that are advantageous to all, even if there is disagreement about the *optimum optimorum.*

The essence of good government is to design institutions that permit solution of the repeated moral-hazard problem. The goal of the models here has been to provide some insight into that problem, and how it affects decisions about policy regimes.

Appendix A

The set of simple two-state Markov equilibria for the money growth model is described in detail here. First the incentive compatibility (IC) constraints for the central bank are derived. Then the equilibria that minimize total expected discounted costs are characterized.

Suppose the economy is in the good state. Then households expect the central bank to permit money growth (net of real output growth) at the Ramsey rate, $\mu = \mu^g = 0$, so they set wages at $w^g = 0 + \alpha$. Households then observe the signal s and the actual inflation rate π. Households use a one-sided threshold to decide if the bank has deviated. If the signal lies below the threshold, $s \leq S^g$, they assume that the central bank has behaved as anticipated, and the state next period is good. Otherwise they assume the central bank has deviated, and the state next period is bad. In equilibrium the central bank sets $\mu = \mu^g = 0$, so $s \leq S^g$ if and only if $\varepsilon \leq S^g$.

When the economy is in the bad state, households expect the central bank to set money growth at the Nash rate, $\mu = \mu^b = \mu^N$, so they set wages at $w^b = \mu^N + \alpha$. They then observe s and π. Households use a two-sided test in the bad state, so the state next period is good if and only if the signal lies within the tolerance level set by the test, $s \in [\mu^N - \varepsilon^b, \mu^N + \varepsilon^b]$. Since the central bank sets money growth at the Nash rate μ^N, the signal s lies in the acceptable region if and only if $\varepsilon \in [-\varepsilon^b, \varepsilon^b]$.

In a two-state Markov equilibrium the expected discounted value of current and future losses from any period on depends only on the current state. Let c^g and c^b denote those expected values, and let

$$\Delta \equiv c^b - c^g$$

denote the difference between the two.

Fix $S^g \geq 0$, and suppose that the economy is in the good state. The Ramsey rate of money growth $\mu^g = 0$ is incentive-compatible for the central bank if and only if any other growth rate $\mu \neq 0$ leads to a (weakly) greater sum of current and future expected losses. If money grows at the rate μ, then the expected loss in the current period is $\Lambda(0, \mu)$. The state next period, and hence the payoffs from next period on, depend only on the signal s that is observed today. If money grows at the rate at μ and the signal s is observed, the error term is $\varepsilon = s - \mu$. Therefore, if S^g is the threshold for the accept region, the expected cost from next period on is

$$\begin{aligned} &c^g \quad \text{if } \varepsilon \leq S^g - \mu, \\ &c^b \quad \text{if } \varepsilon > S^g - \mu. \end{aligned}$$

Hence the Ramsey growth rate $\mu = 0$ is incentive compatible for the central bank in the good state if and only if

$$\begin{aligned} &\Lambda(0, \mu) + \beta\{F(S^g - \mu)c^g + [1 - F(S^g - \mu)]c^b\} \\ &\quad \geq \Lambda(0, 0) + \beta\{F(S^g - 0)c^g + [1 - F(S^g - 0)]c^b\}, \qquad \text{all } \mu. \end{aligned}$$

Rearranging terms and using the definition of Δ, we can write this constraint as

$$\begin{aligned} \beta\Delta[F(S^g) - F(S^g - \mu)] &\geq \Lambda(0, 0) - \Lambda(0, \mu) \\ &= b\alpha\mu - \mu^2, \qquad \text{all } \mu. \end{aligned} \tag{11}$$

The expression on the left side of (11) is zero at $\mu = 0$. Under Assumption 1 it is continuous and increasing in μ, convex for $\mu < S^g$, and concave for $\mu > S^g$. The expression on the right side is also zero at $\mu = 0$, and from (1) we see that it is increasing for $\mu < b\alpha$, decreasing for $\mu > b\alpha$, and everywhere concave. Hence (11) holds near $\mu = 0$ if and only if

$$\beta\Delta f(S^g) \geq b\alpha. \tag{12}$$

Condition (12) is the basic IC constraint that equilibria must satisfy. The interpretation is straightforward: $b\alpha$ is the marginal gain from increasing the expected inflation rate in the current period, $f(S^g)$ is the marginal increase in the probability of reversion to the bad state from that change, and $\beta\Delta$ is the discounted expected loss if a reversion occurs.

Similarly, the Nash rate of money growth, $\mu^b = \mu^N$, is incentive-compatible for the central bank in the bad state if and only if any other growth rate $\mu \neq \mu^N$ leads to (weakly) greater expected losses. Hence the required condition is

$$\beta\Delta\{[F(\varepsilon^b) - F(-\varepsilon^b)] - [F((\mu^N - \mu) + \varepsilon^b) - F((\mu^N - \mu) - \varepsilon^b)]\} \\ \geq \Lambda(\mu^N, \mu^N) - \Lambda(\mu^N, \mu), \qquad \text{all } \mu.$$

Under Assumption 1 the left-hand side is strictly positive for all $\mu \neq \mu^N$, and it follows immediately from the definition of μ^N that the right-hand side is negative for all $\mu \neq \mu^N$. Hence any value $\varepsilon^b \geq 0$ satisfies the incentive constraint for the bad state.

For an equilibrium characterized by (S^g, ε^b), the probability of a reversion to the bad state (along the equilibrium outcome path) is $p = 1 - F(S^g)$, and the probability of a return to the good state is $q = F(\varepsilon^b) - F(-\varepsilon^b)$. Under Assumption 1 the relationship between the thresholds (S^g, ε^b) and the probabilities (p, q) is invertible, so we can formulate the problem in terms of the latter. Thus, the next step is to solve for Δ in terms of the probabilities (p, q).

Suppose (p, q) is an equilibrium pair. In equilibrium, the money growth rate in the good state is $\mu^g = 0$, and the probability of an accidental reversion is p. Hence the expected discounted value of current and future losses satisfies the recursive relation

$$c^g = \Lambda(0, 0) + \beta[c^g + p\Delta]. \tag{13}$$

Similarly, the money growth rate in the bad state is $\mu^b = \mu^N$, and $1 - q$ is the probability of remaining in the bad state. Hence c^b satisfies the recursive relation

$$c^b = \Lambda(\mu^N, \mu^N) + \beta[c^g + (1 - q)\Delta].$$

The difference between the two is

$$\begin{aligned}\Delta(p, q) &= \frac{\delta}{1 - \beta(1 - p - q)} \\ &= \delta\psi(p, q),\end{aligned} \tag{14}$$

where $\delta = \Lambda(\mu^N, \mu^N) - \Lambda(0, 0)$.

Since $f(S^g) = \gamma(p)$, it then follows from (12) that the Ramsey rate of money growth is incentive-compatible in the good state for any $p, q \in [0,1]$ satisfying

$$\gamma(p)\beta\delta\psi(p, q) \geq b\alpha. \tag{15}$$

Since f is symmetric, so is γ. Hence if (15) holds for $\hat{p} \geq 1/2$, it also holds for $p = 1 - \hat{p} \leq 1/2$, and we can limit attention to p-values in the upper half of the distribution, $p \in (0, 1/2]$.

