Rational Expectations of Government Policy: An Application of Newcomb's Problem*

ROMAN FRYDMAN
New York University
New York City

GERALD P. O'DRISCOLL, JR.
Federal Reserve Bank of Dallas
Dallas, Texas

ANDREW SCHOTTER
New York University
New York City

I. Introduction

Muth advanced the rational expectations hypothesis that agents predict the value of endogenous variables on the basis of the "relevant economic theory" of the phenomena in question. As long as there is a unique relevant theory, no ambiguity exists as to what is a rational expectation. But what happens when there is more than one relevant economic theory? In this case, unless we know which theory an agent believes, "the" rational expectation is not defined. This ambiguity takes on increased importance in the debate over policy-effectiveness. Sargent and Wallace [10] have argued that if the public holds rational expectations concerning the monetary authority's actions, systematic monetary policy will be unable to affect the value of any real variables in the economy.

In this paper we demonstrate by presenting a recently discovered decision-theoretic problem called Newcomb's problem [2] that if a decision maker (e.g., the monetary authority) must make one of several decisions in such a way that its payoff is affected by whether or not its action was anticipated by another rational agent with good predictive powers, then one uniquely best rational course of action or monetary policy may not exist.

*The authors benefitted from the helpful comments of James M. Buchanan, Clive Bull, Stephen DeCanio, Roy Gardner, Roger Garrison, Richard Langlois, Edmund S. Phelps, James B. Ramsey and Lester Telser.

Frydman's research was supported in part by a grant from the Spencer Foundation. O'Driscoll's research was supported in part by a grant from the Scaife Family Charitable Trusts, and was accomplished while he was at New York University. Schotter's research was supported in part by grant number N0014-78-0598 of the Office of Naval Research and in part by a grant from the Spencer Foundation. The authors gratefully acknowledge this support.

If, however, such a rational action or policy does not exist for the decision maker (monetary authority), it would be impossible for the agents (public) to possess rational expectations concerning this policy. This is true because a rational expectation must include within it the postulate that the decision maker will behave according to some underlying theory [8]; yet no uniquely best theory here exists for him to follow.

This problem arises only if the monetary authority actually believes that the public can anticipate its action with high accuracy. If it believes that the public is less able than that, then a uniquely best course of action for it does exist and its behavior can be predicted.

There also exists a connection between the problem considered here and the problem analyzed by Kydland and Prescott [6]. Formally, our point is preliminary to theirs. They concern themselves with the question whether consistent government policy can be optimal in the face of rational expectations held by the public [6, 475]. Our point is to question whether a consistent policy can even be formulated when the public holds rational expectations. Having established that no consistent policy may be formulated, we then question the existence of rational expectations.

We proceed as follows: In section II we discuss the decision problem that forms the basis of our analysis, called Newcomb's problem; in section III we apply the logic of the decision making problem discussed in section II to the problem of rational expectations of monetary policy; finally, in section IV we conclude our discussion by stating what we feel is the significance of our analysis and summarize our argument by presenting one proposition and its corollary.

II. Newcomb's Problem

Consider the following decision problem first described by William A. Newcomb and discussed by Gardner [2].[1]

Two closed boxes, B1 and B2, are on the table. B1 contains $1,000. B2 contains either $1 million or nothing. It is assumed that the decision maker does not know the contents of B2. The decision maker has an irrevocable choice between actions:

1. Take the contents of both boxes.
2. Take the contents of B2.

Assume that a superior Being has made a prediction about what the decision maker will decide and that the decision maker knows that the Being has "extremely good predictive powers."

If the Being expects the decision maker to choose both boxes, the Being leaves B2 empty. If He expects the decision maker to take only B2, He puts $1 million in it. (If He expects the decision maker to randomize his choice by, say, flipping a coin, the Being also leaves B2 empty.) In all cases, B1 contains $1,000. The decision maker understands the situation fully, the Being knows he understands, the decision maker knows that He knows,

1. Also see the March 1974 *Scientific American* issue, at which time philosopher Robert Nozick summarized reader's responses to the discussion of the paradox in the July 1973 issue.

