Determinacy and Identification with Taylor Rules

John H. Cochrane

University of Chicago

The new-Keynesian, Taylor rule theory of inflation determination relies on explosive dynamics. By raising interest rates in response to inflation, the Fed induces ever-larger inflation, unless inflation jumps to one particular value on each date. However, economics does not rule out explosive inflation, so inflation remains indeterminate. Attempts to fix this problem assume that the government will choose to blow up the economy if alternative equilibria emerge, by following policies we usually consider impossible. The Taylor rule is not identified without unrealistic assumptions. Thus, Taylor rule regressions do not show that the Fed moved from "passive" to "active" policy in 1980.

I. Introduction

How is the price level determined? The new-Keynesian, Taylor rule theory provides the current standard answer to this basic economic question. In this theory, inflation is determined because the central bank systematically raises nominal interest rates more than one-for-one with inflation. This "active" interest rate target is thought to eliminate the

I gratefully acknowledge research support from the Center for Research in Security Prices and the National Science Foundation via a grant administered by the National Bureau of Economic Research. I am grateful for many comments and suggestions in this paper's long gestation. Among many others, I thank five referees, Fernando Alvarez, Marco Bassetto, David Backus, Florin Bilbiie, Mark Gertler, Eric Leeper, Lars Hansen, Peter Ireland, Henrik Jensen, Pat Kehoe, Sophocles Mavroeidis, Edward Nelson, Yoshio Nozawa, Monika Piazzesi, Stephanie Schmitt-Grohé, Christopher Sims, Eric Sims, Robert Shimer (the editor), Lars Svensson, Martín Uribe, Michael Woodford, and participants at the spring 2005 Indiana University Monetary Economics Conference, the fall 2006 NBER Economic Fluctuations and Growth Conference, and the University of Chicago Money Workshop. Online Appendix B is available on the *Journal* Web site and on my Web page (http://faculty.chicagobooth.edu/john.cochrane/research/Papers/).

[*Journal of Political Economy*, 2011, vol. 119, no. 3]
 0022-3808/2011/11903-0005$10.00

indeterminacy that results from fixed interest rate targets and thus to provide the missing "nominal anchor."

Theories ultimately rise and fall on their ability to organize and interpret facts. Keynes wrote the *General Theory* of the Great Depression. Friedman and Schwartz wrote the *Monetary History of the United States*. The central new-Keynesian story is that U.S. inflation was conquered in the early 1980s by a change from a "passive" policy in which interest rates did not respond sufficiently to inflation to an "active" policy in which they do so. Most famously, Clarida, Galí, and Gertler (2000) ran regressions of federal funds rates on inflation and output. They found inflation coefficients below one up to 1980 and above one since then. (Complex new-Keynesian models also "fit the data" well, but so do other models. This observation is not a useful test of a model's basic structure.)

I analyze this theory and its interpretation of the data. First, I conclude that the Taylor rule, in the context of a new-Keynesian model, leaves the same inflation indeterminacy as with fixed interest rate targets. Second, even accepting the model, I show that the parameters of the Fed's policy rule are not identified, so regression evidence does not say anything about determinacy in a new-Keynesian model.

The same key point drives both observations: new-Keynesian models do not say that higher inflation causes the Fed to raise real interest rates, which in turn lowers "demand," which reduces future inflation. That's "old-Keynesian," stabilizing logic. In new-Keynesian models, higher inflation leads the Fed to set interest rates in a way that produces even higher future inflation. For only one value of inflation today will inflation fail to explode or, more generally, eventually leave a local region. Ruling out nonlocal equilibria, new-Keynesian modelers conclude that inflation today must jump to the unique value that leads to a locally bounded equilibrium path.

But there is no reason to rule out nominal explosions or "nonlocal" nominal paths. Transversality conditions can rule out *real* explosions but not *nominal* explosions. Since the multiple nonlocal equilibria are valid, the new-Keynesian model does not determine inflation.

Furthermore, if we do rule out the nonlocal paths, interest rates that generate explosive inflation are an outcome that is never realized in the observed equilibrium, so that response cannot be measured.

A. Responses: Determinacy and Dilemma

Many authors have advanced proposals to trim new-Keynesian multiple equilibria by adding additional provisions to the policy description, describing actions that the government would take if the undesired equilibrium were to occur. I analyze these proposals, asking several questions: Do they, in fact, rule out the undesired equilibrium? Many do not.

Unpleasant outcomes, such as infinite inflation, can be an equilibrium. Does the future policy lead people to any change in behavior today? In many cases, the answer is no. Knowledge of the future policy and its outcome do not change consumption or asset demands or give any supply-demand pressure toward a different inflation rate today. In a game, a "blow-up-the-world" threat can induce the other player to change earlier behavior. But here the private sector is atomistic. Is the proposal an even vaguely plausible description of what people currently believe our government would do, and not wildly at odds with what governments actually do in similar circumstances? Or is it a suggestion for commitments that future governments might make? We need the former case in order to use the theory as a positive description of current data.

Many proposals to trim equilibria sound superficially like sensible descriptions of what governments do to stop extreme inflation or deflation: switch to a commodity standard, exchange rate peg, or money growth rule, or undertake a fiscal expansion or reform. However, *stopping an inflation* or deflation is completely different from *disallowing an equilibrium.* If an inflation-stopping policy still describes how an equilibrium forms at each date, then the inflation or deflation and its end remain an equilibrium path and we have ruled nothing out.