Hence if the density $f(\varepsilon)$ satisfies Assumption 1, there exists a simple two-state Markov equilibrium for any pair $p \in (0, 1/2]$ and $q \in [0, 1]$ satisfying (15). Since $\psi(p, q)$ is decreasing in q, for fixed p, if (15) holds anywhere, it holds for $q = 0$, establishing parts (a) and (b) of the proposition.

Substituting from (14) into (13), we find that expected cost per period in the (p, q) equilibrium is

$$(1 - \beta)c^g(p, q) = \Lambda(0, 0) + \beta p\delta\psi(p, q).$$

Hence expected costs are minimized if and only if (p, q) solves (3), establishing part (c) of the proposition.

Appendix B

Ramsey behavior for the tax model is characterized here. The cases $\lambda = 0$ and $\lambda > 0$ are similar and can be treated together.

It is convenient first to reformulate the problem as one of choosing the allocation $\{(c_i, \ell_i), i = 1, 2\}$ and the capital tax rates $\{\tau_{ki}, i = 1, 2\}$. To this end, note that the condition for the consumer's maximum is

$$(1 - \tau_{\ell i})wu'(c_i) - 1 = 0, \qquad i = 1, 2. \tag{16}$$

Multiply the budget constraint (7) by $u'(c_i)$ and substitute from (16) to get

$$u'(c_i)[c_i - (1 + r - \tau_{ki})\omega] - \ell_i = 0, \qquad i = 1, 2. \tag{17}$$

The constraints for the Ramsey government's problem are (8), (9) or (10), and (17). Let $\pi_i\mu_i$, ϕ, and $\pi_i\lambda_i$ be the multipliers for the three constraints. Then the conditions for a maximum are

$$0 = u_i' + \lambda_i\{u_i'' \cdot [c_i - (1 + r - \tau_{ki})\omega] + u_i'\} - \mu_i - \phi u_i''\omega[\tau_{ki} - r],$$

$$0 = -1 - \lambda_i + \mu_i w,$$

$$0 = (\lambda_i - \phi)u_i'\omega,$$

$i = 1, 2$. The last two equations imply

$$\lambda_i = \phi \text{ and } \mu_i = \frac{1 + \phi}{w}, \qquad i = 1, 2.$$

Then substituting into the first equation gives

$$(1 + \phi)\left(u_i' - \frac{1}{w}\right) + \phi(c_i - \omega)u_i'' = 0, \qquad i = 1, 2,$$

which suggests a solution with the same private consumption level in the two states, $c_1 = c_2 = c^R$. The remaining task is to find values for $(c^R, \{\ell_i^R, \tau_{ki}^R, i = 1, 2\})$ that satisfy the constraints, (8), (9) or (10), and (17).

Use the market-clearing condition (8) to write labor supply in the two states as

$$\ell_i^R = \frac{1}{w}[c^R - (1 + r)\omega + g_i], \qquad i = 1, 2. \tag{18}$$

Then use this fact and the budget constraint (17) to obtain

$$\tau_{ki}^R\omega u' = (1 - wu')\ell_i^R + g_i u', \qquad i = 1, 2. \tag{19}$$

The rate of return constraint (9) or (10) holds with equality.

Since consumption is the same in the two states, the rate of return constraint is

$$u'(c^{Rx})\omega\left(r - \sum_i \tau_{ki}^{Rx}\right) \geq \frac{\lambda}{1-\lambda}L.$$

Notice that since consumption is different under the two types of government, expected returns must be weighted by marginal utilities as well as probabilities.

Hence

$$\omega u' \sum_i \pi_i \tau_{ki}^R = \omega u' r - \frac{\lambda}{1-\lambda}L. \tag{20}$$

Taking the probability-weighted sum of the two equations in (19), substituting from (20), and using the fact that $\ell_2 = \ell_1 + g/w$, we obtain

$$r\omega u' - \Lambda = (1 - wu')\ell_1^R + \pi_2 \frac{g}{w}.$$

Using (18), we find that c^R satisfies

$$r\omega u' - \Lambda = \left(\frac{1}{w} - u'\right)[c^R - (1+r)\omega] + \pi_2 \frac{g}{w}.$$

The labor supplies ℓ_1^R, ℓ_2^R can then be determined from (18), the capital tax rates from (19), and the labor tax—which is the same in both states—from (16).

REFERENCES

Atkeson, A., and P. J. Kehoe. (2001). The advantage of transparent instruments of monetary policy. Federal Reserve Bank of Minneapolis Staff Report 297.

Barro, R. J., and D. B. Gordon. (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics* 12:101–121.

Canzoneri, M. B. (1985). Monetary policy games and the role of private information. *American Economic Review,* 75:1056–1070.

Chari, V. V., L. J. Christiano, and P. J. Kehoe. (1994). Optimal fiscal policy in a real business cycle model. *Journal of Political Economy* 102:617–652.

———, and P. J. Kehoe. (1990). Sustainable plans. *Journal of Political Economy* 98: 783–802.

———, ———, and E. C. Prescott. (1989). Time consistency and policy. In *Modern Business Cycle Theory*, R. J. Barro (ed.). Harvard University Press.

Clarida, R., J. Gali, and M. Gertler. (1999). The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature* 37:1661–1707.

Cole, H., and L. Ohanian. (2001). New Deal policies and the persistence of the Great Depression. Federal Reserve Bank of Minneapolis.

Faust, J., and L. E. O. Svensson. (2001). Transparency and credibility: Monetary policy with unobserved goals. *International Economic Review* 42.

Fischer, S. (1980). Dynamics inconsistency, cooperation and the benevolent dissembling government. *Journal of Economic Dynamics and Control* 2:91–108.

Friedman, M. (1948). A monetary and fiscal framework for economic stability. *American Economic Review* 38:245–264. Reprinted in M. Friedman, *Essays in Positive Economics*, University of Chicago Press, 1953.

Goodfriend, M. (1986). Monetary mystique: Secrecy and central banking. *Journal of Monetary Economics* 17:63–92.

Green, E. J., and R. H. Porter. (1984). Noncooperative collusion under imperfect price information. *Econometrica* 52:87–100.

Ireland, P. N. (1997). Sustainable monetary policies. *Journal of Economic Dynamics and Control* 22:87–108.

———. (1999). Does the time-consistency problem explain the behavior of inflation in the United States? *Journal of Monetary Economics* 44:279–291.

Kydland, F. E., and E. C. Prescott. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85:473–491.

Lucas, R. E., and N. L. Stokey. (1983). Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12:55–93.

McCallum, B. T. (1999). Issues in the design of monetary policy rules. In *Handbook of Macroeconomics*, J. B. Taylor and M. Woodford (eds.). Elsevier.