Matrix I

		BEING	
		Move 1 (Predicts you take only B2)	Move 2 (Predicts you take B1 and B2)
DECISION MAKER	Move 1 (Take only B2)	1,000,000	0
	Move 2 (Take B1 and B2)	1,001,000	1,000

and so on. (In all that we say, we assume that utility is linear in money; thus the numbers below are utilities.)

The question is which box the decision maker should pick. The conundrum is that there is a strong argument for either decision, yet they cannot both be correct. To further explain the situation, we illustrate the decision problem by Matrix I.

If the decision maker takes B2 alone and that was predicted by the superbeing, then his payoff is \$1,000,000 (the Being having put that amount in B2). On the other hand, if the decision maker took B2 alone but the Being predicted he would take both boxes, then the superbeing would leave B2 empty; thus the payoff would be 0. Likewise, if he chooses both boxes and the Being predicted that he would choose only B2, the payoff would be \$1,001,000 (\$1,000 is always in B1 and the Being would have placed \$1,000,000 in B2 in this case). Finally, if the decision maker chooses both boxes and the Being predicted it, the payoff would only be \$1,000 (B2 being left empty).

Consider the decision maker's dilemma. If he feels that trying to outguess the superbeing is useless, the latter being practically omniscient, he must choose B2 only. As an illustration, assume that the Being anticipates correctly 9 out of 10 times. The expected utility (*EU*) from choosing B2 would be:

$$EU\,(B2 \text{ only}) = .9(1{,}000{,}000) + .1(0) = 900{,}000\,,$$

while the expected utility of choosing both boxes is:

$$EU\,(B1\ \&\ B2) = .1(1{,}001{,}000) + .9(1{,}000) + 101{,}000.$$

Consequently, in terms of the Expected Utility hypothesis, the decision to take B2 is clearly preferable. We can, however, view the problem in a different way.

If the Being has already predicted what the decision maker will do, then either he has placed \$1,000,000 in B2 or he has not. In this case, it makes sense to choose both boxes for the following reason: either the Being predicted the decision maker would take B2, in which case he put \$1,000,000 in B2, or he predicted the decision maker would take both boxes. In the first case, the payoff is \$1,001,000; in the second case, the payoff is \$1,000. But in either case, the decision to take both boxes dominates the decision to take only B2. This can be seen by the fact that row 2 is element by element greater than row 1, so that no matter what the superbeing predicts it is better to choose row 2. Consequently, from the perspective of the dominance principle, the decision to take both boxes is superior to that of taking only B2.

At this point the reader might start to question the robustness of the problem by questioning how sensitive it is to the actual numbers used in the matrix and the predictive powers of the superbeing. Surprisingly, the problem or decision dilemma exists for a variety of worlds. For instance, if we let $U(x_1)$, $U(x_2)$, $U(x_3)$ and $U(x_4)$ be the von Neumann-Morgenstern utility indices for the best, second best, third best, and least attractive outcomes, and if we let α be the probability that the superbeing guesses correctly, then simple algebra indicates that the dilemma holds as long as:

$$\alpha / 1-\alpha > (U(x_1) - U(x_4)) / (U(x_2) - U(x_3)).$$

(This inequality is independent of linear transformations of the utility index.) Consequently, for example, if $U(x_1) = 40$, $U(x_2) = 30$, $U(x_3) = 20$ and $U(x_4) = 10$, then the problem will exist as long as the decision maker thinks that the superbeing has even as little as a 75 percent chance of being correct. Hence, the superiority of the superbeing may be quite limited and yet the decision maker may still fall victim to Newcomb's problem.

The consequence of these considerations is that in Newcomb's problem or in situations isomorphic to it, there is a conflict between the expected utility hypothesis and the dominance principle. In such situations, while either decision can be justified, no rational action exists because no matter which choice is made there is a strong argument for choosing otherwise. Rational expectations may be "the same as the predictions of the relevant economic theory" [8, 315], but here there is a theory-conflict with no unambiguous method for deciding between theories. The choice that maximizes the decision maker's expected utility is dominated by the one that does not.