To rule out equilibria, the government must specify policy so that it is impossible for an equilibrium to form somewhere along the path. Some proposals specify a commodity standard, which implies zero inflation, but *also* a very high interest rate. Others specify a commodity standard but also a limit on money supply that precludes the price level set by the commodity standard. Still others specify inconsistent fiscal and monetary policies or introduce arbitrage opportunities. It is these inconsistent or overdetermined policies, not the inflation or deflation stabilization, that trim equilibria.

Such assumptions seem wildly implausible, as descriptions of government behavior and especially as descriptions of people's current beliefs about government behavior. A policy configuration for which "no equilibrium can form" or "private first-order conditions cannot hold" means a threat to blow up the economy. Furthermore, in these models, there are policies that the government *can* follow that stop the inflation without blowing up the economy, allowing an equilibrium to form at each date. Why would a government choose to blow up the economy when tested policies that stop inflation or deflation, while allowing equilibria to form at each date, are sitting on the shelf? Actual governments that stop inflations do not also insist that the real quantity of money remain at a low level, do not try to target hyperinflationary interest rates, do not introduce arbitrage opportunities, and do coordinate fiscal and monetary policies.

In fact, in most (Ramsey) analyses of policy choices, we label such policy configurations as "impossible," not just "implausible." We think of governments choosing policy configurations while taking private first-order conditions as constraints; we think of governments acting in markets. We do not think governments *can* set policies for which private first-order conditions do not hold. For example, to operate a commodity standard, we would say that a government must offer to exchange currency for the commodity freely at the stated price; it simply *cannot* also maintain a low limit on money supply or a very high interest rate target.

The logical dilemma is unavoidable. If we specify that the government will stop an inflation or deflation in such a way that equilibria can form on each date, we get quite sensible proposals and descriptions of what governments might do, can do, and have done to end inflations or deflations, but we do not rule out any equilibria. To rule out equilibria, people must believe that the government will choose to blow up the economy. Whether the rest of the policy description resembles historically successful stabilizations is irrelevant. Whether the impossible policies occur on the date of stabilization or at any other point on the path is irrelevant. I survey the extensive literature and do not find any successful escape from this dilemma.

There is an important distinction here between "eliminating multiple equilibria" and "defining one equilibrium." The government does set policies for which market-clearing conditions may not hold at off-equilibrium prices. For example, in a commodity standard, there is an arbitrage opportunity if the market price differs at all from the government price. This policy gives a strong supply-demand force toward the equilibrium price. But there is nothing infeasible or incredible about a commodity standard. Non-Ricardian fiscal commitments work the same way.

B. Responses: Identification

The literature also contains many attempts to rescue identification. But we can and must ask whether identification assumptions are reasonable, as a description of Fed behavior, of people's expectations of Fed behavior, and of the theory with which the regressions are interpreted.

The central identification problem is that the theory predicts there is no movement in the crucial right-hand variable, the difference between actual inflation and inflation in the desired equilibrium. (Here, too, the issue is not "in" vs. "out of" equilibrium; the issue is selection between multiple equilibria.) At a deep level, then, we must assume that the correlations between interest rates and inflation *in* the equilibrium are the same as the Fed's unobservable interest rate response to movements of inflation *away* from that equilibrium. In the theory, the

right "natural rate," the behavior of interest rates *in* the desired equilibrium, is a completely different issue from determinacy, how the interest rate should vary if inflation veers *away* from the desired equilibrium. To identify the latter from the former, we must assume they are the same.

But then a second classic problem arises. In the desired equilibrium, the Taylor rule right-hand variables (inflation, output) and all potential instruments are correlated with the monetary policy disturbance term. This correlation is central to the theory: If a monetary policy shock occurs, then inflation and other right-hand variables are supposed to jump to the unique values that lead to a locally bounded equilibrium.

Furthermore, new-Keynesian theory also advocates a "stochastic intercept": the central bank should vary the interest rate in response to structural (IS, cost, etc.) disturbances in order to follow variations in the natural rate. These interest rate movements become part of the empirical monetary policy disturbance. Therefore, the theory predicts that the structural disturbances to other equations, and endogenous variables that depend on them, cannot be used as instruments.

Lags do not help either. If the structural disturbances are serially correlated, lagged endogenous variables are correlated with the monetary policy error term. If the structural disturbances are not serially correlated, lagged endogenous variables are uncorrelated with the right-hand side of the monetary policy rule.

In sum, new-Keynesian models specify policy rules that are a snake pit for econometricians. There is no basis for all the obvious devices, such as excluding variables from the policy rule, using instruments, assuming that the right-hand variables of policy rules are orthogonal to the disturbance, or restricting lag length of disturbances. (Lag-length and exclusion restrictions as approximations are not a big problem; restrictions to produce identification are.) Not only *might* these problems exist, but theory predicts that most of them *do* exist. Empiricists must throw out important elements of the theory in order to identify parameters.

Finally, even if one could identify parameters from a determinate new-Keynesian equilibrium (1980s), what does one measure if the world is indeterminate, as supposedly was the case in the 1970s? The *change* in coefficients is a crucial part of the story, and one must measure the earlier coefficient to measure a change.

C. If Not This, Then What?

If not this theory, what theory can account for price-level determination in a modern fiat money economy whose central bank follows an interest rate target? This paper is entirely negative, and long enough, so I do

not exposit or test an alternative theory. But it is worth pointing out a possibility.