Persson, T., and G. Tabellini. (1994). Introduction. In *Monetary and Fiscal Policy, Volume 1: Credibility*. Cambridge, MA: The MIT Press.

Phelan, C. (2001). Public trust and government betrayal. Federal Reserve Bank of Minneapolis Working Paper.

Poole, W. (1970). Optimal choice of monetary instruments in a simple stochastic macro model. *Quarterly Journal of Economics* 84:197–216.

———. (1985). The optimal degree of commitment to a monetary target. *Quarterly Journal of Economics* 100:1169–1190.

———. (1987). Reputational constraints on monetary policy. *Carnegie-Rochester Series on Public Policy* 26:141–182.

Rogoff, K. (1989). Reputation, coordination, and monetary policy. In *Modern Business Cycle Theory*, R. J. Barro (ed.). Harvard University Press, pp. 83–96.

Stokey, N. L. (1991). Credible public policy. *Journal of Economic Dynamics and Control* 15:627–656.

Zarazaga, C. E. J. M. (1995). Recurrent hyperinflations in a dynamic game with imperfect monitoring in the appropriation of seignorage. Working Paper, Federal Reserve Bank of Dallas.

Zhu, Xiaodong. (1992). Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58:250–289.

Comment

PETER N. IRELAND
Boston College and NBER

1. Introduction

In this paper, Nancy Stokey presents two examples in which the time-consistency problem arises in a macroeconomic policymaking context. I consider these two examples to be extremely well chosen, for several reasons. First, each example deals with an important problem—the choice of a monetary policy instrument or the choice of a capital income tax rate—that is of considerable interest in and of itself. Second, each serves to introduce us to some powerful analytic techniques that recently have been developed by researchers working at the frontiers of economic science. The examples show us how far this branch of the literature has come, from a technical perspective, in the twenty-five years since the publication of Kydland and Prescott's (1977) original paper.

But, third and perhaps most important of all, I consider these two examples to be well chosen because each uses a model that shares its most basic features with all of the other models that have been developed in the literature that builds on Kydland and Prescott (1977). Thus, each of Nancy's models has implications for a wide range of issues that have already been studied extensively, and, by the same token, each remains silent on some important issues that have yet to be fully discussed, much less satisfactorily resolved, in the literature as it stands now.

In fact, because Nancy's models are so representative of others from this branch of the literature, I will be able to use one of them here in my discussion to provide answers of my own to two more basic and fundamental questions. First, what have we learned about the time-consistency problem in the twenty-five years since Kydland and Prescott (1977)? And second, what more might we hope to learn about the problem over the next twenty-five years?

2. The Model

My model is a simplified version of one of Nancy's: the model that she uses to study the trade-off between tightness and observability in the choice of a monetary policy instrument. My version of the model is simplified in that it eliminates the random and unobservable elements that play a key role in Nancy's analysis, but are less essential for my purposes. For the most part, I borrow my notation directly from Nancy's paper, although I make a few minor changes here and there, when they serve to make the results cleaner and easier to understand.

My model describes the behavior of a central bank, which chooses the rate of price inflation π, and a representative household, whose actions determine the rate of wage inflation w. As explained in more detail below, the representative household sets the rate of wage inflation based on its expectations of the central bank's choice of π. In this model, therefore, the variable w also serves as a convenient proxy for expected inflation.

The central bank's objectives are summarized by the single-period return or payoff function

$$R(\pi; w) = -\left(\frac{1}{2}\pi^2 + \frac{b}{2}(w - \pi + \alpha)^2\right),$$

where the parameters b and α are both nonnegative. The first term in this objective function captures the costs of inflation or, more precisely, penalizes deviations of the inflation rate π from the central bank's target of zero. To interpret the second term, consider the expectational Phillips curve

$$U = U^n - (\pi - w),$$

where U denotes the actual rate of unemployment, where U^n denotes the natural rate of unemployment, and where, since w measures expected inflation, $\pi - w$ serves as a measure of surprise inflation. Then

$$w - \pi + \alpha = U - (U^n - \alpha),$$

indicating that, according to the objective function R, the central bank sets a target $U^n - \alpha$ for the unemployment rate that lies *below* the natural rate. The parameter b measures the weight that the central bank places on achieving this goal for unemployment, relative to its goal for inflation.

The representative household in this model has a very simple objective: it wishes to set w as close as possible to the central bank's choice of π. The representative household has rational expectations, which here in the absence of shocks translates into perfect foresight. Thus, the household always accomplishes its goal by setting w exactly equal to π.

This condition, $w = \pi$, must always hold in equilibrium: it summarizes the implications of the household's optimizing behavior. What Kydland and Prescott's (1977) original paper teaches us is that macroeconomic outcomes depend critically on whether or not the central bank also views this equilibrium condition as a constraint that links its choice of π to the representative household's choice of w. To see this, let's consider the two basic cases.

CASE 1: COMMITMENT In this first case, the central bank has the willingness and the ability to precommit to a choice for π at the beginning of the period, before the household embeds its expectations into a particular choice of w. Since the central bank moves first, it views the equilibrium condition $w = \pi$ as a constraint that links its choice of π to a subsequent setting for w. In this case, therefore, the central bank solves the constrained optimization problem

$$\max_{\pi} R(\pi; w) \qquad \text{subject to} \quad w = \pi.$$

The first-order condition for this problem implies that the optimal inflation rate with commitment, denoted by π^c, equals zero:

$$\pi^c = 0.$$

When the central bank precommits to a choice for π, it recognizes that it is losing any ability it might otherwise have to surprise the representative household and thereby exploit the Phillips curve. Hence, in this case with commitment, the central bank abandons any idea of pushing unemployment below the natural rate, and instead focuses exclusively on achieving its goal of zero inflation.

CASE 2: NO COMMITMENT In this second case, the central bank is either unwilling or unable to precommit, and effectively makes its choice of π after the representative household has embedded its expectations into a particular choice of w. Since the central bank moves second, it no longer perceives $w = \pi$ as a constraint. Instead, the central bank simply takes w as given, and solves the unconstrained optimization problem

$$\max_{\pi} R(\pi; w).$$

The first-order condition for this problem dictates that the central bank's optimal choice without commitment, denoted by π^{nc}, is given by

$$\pi^{nc} = \frac{b}{1 + b}(w + \alpha).$$

In equilibrium, however, the condition $w = \pi$ must still hold: in particular, the representative household perfectly anticipates the central bank's actions, and sets $w = \pi^{nc}$. Combining this equilibrium condition with the central bank's first-order condition reveals that in this case without commitment,

$$\pi^{nc} = b\alpha \geq 0.$$

Comparing the outcomes with and without commitment, $\pi^c = 0$ and $\pi^{nc} = b\alpha \geq 0$, serves to crystalize Kydland and Prescott's (1977) original message. The central bank that is either unwilling or unable to precommit to a choice of π finds itself tempted to exploit the expectational Phillips curve, in an effort to achieve its goal of pushing unemployment below the natural rate. The representative household has rational expectations, however, and understands that the central bank faces this temptation to inflate. The household, therefore, builds these inflationary expectations into its wage-setting decisions, so that unemployment remains at its natural rate. The central bank's efforts to exploit the Phillips curve lead only to a suboptimally high rate of inflation.