In the next section we argue that if the monetary authority believes that the public has good predictive powers (where good predictive powers in our model means a sufficiently high probability of guessing its actions; i.e., if $\alpha / 1-\alpha > (U(x_1) - U(x_4)) / (U(x_2) - U(x_3))$, then it becomes a decision maker caught in a Newcomb's-problem-like situation. Consequently, it has no rational course of action or policy prescription. We then investigate the consequences of this dilemma for the existence and consistency of rational expectations of government's action.

III. Newcomb's Problem and Monetary Policy

Consider a monetary authority who wishes to influence the level of two economic variables—the inflation rate $\dot{P}$ and the unemployment rate U, through adjusting the rate of growth of the money supply. Let us assume that this monetary authority faces a public, which we will model here as a monolith with one identical mind. The impact that monetary policy has on the existing $(\dot{P} - U)$ configuration will depend upon whether its actions are correctly anticipated by the public or not. Obviously, by analogy to Newcomb's problem of Section II, we are depicting the monetary authority here as the decision maker. The public is the superbeing, whose ability to predict the monetary authority's action is summarized by the probability with which the monetary authority thinks that the public will correctly anticipate its policy choice.

Let the monetary authority have two possible actions: increase the rate of growth of the money supply (M_H) or decrease the rate of growth of the money supply (M_L). What we want to describe is what effects these monetary policies will have on the existing $\dot{P} - U$ configuration as a consequence of whether they are anticipated or unanticipated by the public.

Consider the following simple macro model proposed by Lucas [7] and widely utilized in the literature on rational expectations:[2]

$$y_t^s = y_t^* + \alpha(P_t - P_t^e)\,;\ \alpha > 0 \tag{1}$$

$$y_t^d = m_t - P_t + V_t \tag{2}$$

$$y_t^s = y_t^d \tag{3}$$

where y_t^s, y_t are logs of aggregate supply and demand, y_t^* = log of "full-employment" or "natural" level of output, P_t = log of the price level, P_t^e = log of the price level *expected by the public*, m_t = log of the money stock V_t = constant − log of velocity.

Equation (1) is the Lucas supply function. Note that P_t^e is the price level expected by the public, since it is assumed that labor supply decisions determine the level of output supplied. Equation (2) is the simple quantity theory representation of aggregate demand. For simplicity we take $V_t = 0$.

Assuming that P_t is such as to equate y_t^s in (1) to y_t^d in (2), we obtain:

$$P_t = (1/1{+}\alpha)\, m_t + (\alpha/1{+}\alpha)\, P_t^e - (1/1{+}\alpha)\, y_t^*\ . \tag{4}$$

It is usually assumed that the public knows the coefficients of the model and therefore in forming a rational expectation of P_t it uses equation (4):

$$P_t^e = (1/1{+}\alpha)\, m_t^e + (\alpha/1{+}\alpha)\, P_t^e - (1/1{+}\alpha)\, y_t^*\ . \tag{5}$$

Solving (5) for P_t^e yields:

$$P_t^e = m_t^e - y_t^*\ . \tag{6}$$

Equation (6) makes it clear that the expectation of m_t^e is crucial for the formation of rational expectation of P_t. Hence forming rational expectations in this model includes forming a rational expectation of the monetary authority's policy action. Using (5) in (4):

$$P_t = (1/1{+}\alpha)\, m_t + (\alpha/1{+}\alpha)\, m_t^e - y_t^*\ . \tag{7}$$

Finally, using (6) and (7) in (1) yields:

$$y_t = y_t^* + (1/1{+}\alpha)\,(m_t - m_t^e). \tag{8}$$

To facilitate further discussion we assume that the deviations $y_t - y_t^*$ are related to unemployment via Okun's Law:

$$y_t - y_t^* = k(U_t^* - U_t) \tag{9}$$

2. We use this particular model because it is widely used in the rational expectations literature. In models of this class, expectations of monetary policy are often assumed to be rational.