The valuation equation for government debt states that the real value of nominal debt equals the present value of real primary surpluses. The new-Keynesian Taylor rule model fulfills this equilibrium condition by assuming that the government will always adjust taxes or spending ex post to validate any change in the price level. If deflation doubles the real value of nominal debt, the government doubles taxes to pay off that debt. It is an "active money, passive fiscal" regime, in Leeper's (1991) terminology.

The active fiscal, passive money regime is an alternative possibility. In this case, the valuation equation for government debt determines the price level, and the central bank follows an interest rate rule that does not destabilize the economy. Since this model of price-level determination relies on ruling out real rather than nominal explosions through the consumer's transversality condition, it uniquely determines the price level. This model is not inconsistent with empirical Taylor rule regressions. It therefore provides a coherent economic theory of the price level that can address current institutions.

This paper is not a criticism of new-Keynesian economics in general. In particular, I do not have anything to say here that criticizes its basic ingredients: an intertemporal, forward-looking "IS" curve or an intertemporally optimizing, forward-looking model of price setting subject to frictions, as captured in the "new-Keynesian Phillips curve." The passive-money, active-fiscal regime of such a model can determine inflation.

D. An Acknowledgement

Indeterminacy, multiple equilibria, and identification in dynamic rational expectations models are huge literatures that I cannot possibly adequately cite, acknowledge, or review in the space of one article. The body of the paper reviews specific important contributions in the context of new-Keynesian models. This is not a critique of those specific papers. I choose these papers as concrete and well-known examples of general points, repeated throughout the literature. Appendix A and online Appendix B contain a more comprehensive review, both to properly acknowledge others' efforts and to establish that no, these problems have not been solved.

The equations in this paper are simple and not new. In this field, however, there is great debate over how one should read and interpret simple and fairly well-known equations. This paper's novelty is a contribution to that difficult enterprise.

II. Simplest Model

We can see the main points in a very simple model consisting only of a Fisher equation (consumer first-order conditions) and a Taylor rule describing Fed policy,

$$i_t = r + E_t\pi_{t+1}, \tag{1}$$

$$i_t = r + \phi\pi_t + x_t, \tag{2}$$

where i_t is the nominal interest rate, π_t is inflation, and r is the constant real rate.

The monetary policy disturbance x_t represents variables inevitably left out of any regression model of central bank behavior, such as responses to financial crises, exchange rates, time-varying rules, mismeasurement of potential output, and so on. It is not a forecast error, so it is serially correlated:

$$x_t = \rho x_{t-1} + \varepsilon_t. \tag{3}$$

(Equivalently, the target may be smoothed and react to past inflation.)

We can solve this model by substituting out the nominal interest rate, leaving the equilibrium condition

$$E_t\pi_{t+1} = \phi\pi_t + x_t. \tag{4}$$

A. *Determinacy*

Equation (4) has many solutions. We can write the equilibria of this model as

$$\pi_{t+1} = \phi\pi_t + x_t + \delta_{t+1}; \quad E_t(\delta_{t+1}) = 0, \tag{5}$$

where δ_{t+1} is any conditionally mean zero random variable. Multiple equilibria are indexed by arbitrary initial inflation π_0 and by the arbitrary random variables or "sunspots" δ_{t+1}. This observation forms the classic doctrine (Sargent and Wallace 1975) that inflation, to say nothing of the price level, is indeterminate with an interest rate target.

If $\|\phi\| > 1$, all of these equilibria except one eventually explode; that is, $E_t(\pi_{t+j})$ grows without bound. If we disallow such solutions, then a unique locally bounded solution remains. We can find this solution by solving the difference equation (4) forward or by undetermined coefficients (which assumes a bounded solution, depending only on x_t):

$$\pi_t = -\sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} E_t(x_{t+j}) = -\frac{x_t}{\phi - \rho}. \tag{6}$$

Equivalently, by this criterion we select the variables π_0, $\{\delta_{t+1}\}$, which index multiple equilibria, as

$$\pi_0 = -\frac{x_0}{\phi - \rho}; \quad \delta_{t+1} = -\frac{\varepsilon_{t+1}}{\phi - \rho}. \tag{7}$$

Thus we have it: if the central bank's interest rate target reacts sufficiently to inflation—if $\|\phi\| > 1$—then it seems that a pure interest rate target, with no control of monetary aggregates, no commodity standard or peg, and no "backing" beyond pure fiat, can determine at least the inflation rate, if not quite the price level. It seems that making the peg react to economic conditions overturns the classic doctrine that inflation is indeterminate under an interest rate peg (McCallum 1981).

But what's wrong with nonlocal equilibria? Transversality conditions can rule out real explosions, but not nominal explosions. Hyperinflations are historic realities. This condition did not come from any economics of the model. I conclude that there is nothing wrong with them, and this model does not determine inflation.

This is an example that needs fleshing out. First, I need to write down a fully specified model, to show that there truly is nothing wrong with nonlocal equilibria. Second, I need to examine the standard three-equation model, including varying real rates and price stickiness, to verify that this simple frictionless model indeed captures the same issues. Third, haven't the legions of people who have addressed these issues solved all these problems? I review the literature to verify that they have not done so.