3. *What Have We Learned since Kydland and Prescott (1977)?*

But what else have we learned about the time-consistency problem in the twenty-five years since the publication of Kydland and Prescott's (1977) original paper? Comparing the outcomes $\pi^c = 0$ and $\pi^{nc} = b\alpha$ immediately reveals that in this simple version of Nancy's model, $b\alpha$ conveniently measures the inflationary bias that results when the central bank does not precommit to its choice for π. This expression, $b\alpha$, also suggests that there are at least two promising strategies that policymakers can use to minimize the inflationary bias, and thereby improve welfare.

One possibility involves setting the parameter α equal to zero, that is, instructing the central bank to stop targeting an unemployment rate that lies below the natural rate. McCallum (1995) argues passionately in favor of this solution to the central bank's time-consistency problem, and Blinder (1997) suggests that in practice, Federal Reserve officials have acted to minimize the importance of the time-consistency problem by behaving as if $\alpha = 0$. In fact, when $\alpha = 0$ in the simple model considered here, the time-consistency problem vanishes: outcomes with and without commitment coincide.

A second possibility involves setting the parameter b equal to zero, that is, instructing the central bank to stop caring so much about unemployment in the first place. This proposed solution to the time-consistency problem corresponds, of course, to Rogoff's (1985) suggestion that the appointment of a conservative central banker can lead to preferred outcomes in cases where monetary precommitment is impossible. And again, in the context of this simple model, outcomes with and without commitment coincide when $b = 0$.

Much of the recent literature that builds on Kydland and Prescott's (1977) original study focuses on the choice between these two solutions to the time-consistency problem for monetary policymaking. As noted above, both solutions work perfectly well in the context of the simple nonstochastic model used here. However, in more complicated models where random shocks give rise to a trade-off between the *variability* as well as the *levels* of inflation on one hand and unemployment on the other, Clarida, Gali, and Gertler (1999), Herrendorf and Lockwood (1997), and Svensson (1997) find that in addition to the *inflationary bias* that arises here, a *stabilization bias* also emerges in the absence of commitment: the discretionary central bank works too hard to stabilize unemployment, and not hard enough to stabilize inflation, in response to the shocks that hit the economy.

All three of these recent papers demonstrate that while the inflationary bias vanishes when $\alpha = 0$, so that the central bank's target for unemployment coincides with the natural rate, the stabilization bias remains. All three of these papers also suggest that the alternative solution of appointing a conservative central banker, with a lower value of b, can work to minimize both the inflationary bias and the stabilization bias, especially in cases where the conservative central banker is also offered an inflation contract of the kind first proposed by Walsh (1995). This, in my view, represents one of the most important lessons to have come out of the literature that builds on Kydland and Prescott (1977): that in situations where the time-consistency problem arises, it can be desirable to appoint policymakers whose preferences or incentives differ systematically from those of society as a whole.

In the U.S. economy, therefore, consider Federal Reserve Chairmen Volker and Greenspan, both of whom might reasonably be described as conservative in the Rogoffian sense of caring more about inflation, and less about unemployment, than the average American consumer or worker. It is certainly legitimate to ask whether, in a representative democracy like ours, it is really appropriate to give men like Volker and Greenspan power over such an important component of macroeconomic policy. The literature that builds on Kydland and Prescott (1977), however, provides us with a compelling response to this concern, by demonstrating that in situations like monetary policymaking, where the time-consistency problem may arise, it makes sense to appoint conservative central bankers—even when the ultimate goal is to maximize the welfare of the economy's representative household.

4. What More Can We Learn?

And what additional lessons might we hope to learn over the next twenty-five years? As a first step in answering this question, consider following Barro and Gordon (1983) and Ireland (1997) in allowing the monetary policymaking game described above to be repeated over an infinite horizon, where time periods are indexed by $t = 0, 1, 2, \ldots$. Suppose, as in Case 2 above, that the central bank does not precommit to its choice for inflation; but suppose, also, that the behavior of the representative household's expectations provides the central bank with an incentive to maintain a reputation for keeping inflation low.

More specifically, suppose that at the beginning of period $t = 0$, the representative household expects the central bank to choose an inflation rate $\pi_0 = \pi^{\text{rep}}$ for that period, where π^{rep} lies somewhere between $\pi^{c} = 0$ and $\pi^{\text{nc}} = b\alpha$. Suppose, in addition, that in each period $t = 1, 2, 3, \ldots$, the household continues to expect the central bank to choose $\pi_t = \pi^{\text{rep}}$ so long as it has always done so in the past. If, however, the central bank deviates during some period $t = 0, 1, 2, \ldots$, by choosing an inflation rate π_t that differs from π^{rep}, then the household's expectations permanently shift, so that the no-commitment choice $\pi^{\text{nc}} = b\alpha$ is expected forever after. Given the household's objective of setting w in line with expected inflation, these assumptions imply that for $t = 0$, $w_0 = \pi^{\text{rep}}$, while for all $t = 1, 2, 3, \ldots$,

$$w_t = \begin{cases} \pi^{\text{rep}} & \text{if } \pi_s = \pi^{\text{rep}} \text{ for all } s = 0, 1, \ldots, t-1, \\ \pi^{\text{nc}} = b\alpha, & \text{otherwise.} \end{cases}$$

The question now becomes: given this behavior of private-sector expectations, will the central bank choose to maintain its reputation for selecting the lower inflation rate π^{rep}?

In the case where the central bank does maintain its reputation by choosing $\pi_t = \pi^{\text{rep}}$ for all $t = 0, 1, 2, \ldots$, its total discounted return over the infinite horizon is given by

$$\frac{1}{(1-\beta)} R(\pi^{\text{rep}}; \pi^{\text{rep}}),$$

where β, the central bank's discount factor, lies between zero and one. In the alternative case, where the central bank deviates, it will always find it optimal to do so immediately, during period $t = 0$, by choosing π_0 to solve the problem

$$\max_{\pi} R(\pi; \pi^{\text{rep}}).$$

Hence, in this case, the central bank chooses $\pi_0 = \pi^{\text{dev}}$, where

$$\pi^{\text{dev}} = \frac{b}{1+b}(\pi^{\text{rep}} + \alpha).$$

During each period thereafter, having lost its reputation, the best that the central bank can do is to select $\pi_t = \pi^{\text{nc}} = b\alpha$ for all $t = 1, 2, 3, \ldots$. Its total discounted return from deviating is therefore

$$R(\pi^{\text{dev}}; \pi^{\text{rep}}) + \frac{\beta}{1-\beta} R(\pi^{\text{nc}}; \pi^{\text{nc}}).$$