where $k > 0$ is a constant. Equation (9) assumes that when $y_t = y_t^*$ unemployment is equal to natural rate U_t^*. Using (9), we can rewrite (8) as:

$$U_t = U_t^* + \beta(m_t^e - m_t) \tag{10}$$

where $\beta = 1/k(1+\alpha)$, or

$$U_t = U_t^* + \beta[(m_t^e - m_{t-1}) - (m_t - m_{t-1})].$$

We can now use this model to analyze how an existing $\dot{P} - U$ configuration will change when the government embarks on one of its two available monetary policies. We examine the effects of each policy for the case when it is correctly or incorrectly anticipated by the public. To do this assume that at time $t - 1$ the system is at the natural rate of unemployment U_t^* and the steady state inflation rate at that unemployment level is $\dot{P}^*$ with initial price level P_{t-1}^*. Using equations (10), (2), (6) and (7) we can construct the following table illustrating the consequences of the government's monetary action as a function of its being correctly or incorrectly anticipated by the public.

These outcomes can be depicted in Figure 1.

In this diagram we depict the decision problem of the monetary authority as one in which the status quo is point (x_0) through which pass two Phillips curves labeled PH' and PH''. The vertical Phillips curve PH' represents the locus of ($\dot{P} - U$) combinations that result when the monetary authority either increases or decreases the rate of monetary expansion and that action is correctly anticipated by the public. It is usually considered to be the long-run Phillips curve but we shall call it the "correctly anticipated Phillips curve." The curve PH'' represents what we shall call the "misperceived Phillips curve" since it represents that locus of ($\dot{P} - U$) combinations that would result when the public *incorrectly anticipates* the monetary action of the government and predicts an increase in the rate of growth of the money supply when the authority actually decreases its rate of expansion and vice versa.

Note that the PH'' curve is not a short-run Phillips curve, since that curve is drawn under the assumption that the public has static expectations of the monetary authority's actions. A true short-run Phillips curve traces out the results of *unanticipated* changes in the rate of growth of the money supply, while our curve traces out the consequences of *incorrectly anticipated* monetary actions. Hence we are calling it the "misperceived Phillips curve."

Table I

Policy			
Expected by the Public	Actual	Description of Outcomes	
M_L	M_H	$P_t > P_t^e$, $U_t < U_t^*$, $P_t - P_{t-1}^* > \dot{P}^*$	(x_1)
M_L	M_L	$P_t = P_t^e$, $U_t = U_t^*$, $P_t - P_{t-1}^* < \dot{P}^*$	(x_2)
M_H	M_H	$P_t = P_t^e$, $Ut = U_t^*$, $P_t - P_{t-1}^* > \dot{P}^*$	(x_3)
M_H	M_L	$P_t < P_t^e$, $U_t > U_t^*$, $P_t - P_{t-1}^* < \dot{P}^*$	(x_4)