This simple example also makes clear the stark difference between "indeterminacy" and "inflationary and deflationary spirals" and the difference between "determinacy" and the "stabilizing" stories, common in policy analysis and Federal Reserve statements. Authors at least since Friedman (1968) have worried that if the Fed follows an interest rate target, inflation could rise, real rates would fall (for Friedman, money growth would increase), this would cause higher future inflation, and the spiral would continue. Many analyses of the 2008–11 situation worry about an opposite deflationary spiral, especially with nominal interest rates stuck at zero. Many explanations of the Taylor rule say that it cuts off such spirals: nominal interest rates rise more than inflation, so real rates rise, which cools off future inflation.

All of this is "old-Keynesian" logic. Whether right or wrong, the issues are completely different. The spirals describe a single but undesirable equilibrium. The Taylor rule induces a stable root, not an unstable root, to the system dynamics. All of these stories require at least nominal effects on real interest rates or price stickiness absent in this analysis. As King (2000) emphasizes, $\phi < 1$, oscillating hyperinflation and deflation, works just as well as $\phi > 1$ to ensure determinacy. That example is hard to describe by "stabilizing" intuition.

B. Identification

Now, suppose the solution (6) is in fact correct; if so, what are its observable implications? Since π_t is proportional to x_t, the dynamics of equilibrium inflation are simply those of the disturbance x_t,

$$\pi_t = \rho\pi_{t-1} + w_t \tag{8}$$

($w_t \equiv -\varepsilon_t/[\phi - \rho]$, but ε_t and x_t are not directly observed, so we can summarize observable dynamics with the new error w_t). Using (1) and (8), we can find the equilibrium interest rate

$$i_t = r + \rho\pi_t. \tag{9}$$

Equation (9) shows that *a Taylor rule regression of i_t on π_t will estimate the disturbance serial correlation parameter ρ rather than the Taylor rule parameter ϕ.*

What happened to the Fed policy rule, equation (2)? The solution (6) shows that the right-hand variable π_t and the disturbance x_t are correlated—perfectly correlated in fact. That correlation is no accident or statistical assumption; it is central to how the model behaves. The whole point of the model, the whole way it generates responses to shocks, is that endogenous variables (π_t) "jump" in response to shocks (ε_t) so as to head off expected explosions.

Perhaps we can run the regression by instrumental variables? Alas, the only instruments at hand are lags of π_t and i_t, themselves endogenous and thus invalid. For example, if we use all available lagged variables as instruments, we have from (8) and (9)

$$E(\pi_t|\pi_{t-1},\ i_{t-1},\ \pi_{t-2},\ i_{t-2},\ \ldots) = \rho\pi_{t-1},$$

$$E(i_t|\pi_{t-1},\ i_{t-1},\ \pi_{t-2},\ i_{t-2},\ \ldots) = r + \rho^2\pi_{t-1}.$$

Thus the instrumental variables regression gives exactly the same estimate:

$$E(i_t|\pi_{t-1},\ i_{t-1},\ \pi_{t-2},\ i_{t-2},\ \ldots) = r + \rho E(\pi_t|\pi_{t-1},\ i_{t-1},\ \pi_{t-2},\ i_{t-2},\ \ldots).$$

If the disturbance x_t were independently and identically distributed (i.i.d.), then the correlation of instruments with errors would be removed, but so would the correlation of instruments with right-hand variables.

Is there nothing clever we can do? No. The equilibrium dynamics of the observable variables $\{i_t, \pi_t\}$ are completely described by equations (8) and (9). The equilibrium dynamics, and the resulting likelihood function, do not involve ϕ: *ϕ is not identified from data on $\{i_t, \pi_t\}$ in the equilibrium of this model.* Inflation is supposed to jump to the one value for which accelerating inflation at rate ϕ is not observed. If inflation does jump, there is no way to measure how fast the inflation would accelerate if it did not jump.

Absence of ϕ from equilibrium dynamics and the likelihood function means that we cannot even test whether the data are generated from the *region* of determinacy $\|\phi\| > 1$, abandoning hope of measuring the precise value of ϕ, as Lubik and Schorfheide (2004) try to do. For every equilibrium generated by a ϕ^* with $\|\phi\| > 1$, the same equilibrium dynamics (8) and (9) can be generated by a different ϕ with $\|\phi\| < 1$. Online Appendix B elaborates this point.

Again, this is the beginning. I need to show that the same problems occur in more complex models, including the standard three-equation new-Keynesian model that Clarida et al. (2000) and other authors use, and that the many attempts at identification do not convincingly surmount them.

III. An Explicit Frictionless Model

A. *The Model*

To keep the discussion compact and consistent with the literature, I simplify standard sources, Benhabib, Schmitt-Grohé, and Uribe (2002) and Woodford (2003). Consumers maximize a utility function

$$\max E_t \sum_{j=0}^{\infty} \beta^j u(C_{t+j}).$$

Consumers receive a constant nonstorable endowment $Y_t = Y$; markets clear when $C_t = Y$. Consumers trade in complete financial markets described by real contingent claims prices $m_{t,t+1}$ and hence nominal contingent claims prices,

$$Q_{t,t+1} = \frac{P_t}{P_{t+1}} m_{t,t+1}.$$

The nominal interest rate is related to contingent claim prices by

$$\frac{1}{1 + i_t} = E_t(Q_{t,t+1}).$$

The government issues one-period nominal debt; B_{t-1} denotes the face value issued at time $t - 1$ and coming due at date t. The government levies lump-sum taxes S_t, net of transfers. The term S_t denotes the real primary surplus. I follow Benhabib et al. (2002), Woodford (2003), Cochrane (2005), and many others in describing a frictionless economy. The dollar can be a unit of account even if, in equilibrium, nobody chooses to hold any dollars overnight.