It follows that the policy choice $\pi_t = \pi^{\text{rep}}$ for all $t = 0, 1, 2, \ldots$ is sustainable in this type of reputational equilibrium if and only if the incentive-compatibility constraint

$$R(\pi^{\text{rep}}; \pi^{\text{rep}}) \geq (1 - \beta)R(\pi^{\text{dev}}; \pi^{\text{rep}}) + \beta R(\pi^{\text{nc}}; \pi^{\text{nc}})$$

holds. Using the solutions $\pi^{\text{dev}} = [b/(1+b)](\pi^{\text{rep}} + \alpha)$ and $\pi^{\text{nc}} = b\alpha$, this incentive constraint can be rewritten as

$$[\beta(2+b) - 1](b\alpha)^2 + 2(1-\beta)b\alpha\pi^{\text{rep}} - (\beta b + 1)(\pi^{\text{rep}})^2 \geq 0.$$

This last expression indicates that the zero-inflation policy that is optimal under commitment can be supported in a reputational equilibrium whenever $\beta \geq 1/(2+b)$. And since b is nonnegative, this condition almost certainly holds: it is satisfied for any value of β exceeding $1/2$. Here, therefore, we have another lesson to have emerged from the literature since Kydland and Prescott (1977): a central bank that is sufficiently patient, and that is lucky enough to be endowed with a reputation for keeping inflation low, will find it optimal to maintain its reputation even if it lacks the ability to commit.

One can also show, however, that if $\beta \geq 1/(2+b)$, so that the central bank's incentive constraint holds with $\pi^{\text{rep}} = 0$, then the incentive constraint also holds for *any* value of π^{rep} between $\pi^{\text{c}} = 0$ and $\pi^{\text{nc}} = b\alpha$. In this case, therefore, the model features multiple equilibria, supporting inflation rates that range all the way from zero to $b\alpha$. To see why this is a problem, consider a reputational equilibrium in which π^{rep} lies below $\pi^{nc} = b\alpha$, but closer to $\pi^{nc} = b\alpha$ than to $\pi^{c} = 0$. In such an equilibrium, the central bank benefits from maintaining its reputation: it achieves an outcome that improves upon the endless repetition of the one-shot outcome without commitment. At the same time, however, the central bank knows that even better equilibria exist, with even lower inflation rates. Yet the model provides absolutely no advice as to how the central bank might steer the economy towards these preferred, low-inflation equilibria.

Taylor (1982) suggests that a central bank ought to build credibility for a low-inflation policy by adopting that policy unilaterally and by demonstrating that it will stick with the policy, even if it imposes short-run costs on the economy. Taylor's pragmatic approach has considerable intuitive appeal, and may be a good strategy for any real-world central bank to follow. But it simply will not work in the context of the example considered here. In fact, the triggerlike behavior of the representative household's expectations that help support the reputational equilibria with $\pi^{rep} < \pi^{nc} = b\alpha$ dictates that expected inflation will actually jump *higher*, to $\pi^{nc} = b\alpha$, should the central bank deviate from π^{rep} by unexpectedly trying to disinflate.

How can a central bank establish a reputation for fighting inflation, or build credibility for a welfare-improving disinflationary program? In the literature that follows Kydland and Prescott (1977), work towards answering this question has only just begun. Significantly, providing answers to this question would seem to require departing in some way from the rational-expectations hypothesis, since, after all, the reputational equilibria in which inflation is stuck forever between $\pi^{c} = 0$ and $\pi^{nc} = b\alpha$ are bona fide rational-expectations equilibria. Cho and Matsui (1995) and Ireland (2000), for instance, both develop models of macroeconomic policymaking in which private agents are assumed to be boundedly rational. The objective of both of these papers is to identify restrictions on private-expectation formation that are weak enough to allow Taylor's (1982) pragmatic approach to work, but strong enough to prevent the policymaker from repeatedly fooling the boundedly rational agents. Still, much more work needs to be done along these lines: we have much to learn, over the next twenty-five years, about how governments can build credibility for the policies—like low inflation and low capital income tax rates—that we'd like them to pursue.

REFERENCES

Barro, R. J., and D. B. Gordon. (1983). Rules, discretion, and reputation in a model of monetary policy. *Journal of Monetary Economics* 12(July):101–121.

Blinder, A. S. (1997). Distinguished lecture on economics in government: What central bankers could learn from academics—and vice versa. *Journal of Economic Perspectives* 11(Spring):3–19.
Cho, I.-K., and A. Matsui. (1995). Induction and the Ramsey policy. *Journal of Economic Dynamics and Control* 19 (July–September):1113–1140.
Clarida, R., J. Gali, and M. Gertler. (1999). The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature* 37(December):1661–1707.
Herrendorf, B., and B. Lockwood. (1997). Rogoff's "conservative" central banker restored. *Journal of Money, Credit, and Banking* 29(November, Part 1):476–495.
Ireland, P. N. (1997). "Sustainable monetary policies." *Journal of Economic Dynamics and Control* 22(November):87–108.
———. (2000). Expectations, credibility, and time-consistent monetary policy. *Macroeconomic Dynamics* 4(December):448–466.
Kydland, F. E., and E. C. Prescott. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85(June):473–491.
McCallum, B. T. (1995). Two fallacies concerning central-bank independence. *American Economic Review* 85(May):207–211.
Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100(November):1169–1189.
Svensson, L. E. O. (1997). Optimal inflation targets, "conservative" central banks, and linear inflation contracts. *American Economic Review* 87(March):98–114.
Taylor, J. B. (1982). Establishing credibility: A rational expectations viewpoint. *American Economic Review* 72(May):81–85.
Walsh, C. E. (1995). Optimal contracts for central bankers. *American Economic Review* 85(March):150–167.

Comment[1]

LARS E. O. SVENSSON
Princeton University, NBER, and CEPR

Nancy Stokey's interesting and thought-provoking paper has two main parts. Section 2, "Reputation Building," discusses the choice of monetary-policy instruments by relying on Atkeson and Kehoe (2001). This discussion is in terms of a trade-off between observability and "tightness" (the correlation with the monetary-policy goal). Section 3, "Robustness," discusses the choice between discretion and commitment to a simple rule. This discussion is in terms of a trade-off between flexibility and myopia on the one hand and rigidity and farsightedness on the other.

I believe Section 2 is better described as concerned with the choice of an *intermediate target* for monetary policy rather than a monetary-policy *instrument*. The setting of the monetary-policy instrument (the Fed funds rate in the United States) is usually directly observable, whereas the rela-

1. I thank Annika Andreasson for secretarial and editorial assistance.

tion between the instrument setting and the monetary-policy goals is complex, making it difficult to infer the central bank's intentions from its instrument setting. Thus, I interpret Section 2 as a discussion of the pros and cons of either an exchange-rate target or a money-growth target as intermediate targets, when the final target (the goal) is inflation.

The choice of an intermediate target is a classic problem in the design of monetary policy. An ideal intermediate target is (1) highly correlated with the goal, (2) easier to control than the goal, and (3) easier to observe than the goal. The idea is that, if such an ideal intermediate target can be found, it may be better to aim for the intermediate target rather than to aim directly for the goal, and this way indirectly achieve the goal.