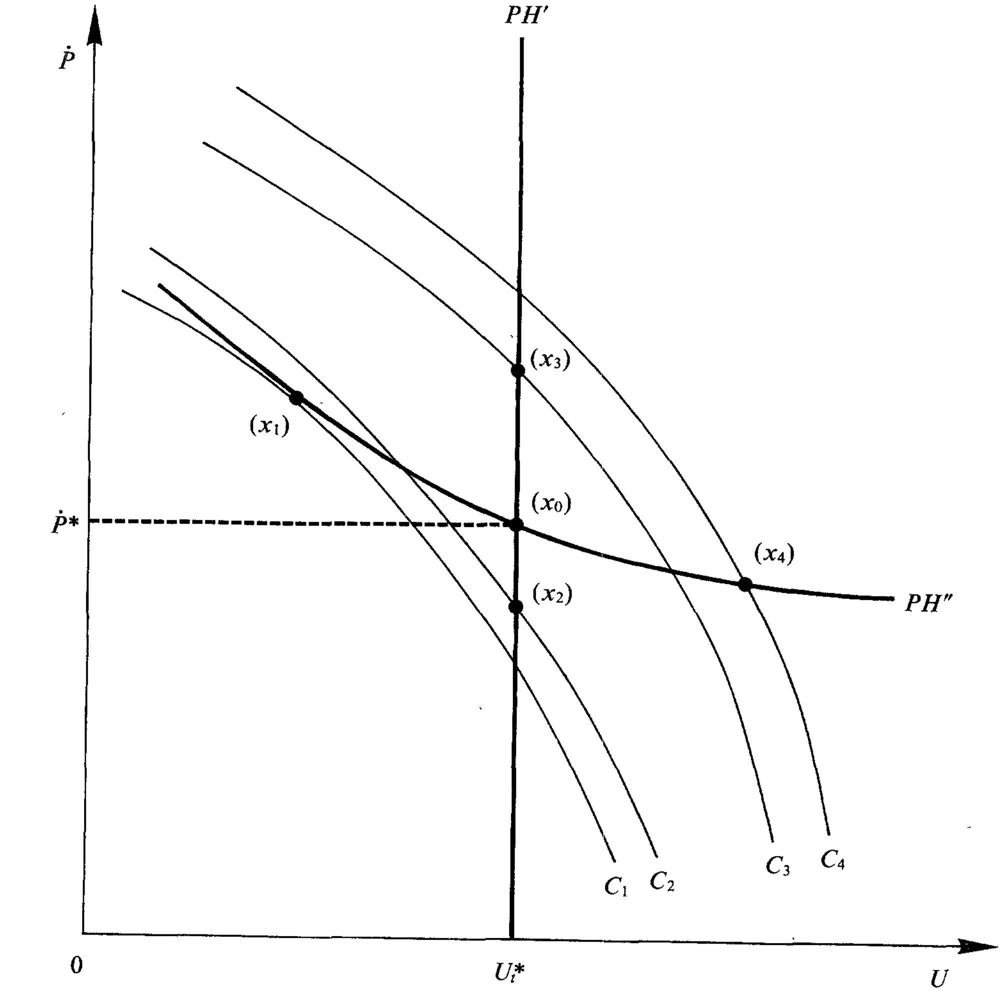

Figure 1.

The preferences of the monetary authority are given by the concave (to the origin) curves C_1, C_2, C_3, C_4, where curves closer to the origin represent higher levels of utility.

(x_1), (x_2), (x_3) and (x_4) are the points described in Table I and depict the consequences of correctly and incorrectly anticipated monetary policy.

Given the preferences of the monetary authority depicted in the diagram it is clear that:

$$U(x_1) > U(x_2) > U(x_3) > U(x_4),$$

where $U(x_i)$ is the von Neumann-Morganstern utility associated with the outcome x_i, $i = 1,2,3,4$.

Our analysis can now be represented by the following decision matrix, where the payoffs in each cell are the payoffs to the monetary authority:

Matrix II

		PUBLIC: Anticipate M_L	PUBLIC: Anticipate M_H
MONETARY AUTHORITY	Choose M_L	$U(x_2)$	$U(x_4)$
	Choose M_H	$U(x_1)$	$U(x_3)$

At this point Newcomb's problem appears. If the monetary authority feels that the public is a sufficiently good predictor of its actions (i.e., if it thinks that $\alpha / 1-\alpha > (U(x_1) - U(x_4)) \,/\, (U(x_2) - U(x_3))$, then by the expected utility theory it should clearly choose to decrease the rate of growth of the money supply (action M_L) since the expected return it can expect from this action is greater than what it can expect to get from action M_H. However, M_H dominates M_L. Consequently, the same theory conflict exists here as exists in Newcomb's problem and no rational course of action exists for the monetary authority. Hence, no rational prediction of its policy action exists.

It should also be emphasized (see equation 6) that for the public to form rational price expectations, it must form rational expectations about monetary policy. Hence, in this standard rational expectations model, the public's inability to form rational expectations of monetary policy makes it unable to form rational expectations about any other relevant variable in the model.