The consumer faces a present-value budget constraint

$$E_t \sum_{j=0}^{\infty} Q_{t,t+j} P_{t+j} C_{t+j} = B_{t-1} + E_t \sum_{j=0}^{\infty} Q_{t,t+j} P_{t+j} (Y_{t+j} - S_{t+j}) \qquad (10)$$

or, in real terms,

$$E_t \sum_{j=0}^{\infty} m_{t,t+j} C_{t+j} = \frac{B_{t-1}}{P_t} + E_t \sum_{j=0}^{\infty} m_{t,t+j}(Y_{t+j} - S_{t+j}). \tag{11}$$

B. *Equilibria*

The consumer's first-order conditions state that marginal rates of substitution equal contingent claims price ratios, and equilibrium $C_t = Y$ implies a constant real discount factor

$$\beta \frac{u_c(C_{t+1})}{u_c(C_t)} = m_{t,t+1} = \beta \frac{u_c(Y)}{u_c(Y)} = \beta. \tag{12}$$

Therefore, the real interest rate is constant,

$$\frac{1}{1+r} = E_t(m_{t,t+1}) = \beta,$$

and the nominal discount factor is

$$Q_{t,t+1} = \frac{P_t}{P_{t+1}} m_{t,t+1} = \beta \frac{P_t}{P_{t+1}}. \tag{13}$$

The interest rate follows a Fisher relation,

$$\frac{1}{1+i_t} = E_t(Q_{t,t+1}) = \beta E_t\left(\frac{P_t}{P_{t+1}}\right) = \frac{1}{1+r} E_t\left(\frac{1}{\Pi_{t+1}}\right). \tag{14}$$

The usual relation (1) follows by linearization.

From the consumer's present-value budget constraint (10) and with contingent claim prices from (13), equilibrium $C_t = Y$ also requires

$$\frac{B_{t-1}}{P_t} = \sum_{j=0}^{\infty} \frac{1}{(1+r)^j} E_t(S_{t+j}). \tag{15}$$

The value of government debt is the present value of future net tax payments. This is not a "government budget constraint"; it is an equilibrium condition, an implication of supply = demand or $C_t = Y_t$ in goods markets, as you can see directly by looking at (11). Cochrane (2005) discusses this issue in more detail. I assume that the present value of future primary surpluses is positive and finite:

$$0 < \sum_{j=0}^{\infty} \frac{1}{(1+r)^j} E_t(S_{t+j}) < \infty.$$

The Fisher equation (14) and the government debt valuation equation (15) are the only two conditions that need to be satisfied for the price sequence $\{P_t\}$ to represent an equilibrium. If they hold, then the allocation $C_t = Y$ and the resulting contingent claims prices (13) imply that markets clear and the consumer has maximized subject to his budget constraint. The equilibrium is not yet unique in that many different

price or inflation paths will work. Unsurprisingly, we need some specification of monetary and fiscal policy to determine the price level.

C. *New-Keynesian Policy and Multiple Equilibria*

The new-Keynesian Taylor rule analysis maintains a "Ricardian" fiscal regime; net taxes S_{t+j} are assumed to adjust so that the government debt valuation equation (15) holds given any price level (Woodford 2003, 124). It also specifies a Taylor rule for monetary policy.

We have answered the first question needed from this explicit model: Yes, solutions of the simple model consisting of a Fisher equation and a Taylor rule (1)–(2), as I studied above, do in fact represent the full set of (linearized) equilibrium conditions of this explicit model, if we assume a Ricardian fiscal regime. My simple model did not leave anything out.

Are the nonlocal equilibria really globally valid? Here I follow the standard sources, in part to emphasize agreement that they are (Benhabib et al. 2002; Woodford 2003, chap. 2.4, starting on 123, and chap. 4.4, starting on 311, with a review).

Restrict attention to perfect-foresight equilibria. Adding uncertainty (sunspots) can only increase the number of equilibria. Consider an interest rate (Taylor) rule

$$1 + i_t = (1 + r)\Phi(\Pi_t); \quad \Pi_t \equiv P_t/P_{t-1}, \tag{16}$$

where $\Phi(\cdot)$ is a function allowing nonlinear policy rules. The consumer's first-order condition (14) reduces to

$$\Pi_{t+1} = \beta(1 + i_t). \tag{17}$$

We are looking for solutions to the pair (16) and (17). As before, we substitute out the interest rate and study the equation

$$\Pi_{t+1} = \Phi(\Pi_t). \tag{18}$$

We have a nonlinear, global, perfect-foresight version of the analysis in Section II.