In current real-world monetary policy, the idea of intermediate targeting has largely been abandoned (except in a specific sense mentioned below). Instead, central banks nowadays aim directly for their goals, typically low inflation and (to some extent) stable output gaps, as in (flexible) inflation targeting. The main problem with inflation targeting is that the control of inflation (and the output gap) is very imperfect, due to the lags in, and different strengths of, the various channels in the transmission process from instrument adjustments to actual inflation and output. This makes it difficult to judge whether current policy settings are, and past policy settings were, appropriate. The best solution to this problem is to regard inflation and output-gap forecasts as intermediate targets.

Indeed, as discussed in Svensson (1997a), the inflation forecast is an ideal intermediate target variable when inflation is the final target variable. The inflation forecast is by definition the current variable that has the highest correlation with future inflation. It is easier to control than actual inflation, for instance, because it leaves out a number of unanticipated shocks that will later affect actual inflation. It is in principle easier to observe than actual inflation, since it is a variable currently available, whereas the corresponding actual inflation will only be observed some two years later (due to the lags) and then be contaminated by a number of intervening shocks. In particular, transparent inflation-targeting central banks *make* their inflation forecasts observable, by issuing detailed inflation reports where the forecast is presented and motivated. (Thus, arguably, the only ideal intermediate target variables are the forecasts of the final target variables.)

Section 3, "Robustness," discusses the trade-off between flexibility and an inflation bias on the one hand and rigidity and no inflation bias on the other. This is a well-known and classic issue. For instance, the purpose of a fixed exchange rate or a currency board in a country may be to avoid inflation bias by importing a less inflationary monetary policy from an anchor country. But this is a second-best solution, since monetary policy can then no longer be independent and respond to the specific shocks hitting the country.

In real-world monetary policy, however, it seems possible in many cases to get rid of any inflation bias *without* losing flexibility and stabilization. In order to discuss this, let us go back to the classic treatments in Kydland and Prescott (1977) and Barro and Gordon (1983a) of rules vs. discretion and the time-consistency problem. Although these issues in principle apply to a number of different policies, monetary policy provides the best examples, having arguably suffered the largest problems and benefited the most from their solutions.

The main result in the classic treatment was that discretion may result in an average inflation bias. The simplest way to illustrate this result is with the help of a simple Lucas-type Phillips curve,

$$y_t = \bar{y} + \alpha(\pi_t - \mathrm{E}_{t-1}\pi_t) + \varepsilon_t,$$

where y_t is output in period t, $\bar{y}$ is potential output (the natural output level), α is a positive coefficient, π_t is inflation in period t, $\mathrm{E}_{t-1}\,\pi_t$ denotes rational expectations of inflation in period t conditional on information available in period $t - 1$, and ε_t is a zero-mean i.i.d. shock. The central bank is assumed to control either inflation or output, and has a quadratic loss function,

$$L_t = (\pi_t - \pi^*)^2 + \lambda(y_t - \bar{y} - k)^2,$$

where π^* is an inflation target, λ is a positive weight, and k is a positive parameter. This formulation implies that the output target, $\bar{y} + k$, is larger than potential output, $\bar{y}$.

Discretionary optimization of the central bank implies the first-order condition

$$\pi_t - \pi^* + \lambda\alpha(y_t - \bar{y} - k) = 0.$$

Combining this condition and the Phillips curve gives the equilibrium outcome for inflation and output,

$$\pi_t = \pi^* + \lambda\alpha k - \frac{\lambda\alpha}{1 + \lambda\alpha^2}\varepsilon_t,$$

$$y_t = \bar{y} + \frac{1}{1 + \lambda\alpha^2}\varepsilon_t.$$

In particular, there is an average inflation bias, in that the unconditional mean of inflation exceeds the inflation target:

$$E[\pi_t] - \pi^* = \lambda\alpha k > 0.$$

Numerous solutions to the problem of average inflation bias under discretion have been suggested. One solution is a commitment to an optimal reaction function. In the absence of a commitment mechanism, this solution is not realistic. In particular, in any realistic problem, the optimal reaction function is quite complex and in practice unverifiable, making a commitment to it very difficult or impossible.

Another solution is by extension to non-Markov trigger-strategy equilibria, following Barro and Gordon (1983b). These have the inherent problem that follows from the folk theorem: there is no unique equilibrium. Furthermore, in the realistic situation with an atomistic private sector, there is no coordination mechanism by which a particular equilibrium could be achieved.[2] In addition, these equilibria are sometimes (and in Stokey's paper) referred to as having to do with "reputation." I think that is a (very common) misnomer. There is no uncertainty about the characteristics of the players in these settings. I think "reputation" is much more naturally associated with a situation of incomplete information, when the preferences of the central bank are not directly observable, as is the case in classic papers by Backus and Driffill (1985) and Cukierman and Meltzer (1986), and in the recent extension of the latter by Faust and Svensson (2001). In these papers, "reputation" is the private sector's best estimate of the preferences of the central bank.

A much-noted suggestion is McCallum's (1995) "just do it." This assumes that the central bank, in the absence of a commitment mechanism, just ignores the incentives to deviate from the socially optimal outcome that arises under discretion. I find this suggestion problematic because, to my knowledge, neither McCallum nor anyone else has presented a model where "just do it" is an equilibrium outcome. The best rationale for "just do it" that I am aware of is in Faust and Svensson (2001): There, increased transparency about the bank's actions makes the bank's "reputation" (the private sector's estimate of the bank's unobservable internal time-varying objectives) more sensitive to its actions. This increases the cost for the bank of deviating from its announced social objective and pursuing its internal objectives, and thus works as an implicit mechanism for commitment to the announced objective.

Many papers have fruitfully applied a principal–agent approach to the time-consistency problem. Here society (the principal) can assign loss functions to the central bank (the agent) that may differ from society's

2. Problems with trigger-strategy equilibria are further discussed in Ireland's comment preceding this one.

loss function, in order to improve the discretionary problem.[3] That is, it is assumed that it is possible to commit the central bank to a particular loss function, whereas the minimization of that loss function occurs under discretion. A well-known suggestion is Rogoff's (1985) "weight-conservative" central bank, where the central bank is assigned a relative weight λ on output stabilization that is less than that of society. This reduces average inflation and inflation variance, but increases output variance. This is often described as a necessary trade-off between inflation bias and "flexibility." However, this potential explanation of low inflation in some countries is rejected by the data: Countries with lower average inflation do not have higher output variability.

Another suggestion is an "inflation contract," by Walsh (1995) and Persson and Tabellini (1994), further discussed in Svensson (1997b), where lower inflation is assumed to be accompanied by an increased bonus to the central bank or its governor. This idea has never been tried in the real world (not even in New Zealand, counter to common misperceptions).