IV. Conclusions

The argument presented in this paper can be summarized by the following proposition and its corollary:

Proposition 1: In the model presented above, if the government has a finite number of monetary actions that it can take (M_L and M_H) and if the utility consequences of its actions can be depicted by matrix II and Figure 1, then if the monetary authority believes that the public can predict its behavior with "good predictive power" (i.e., if $\alpha / 1-\alpha > (U(x_1) - U(x_4)) \,/\, (U(x_2) - U(x_3))$, no rational expectation of the government action can be formed.

Proof:
If $\alpha / 1-\alpha > (U(x_1) - U(x_4)) \,/\, (U(x_2) - U(x_3))$ then $\alpha U(x_2) + (1 - \alpha)\, U(x_4) > \alpha U(x_3) + (1 - \alpha)\, U(x_1)$ and the expected utility of M_L is greater than the expected utility of M_H. But M_H dominates M_L by construction. Hence by Newcomb's problem no uniquely best theory of behavior exists for monetary authority. Since by definition, rational expectations are expectations based on *the* relevant theory of the phenomena, if no such theory exists, rational expectations cannot be formed. Q.E.D.

Corollary: In the model discussed above, rational expectations of government policy are inconsistent.

Proof:

Rational expectations of government policy by the public imply, by definition, that $\alpha/1-\alpha > (U(x_1) - U(x_4))/(U(x_2) - U(x_3))$. But $\alpha/1-\alpha > (U(x_1) - U(x_4))/(U(x_2) - U(x_3))$ implies "good predictive powers" on the part of the public which, by the Proposition imply that no rational expectation can be formed. Q.E.D.

The reason why this problem exists in our model stems from a common problem involved in forming rational expectations in models in which one or more of the endogenous variables are exclusively controlled by one or at most several decision makers. In such situations, a rational expectation about the level of such a variable becomes a prediction of the behavior of the decision maker who controls its value. The problem is that if the decision maker's payoffs depend upon whether its behavior is anticipated or not, then the decision maker will be forced to enter into a series of higher order expectations in his effort to try to second-guess those agents who are forming rational expectations over his behavior. The knowledge that others are trying to predict a decision maker's behavior in a systematic (i.e., rational) way may force the decision maker to behave non-systematically and hence eliminate any hope of forming a rational expectation about his behavior.[3, 4]

3. Such problems are less likely to arise in models in which no endogenous variable is under the control of a single decision maker, such as [3; 4; 5; 7]. Different types of decision-theoretic problems can, however, arise in these models [1].

4. Similar problems in stabilization policy were investigated in a stimulating paper by Phelps [9].

References

1. Frydman, Roman, "Towards an Understanding of Market Processes: Individual Expectations, Learning and Covergence to Rational Expectations Equilibrium". Forthcoming in *American Economic Review*, 1982.

2. Gardner, Martin, "Mathematical Games." *Scientific American*, July 1973, 104-08.

3. Grossman, Sanford J., "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information." *Journal of Finance*, May 1976, 573-85.

4. ——— . "The Existence of Futures Markets, Noisy Rational Expectations, and Informational Externalities." *Review of Economic Studies*, October 1977, 431-49.

5. Grossman, Sanford J. and Stiglitz, Joseph E., "Information and Competitive Price Systems." *American Economic Review*, May 1976, 246-53.

6. Kydland, Finn E. and Prescott, Edward C., "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, June 1977, 473-91.

7. Lucas, Robert E., Jr., "Some International Evidence on Output-Inflation Tradeoffs." *American Economic Review*, June 1973, 326-34.

8. Muth, John F., "Rational Expectations and the Theory of Price Movements." *Econometrica*, July 1961, 315-35.

9. Phelps, Edmund S. "Obstacles to Curtailing Inflation," in *Essays in Post-Keynesian Inflation*, edited by James H. Gapinski and Charles E. Rockwood. Cambridge, Mass.: Ballinger, 1979, pp. 179-93.

10. Sargent, Thomas J. and Wallace, Neil. *Rational Expectations and the Theory of Economic Policy*, Studies in Monetary Economics. Minneapolis: Federal Reserve Bank of Minneapolis Research Department, 1975.