As Benhabib et al. emphasize, a Taylor rule with slope greater than one should not apply globally to an economy in which consumers can hold money, because nominal interest rates cannot be negative. Thus, if we want to specify a Taylor rule with $\Phi_\pi > 1$ at some point, we should think about the situation as illustrated in figure 1. The equilibrium at Π^* satisfies the Taylor principle and is a unique locally bounded equilibrium. Any value of Π_0 other than Π^* leads away from the neighborhood of Π^* as shown. With a lower bound on nominal interest rates, however, the function $\Phi(\Pi)$ must also have another stationary point, labeled Π_L. This stationary point must violate the Taylor principle. Therefore, many paths lead to Π_L, and there are "multiple local equi-

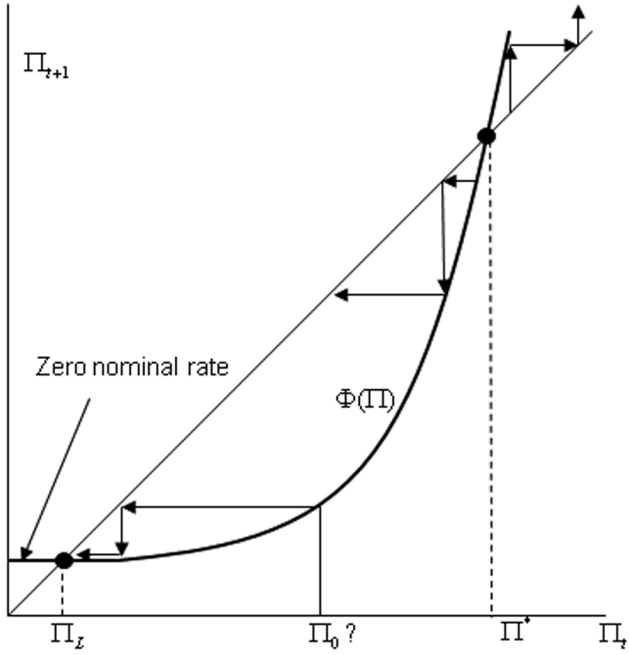

FIG. 1.—Dynamics in a perfect-foresight Taylor rule model

libria" near this point. In addition, the equilibria descending from Π* to Π_L are "bounded" though not "locally bounded."

(Yes, Π* is the "good" equilibrium and Π_L is the "bad" equilibrium. The point is to find determinacy by ruling out multiple equilibria. The value Π* is a unique locally bounded equilibrium. "Stability" near Π_L comes with "indeterminacy.")

All of these paths are equilibria. Since these paths satisfy the policy rule and the consumer's first-order conditions by construction, all that remains is to check that they satisfy the government debt valuation formula (15), that is, that there is a set of ex post lump-sum taxes that can validate them and hence ensure that the consumer's transversality condition is satisfied. There are lots of ways the government can implement such a policy. We need to exhibit only one. If the government simply sets net taxes in response to the price level as

$$S_t = \frac{r}{1+r}\frac{B_{t-1}}{P_t}, \tag{19}$$

then the real value of government debt will be constant, and the valuation formula will always hold.

To see why this is true, start with the flow constraint, proceeds of new

debt sales + taxes = old debt redemption:

$$\frac{B_t}{1 + i_t} + P_t S_t = B_{t-1}.$$

With $1 + i_t = (1 + r)P_{t+1}/P_t$, this can be rearranged to express the evolution of the real value of the debt:

$$\frac{B_t}{P_{t+1}} = (1 + r)\left(\frac{B_{t-1}}{P_t} - S_t\right). \tag{20}$$

Substituting the rule (19), we obtain

$$\frac{B_t}{P_{t+1}} = \frac{B_{t-1}}{P_t}.$$

We're done. With constant real debt and the flow relation (20), the transversality condition holds, and (20) implies (15). All the "explosive" equilibria as in Section II are, in fact, valid.

Deflationary equilibria that approach Π_L are also valid equilibria, as is Π_L itself. If we write the Taylor rule such that $i = 0$ at Π_L (e.g., $i_t = \max(0, r + \phi\pi_t)$), the "liquidity trap" or "Friedman rule" equilibrium $i_t = 0$, $\Pi_L = \beta$ (deflation at the rate r) is also a valid equilibrium.

D. Non-Ricardian Policy

The price level is uniquely determined in this frictionless model if we strengthen, rather than throw out, the government valuation equation—if the government follows a "non-Ricardian" fiscal regime. This is a natural alternative theory to consider, it is the basis for a lot of equilibrium trimming and related discussion that follows, and it clarifies the fundamental issue.

As the simplest example, suppose that fiscal policy sets the path of real net taxes $\{S_t\}$ independently of the price level. (A proportional income tax achieves this result.) The initial face value of one-period government debt B_{t-1} is predetermined at date t. Then, (15) determines the price level P_t,

$$\frac{B_{t-1}}{P_t} = E_t \sum_{j=0}^{\infty} \frac{1}{(1 + r)^j} S_{t+j}. \tag{21}$$

This is the same mechanism by which stock market prices are determined as the present value of dividends (Cochrane 2005).

The government can still follow an interest rate rule. By varying the amount of nominal debt sold at each date, the government can control expected future prices and hence the interest rate. Multiplying (21) at $t + 1$ by $1/(1 + r)$ and taking expectations, we get

$$\frac{B_t}{P_t}E_t\left(\frac{1}{1+r}\frac{P_t}{P_{t+1}}\right) = \frac{B_t}{P_t}\frac{1}{1+i_t} = E_t\sum_{j=1}^{\infty}\frac{1}{(1+r)^j}S_{t+j}.$$

The value of P_t is determined by (21). Then, by changing debt sold at time t, B_t, the government can determine i_t and $E_t(P_t/P_{t+1})$. Alternatively, the government can simply auction bonds at the interest rate i_t, and this equation tells us how many B_t it will sell.