A third suggestion is an "output-conservative" central bank, meaning a loss function for the central bank where the output target is equal to potential output, $k = 0$. This eliminates the average inflation bias without increasing output variability and is hence consistent with the data. This explanation has been suggested by Blinder (1998) for the Fed. I believe this is the best single explanation for the apparent disappearance of average inflation bias in many countries. Indeed, I believe that the flexible inflation targeting currently applied in an increasing number of countries is consistent with central-bank loss functions where there is some modest weight on output-gap stability and the output target equals potential output. Thus, this solution to the inflation-bias problem need not imply any loss in flexibility. It is consistent with the insight that society had better find other policies than monetary policy (such as structural policies improving competition) to increase average and potential output.[4]

Issues of commitment and discretion have been discussed in a more general linear–quadratic model in early papers of Oudiz and Sachs (1985), Currie and Levine [collected in Currie and Levine (1993)], and Backus and Driffill (1986), with the model equations

$$\begin{bmatrix} X_{t+1} \\ \mathrm{E}_t x_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t + \begin{bmatrix} \varepsilon_t + 1 \\ 0 \end{bmatrix}.$$

3. The possibility of improving the discretionary equilibrium by adjusting the parameters of the central-bank loss function was noted in Barro and Gordon (1983a, footnote 19).

4. Ireland, in his comment, interprets, McCallum's "just do it" as modifying the central-bank loss function by setting $k = 0$, but I can't find any support for that interpretation in McCallum (1995) (in McCallum's notation it would amount to setting $k = 1$).

Here, X_t is a vector of predetermined variables (one of these can be unity, in order to handle constants in a convenient way), x_t is as vector of forward-looking variables (jump variables, nonpredetermined variables), i_t is a vector of policy instruments, ε_t is a vector of zero-mean i.i.d. shocks, and A and B are matrices of appropriate dimension. The policymaker's intertemporal loss function in period t is

$$\mathrm{E}_t(1 = \delta) \sum_{\tau=0}^{\infty} \delta^{\tau} L_{t+\tau},$$

where δ, $0 < \delta < 1$, is a discount factor, and the period loss function L_t is quadratic:

$$L_t = (Y_t - \hat{Y})' W (Y_t - \hat{Y}).$$

Here W is a positive semidefinite weight matrix, Y_t is a vector of target variables, and $\hat{Y}$ is a vector of corresponding target levels, which can be written

$$Y_t - \hat{Y} = C \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix},$$

where C is a matrix.

The optimal reaction function under *commitment* (the optimal "instrument rule") can be written

$$i_t = F X_t + \Phi \Xi_{t-1},$$

where F and Φ are matrices and Ξ_t is a vector of Lagrange multipliers for the equations for the forward-looking variables (the lower block of the model equations above), the equilibrium dynamics of which are given by

$$\Xi_t = S X_t + \Sigma \Xi_{t-1},$$

where S and Σ are matrices.

The equilibrium reaction function resulting from optimization under *discretion* can be written

$$i_t = \tilde{F} X_t,$$

where $\tilde{F}$ is a matrix. Compared to the optimal reaction function under commitment, there is generally *stabilization bias* [meaning that the matrix of response coefficients $\tilde{F}$ under discretion differs from the optimal response F under commitment, as discussed in Svensson (1997b), for instance] and *lack of history dependence* [$\Phi \equiv 0$, as discussed in Woodford (1999)]. Optimization under discretion thus results in a higher loss than under commitment.

Several solutions to the problem of how to improve the equilibrium under discretion have been suggested for this more general setting. One solution is a commitment to the optimal reaction function above. Unfortunately, in realistic problems the optimal reaction function is quite complex, making verification and other aspects of a commitment mechanism unrealistic.

A commitment to a *simple* (rather than optimal) instrument rule, such as a Taylor rule, has been suggested as a compromise. A simple instrument rule could be verifiable, and a commitment would in principle be feasible. No central bank has committed to a simple instrument rule, however, and prominent central bankers seem skeptical [see Svensson (2002) for further discussion].

One solution is a *commitment to continuity and predictability,* suggested by Svensson and Woodford (2002), who argue that such a commitment is to some extent consistent with both the rhetoric and the practice of current inflation targeting. It consists in internalizing the cost of deviating from previous expectations, and boils down to a modified period loss function of the form

$$L_t = (Y_t - \hat{Y})'W(Y_t - \hat{Y}) + \Xi_{t-1}(x_t - \mathrm{E}_{t-1}x_t),$$

where Ξ_{t-1} is the vector of Lagrange multipliers from the previous decision period.

Another solution is a commitment to an *optimal targeting rule,* discussed in Svensson and Woodford (2002) and Svensson (2002), and consistent with previous work of Sims (1980), Rogoff (1985), Walsh (1995), and Svensson (1997a). An optimal targeting rule is an Euler condition, an optimal first-order condition for the target variables—essentially, the equality of the marginal rate of transformation between the target variables (given by the model equations) and the corresponding marginal rate of substitution (given by the loss function). One attraction of optimal targeting rules is that they are usually much simpler and more robust than the optimal reaction function, making a commitment to them more realistic. For instance, all additive shocks to the model equations vanish from the optimal targeting rule (but not from the optimal reaction function). In the simple Kydland–Prescott–Barro–Gordon model above, the optimal targeting rule is

$$\pi_t - \pi^* + \lambda\alpha(y_t - \bar{y}) = 0.$$

With lags in the transmission mechanism of monetary policy, the optimal targeting rule involves forecasts of the target variables rather than current values.

Another alternative is a commitment to a *simple targeting rule*, for instance, the simple rule emphasized by the Bank of England and Sweden's Riksbank that the two-year-ahead inflation forecast should be equal to the inflation target.

So what have we learnt about rules and discretion after twenty-five years? I believe the most important things we have learnt are:

The problem of average inflation bias seems to be gone. The single best explanation for its disappearance is probably *output-conservativeness* of central banks—that is, central banks, in addition to an explicit or implicit inflation target, have an explicit or implicit output target equal to (rather than exceeding) potential output. This also means that average inflation bias can be avoided without loss in flexibility or stabilization of the output gap.

Even if no average inflation bias occurs, discretion generally implies stabilization bias and lack of history dependence (although the quantitative importance of these two phenomena remains to be firmly established).

A principal–agent approach to central banking is useful. Commitment to objectives (loss functions) is probably more realistic and relevant than commitment to particular reaction functions (instrument rules).

Targeting rules may be more useful and realistic than instrument rules.

REFERENCES

Atkeson, A., and P. J. Kehoe. (2001). The advantage of transparent instruments of monetary policy. Federal Reserve Bank of Minneapolis Staff Report 297.

Backus, D., and J. Driffill. (1985). Rational expectations and policy credibility following a change in regime. *Review of Economic Studies* 52:211–222.

———, and ———. (1986). The consistency of optimal policy in stochastic rational expectations models. CEPR Discussion Paper No. 124.

Barro, R. J., and D. B. Gordon. (1983a). A positive theory of monetary policy in a natural-rate model. *Journal of Political Economy* 91:589–610.

———, and ———. (1983b). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics* 12:101–121.