Ex post inflation is determined by the ex post value of (21), which we can write in a pretty proportional form:

$$\frac{(E_{t+1}-E_t)[1/(1+\pi_{t+1})]}{E_t[1/(1+\pi_{t+1})]} = \frac{(E_{t+1}-E_t)\sum_{j=1}^{\infty}[1/(1+r)^j]S_{t+j}}{E_t\sum_{j=1}^{\infty}[1/(1+r)^j]S_{t+j}}. \quad (22)$$

Linearizing in the style of Section II, innovations to the present value of surpluses determine ex post inflation, the quantity $\delta_{t+1} = \pi_{t+1} - E_t\pi_{t+1}$, which indexed multiple equilibria in (5).

In this regime, the price level (not just inflation) is determinate, even with a constant interest rate target $i_t = i$. This regime also overturns the doctrine that interest rate targets lead to indeterminacy (Leeper 1991; Sims 1994; Woodford 1995).

Since it is free to set interest rates, the government can follow a Taylor rule. Thus, the empirical finding that a Taylor rule seems to fit well is not inconsistent with this theory, nor is the observation that central banks can and do set interest rates. A Taylor rule with a structural (not necessarily measured) $\phi > 1$ will generically lead to equilibria that are not locally bounded, unless fiscal shocks happen to select the new-Keynesian equilibrium. Thus, we obtain the usual doctrine following Leeper (1991) that "active" fiscal policy should be paired with "passive" ($\phi < 1$) monetary policy.

Similar identification problems remain, discussed in Section C.1 of online Appendix B. However, estimates of ϕ are not particularly important in this regime, as price-level determinacy or the control of inflation does not hinge on ϕ. In fact, problems in measuring ϕ are to some extent welcome. They mean we do not have to take regression estimates as strong evidence for a troublesome structural $\phi > 1$ despite stable inflation. Non-Ricardian models can also generate spurious $\phi > 1$ estimates.

At a minimum, the fiscal regime offers a way to understand U.S. history in periods that even new-Keynesians believe are characterized by passive ($\phi < 1$) monetary policy. This offers an improvement over "indeterminacy" or "sunspots," which place few restrictions on the data. Woodford (2001) applies this regime to the Fed's interest rate peg in the late 1940s and early 1950s. Applying it to the 1970s is an obvious possibility.

E. Ricardian Asymmetry, Asset Prices, and Observational Equivalence

Equations (21) and (22) also describe an equilibrium in which a variable, the price level, is a forward-looking expectation and jumps to avoid an explosive root. Recall the evolution of government debt (20) as

$$\frac{B_t}{P_{t+1}} = (1 + r)\left(\frac{B_{t-1}}{P_t} - S_t\right). \tag{23}$$

Again, we have an unstable root. If P_t is too low, then the real value of government debt explodes. In response to a shock, P_t jumps to the unique value that prevents such an explosion.

How do I accept explosive solutions in the new-Keynesian model while I deny them in the non-Ricardian regime? Why do I solve asset pricing equations $p_{t+1} = R_{t+1}p_t - d_{t+1}$ forward but not $\pi_{t+1} = \phi\pi_t - x_t$? There is a fundamental difference. There is a transversality condition forcing the consumer to avoid *real* explosions, explosions of B_{t-1}/P_t or the real value of assets. There is no corresponding condition forcing anyone to avoid nominal explosions, explosions of P_t itself.

Correspondingly, there is an economic mechanism forcing (21) to hold in a non-Ricardian regime. If the price level is below the value specified by (21), nominal government bonds appear as net wealth to consumers. They will try to increase consumption. Collectively, they cannot do so; therefore, this increase in "aggregate demand" will push prices back to the equilibrium level. Supply equals demand and consumer optimization are satisfied only at the unique equilibrium. Stock prices are pushed to the present value of dividends by the same mechanism. The valuation equation (15) is a market-clearing condition.

There is no corresponding mechanism to push inflation to the new-Keynesian value (6). In the new-Keynesian model we are choosing among equilibria; supply equals demand and consumer optimization hold for any of the alternative paths, any choice of δ_{t+1}; we are finding the unique locally bounded equilibrium, not the unique equilibrium itself.

In asset pricing equations such as (23) and $p_{t+1} = R_{t+1}p_t - d_{t+1}$, we can also measure the explosive eigenvalue, the rate of return, despite the forward-looking solution. This occurs because we can measure the dividend directly. In a deep sense, the reason we cannot measure ϕ is that we have no independent measure of the monetary policy shock.

Alas, passive and active fiscal regimes are observationally equivalent at this general level. All the equilibrium conditions hold in each case. We cannot test whether inflation occurred, and this caused the government to "passively" change taxes ex post, or whether people knew that taxes were going to change, and the price level jumped in their expec-

tation. Canzoneri, Cumby, and Diba's (2001) fiscal test has the same flaw as Clarida et al.'s (2000) monetary test, shown in Cochrane (1998).

The regimes are also not as conceptually distinct as they may appear. For example, if the government runs a commodity standard, offering to buy and sell a commodity at a given price, it must adjust taxes so as always to have sufficient stocks of the commodity on hand. Is this "Ricardian" or "non-Ricardian"? One could say that the government valuation equation (21) "really" determines the price level, and the commodity standard just communicates the necessary fiscal commitment. Since the present value of future surpluses is on its own difficult to forecast, communicating such a fiscal commitment is a useful way to stabilize prices and a central part of any successful monetary-fiscal policy structure. And commodity standards and pegs fall apart precisely when the underlying fiscal commitment is no longer credible.