Blinder, A. S. (1998). *Central Banking in Theory and Practice: The 1996 Robbins Lectures*. Cambridge, MA: The MIT Press.

Cukierman, A., and A. H. Meltzer. (1986). A theory of ambiguity, credibility and inflation under discretion and asymmetric information. *Econometrica* 53:1099–1128.

Currie, D., and P. Levine. (1993). *Rules, Reputation and Macroeconomic Policy Coordination*. Cambridge: Cambridge University Press.

Faust, J., and L. E. O. Svensson. (2001). Transparency and credibility: Monetary policy with unobservable goals. *International Economic Review* 42:369–397.
Kydland, F. E., and E. C. Prescott. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85:473–491.
McCallum, B. (1995). Two fallacies concerning central-bank independence. *American Economic Review* 85, 207–211.
Oudiz, G., and J. Sachs. (1985). International policy coordination in dynamic macroeconomic models. In *International Economic Policy Coordination,* W. H. Buiter and R. C. Marston (eds.). Cambridge: Cambridge University Press.
Persson, T., and G. Tabellini. (1994). Designing institutions for monetary stability. *Carnegie-Rochester Conference Series on Public Policy* 39:53–84.
Rogoff, K. (1985). The optimal degree of commitment to a monetary target. *Quarterly Journal of Economics* 100:1169–1190.
Sims, C. A. (1980). Macroeconomics and reality. *Econometrica* 48:1–48.
Svensson, L. E. O. (1997a). Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review* 41:1111–1146.
———. (1997b). Optimal inflation targets, "conservative" central banks, and linear inflation contracts. *American Economic Review* 87:98–114.
———. (2002). What is wrong with Taylor rules? Using judgment in monetary policy through targeting rules. Princeton University, working paper.
———, and M. Woodford. (2002). Implementing optimal policy through inflation-forecast targeting. Princeton University, working paper.
Walsh, C. E. (1995). Optimal contracts for independent central bankers. *American Economic Review* 85:150–167.
Woodford, M. (1999). Optimal monetary policy inertia. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper 7261.

Discussion

A theme taken up by several of the participants was the different possible interpretations of U.S. inflation history in the 1970s and 1990s. Greg Mankiw suggested as an alternative to Lars Svensson's interpretation that central bankers in the 1970s thought that the natural level of output was higher than it actually was, whereas in the 1990s, central bankers thought that the natural level of output was lower than it actually was. He added that such an alternative interpretation would be bad for the rules-vs.-discretion literature. It would imply that monetary policy in the 1990s was better than 30 years ago not because monetary policy was less discretionary in the 1990s than in the 1970s, but because central bankers were lucky in the shocks that hit the economy.

Robert Barsky proposed a variation on Mankiw's comment. Where Mankiw emphasized the importance of the Fed knowing or not knowing what was the natural level of output, Barsky suggested that an alternative explanation for recent U.S. inflation history is that the Fed learned the natural-rate hypothesis over time.

Bob Hall commented that when Alan Greenspan was asked why he tolerated unemployment below the natural rate in the 1990s, he replied that he had focused not on unemployment but on what was happening to prices. Hall's interpretation of this reply was that Greenspan's success relative to the central bankers of the 1970s can be attributed to a policy of price targeting, and not merely to good luck. Hall also maintained that, contrary to the teaching of Milton Friedman, the idea of an exogenous natural rate of unemployment is not a sensible one. On this point, Greg Mankiw responded that Friedman's idea was merely that the natural rate of unemployment was exogenous to monetary policy. Hall replied that there is strong evidence of hysteresis in the labor market and of monetary policy affecting the labor market, and hence exogeneity of the natural rate of unemployment to monetary policy is unlikely. Greg Mankiw desired clarification on the empirical evidence. He did not dispute that the natural rate of unemployment changes over time, but questioned Hall's certainty that monetary policy can affect the long-run level of unemployment.

Robert Barro suggested that Hall's interpretation of Alan Greenspan's approach is not accurate empirically, as there is clear evidence that the federal funds rate responds not only to inflation, but also to employment and other macroeconomic variables.

On the discussion of the natural-rate hypothesis, Lars Svensson maintained that potential output is a very useful concept both in theory and in practice. He allowed that, as it is an unobservable variable, the Fed can make mistakes in estimating it. His view, however, was that if the Fed had looked at Kalman-filter estimates of potential output in the 1970s, it would have realized that rising inflation meant a reduction in potential output.

Alberto Alesina contributed to the discussion on U.S. inflation history by saying that in looking for empirical evidence for or against the rules-vs.-discretion literature, it might be useful to look beyond the United States to the experience of other countries.

Alesina raised another issue of interest to several participants. He supported Stokey's view of reputation over that of the discussants. As evidence for the relevance of reputation in the real world, he cited the fact that major breakdowns in monetary rules and government default on debt are relatively rare in the developed world. He said this confirms the view that governments desire to maintain a reputation for repayment in order to be able to borrow again. He was of the opinion that multiplicity of reputational equilibria is not a crucial problem in the real world. He suggested that in the example of the Barro–Gordon model, the lowest possible level of inflation is an equilibrium that should be easy to coordinate on.

Nancy Stokey agreed that the dichotomy between rules and discretion can be overdrawn, and that reputation is important. She noted that while Argentina's currency board and peg to the dollar had implied a very strong rule for monetary policy, it was a rule the government had proved unable to maintain. On multiplicity of reputational equilibria, Stokey suggested that the role of the central bank is to be a cheerleader, selecting an equilibrium and persuading the private sector to behave accordingly. She added that empirical research into the means through which central banks do this would be useful.

Jonathan Parker drew attention to a little-known feature of many models with distortionary capital taxation. He explained that in these models, time-consistency problems can usually be eliminated by taxing capital a lot in the initial period. However, he noted that this result is a reversal of good and bad policy as economists usually see it, and that it is generally avoided by assuming that capitalism is better than socialism, or by focusing on stationary Markov-perfect equilibria. He commented that he would like to see a better foundation for the assumption that the government should not own all of capital, by adding to the model reasons why government is not good at running capital. He said that the interaction of this with time-consistency issues is an interesting direction for research over the next 25 years.

Alberto Alesina commented on the contention in Nancy Stokey's paper that while there might be many types of bad government, there is essentially only one type of good government. He proposed instead the view that there can be several types of good government, in particular in a nonrepresentative-agent world. For example, a "good" government representing the interests of capital will choose a different policy on capital and labor taxation from a "good" government representing the interests of labor.

Ken Rogoff questioned Lars Svensson's contention in his discussion that conservative central banks do not lead to higher output variability in practice. He suggested that Japan and Germany might be seen as counterexamples.

In conclusion, Nancy Stokey replied to Lars Svensson's comments on the observability issue in his discussion. She pointed out that in her simple model, the instrument is unobservable, but the target is observable. She commented that in a more realistic model, the observability problem would be whether central bankers are setting what they should be setting, a problem which would be made much more complex if they were trying to hit a moving target.