Similarly, if the new-Keynesian equilibrium selection were successful, one could say that the government valuation equation (21) "really" determines the price level, with interest rate policy merely a way to communicate and enforce that fiscal commitment. In this view, the problem with the new-Keynesian interest rate regime is that it does not communicate a unique fiscal commitment.

IV. New-Keynesian Solutions

Of course, the new-Keynesian literature is aware of these issues. How do new-Keynesian authors pick the locally bounded solution Π^* and get rid of the other ones?

A. *Reasonable Expectations?*

Much of the approach is simply to think about what expectations seem reasonable. For example, Woodford (2003, 128) argues that "the equilibrium [Π^*] . . . is nonetheless *locally* unique, which may be enough to allow expectations to coordinate upon that equilibrium rather than on one of the others." Similarly, King (2000, 58–59) writes that "by specifying [$\phi > 1$] then, the monetary authority would be saying, 'if inflation deviates from the neutral level, then the nominal interest rate will be increased relative to the level which it would be at under a neutral monetary policy.' If this statement is believed, then it may be enough to convince the private sector that the inflation and output will actually take on its neutral level."

This seems a rather weak foundation for the basic economic question, how the price level is determined. It is especially weak in ruling out equilibria between Π_L and Π^*. One might think that expectations of accelerating inflation are not reasonable. But if Π^*, say 2 percent, in-

flation expectations are reasonable, is a path that starts at 1 percent inflation and slowly declines to Π_L near zero really so unreasonable?

Importantly for judging the reasonableness of alternative equilibria, Woodford argues that we should not think of an economy or Fed making a small "mistake" and therefore slipping from Π^* into an explosive equilibrium. Instead we should think of expectations of future inflation driving inflation today:

> Indeed it is often said that . . . the steady state with inflation rate Π^* is "unstable" implying that an economy should be expected almost inevitably to experience either a self-fulfilling inflation or a self-fulfilling deflation under such a regime.
>
> Such reasoning involves a serious misunderstanding of the causal logic of the difference equation [(18)]. This equation does not indicate how the equilibrium inflation rate in period $t + 1$ is determined by the inflation that happens to have occurred in the previous period. If it did it would be correct to call Π^* an unstable fixed point of the dynamics—even if that point were fortuitously reached, any small perturbation would result in divergence from it. But instead, the equation indicates how the equilibrium inflation rate in period t is determined by *expectations* regarding inflation in the following period. . . . The equilibria that involve initial inflation rates near (but not equal to) Π^* can only occur as a result of expectations of *future* inflation rates (at least in some states) that are even *further* from the target inflation rate. Thus the economy can only move to one of these alternative paths if expectations about the future change significantly, something that one may suppose would not easily occur. (Woodford 2003, 128)

A "serious misunderstanding of causal logic" is a strong charge, and I think unwarranted here. The equations of the model do not specify a causal ordering. They are just equilibrium conditions. And a strict opposite causal ordering does not make sense either. If you see a small change today in an unstable dynamic system, your expectations of the future may well change by a large amount. If you see the waiter trip, it is a good bet that the stack of plates he is carrying will crash. In new-Keynesian models, agents might well see a disturbance, know the Fed will feed back on its past mistakes, think "oh no, here we go," and radically change their expectations of the future. They do not need to wake up and think "gee, I think there will be a hyperinflation" before reading the morning paper. The new-Keynesian forward-looking solutions rely exactly on such endogenous expectations: near-term expectations jump in response to a shock, to put the economy back on the saddle path that has no change in asymptotic expectations.

Still, there is some appeal to the argument that expectations of hy-

perinflations seem far-fetched. But expectations that are far-fetched in our intuitive understanding of our world are not necessarily so far-fetched for agents in this model, once we recognize that this model may not represent our world. In this model, the Fed is absolutely committed to raising interest rates more than one-for-one with inflation, forever, no matter what. In this model, real rates are constant, so the rise in nominal rates must correspond directly to a rise in inflation—precisely the opposite of the explicitly stabilizing language in the Federal Reserve's account of its actions. If we lived in such a world, I would confidently expect hyperinflation. If we think that forecast is "unreasonable," it means we do not believe the model describes the world in which we live.

B. Solutions and Dilemma; Stabilizations

Recognizing, I think, the weakness of these arguments—if not, they would not need to go on—new-Keynesian theorists have explored a variety of ways to trim multiple equilibria. Alas, these fall in the logical conundrum explained in the introduction: To trim equilibria, we must assume that the government will blow up the world—to set policy in such a way that private first-order conditions cannot hold—even though such policies cannot be achieved through markets, and even though policies exist that would allow the government to stop inflation or deflation while letting the economy operate.

Woodford (2003, sec. 4.3) studies proposals to cut off inflationary equilibria to the right of Π^*. His main suggestion is that

> Self-fulfilling inflations may be excluded through the addition of policy provisions that apply only in the case of hyperinflation. For example, Obstfeld and Rogoff (1986) propose that the central bank commit itself to peg the value of the monetary unit in terms of some real commodity by standing ready to exchange the commodity for money in event that the real value of the total money supply ever shrinks to a certain very low level. If it is assumed that this level of real balances is one that would never be reached except in the case of a self-fulfilling inflation, the commitment has no effect except to exclude such paths as possible equilibria. . . . [This proposal could] well be added as a hyperinflation provision in a regime that otherwise follows a Taylor rule. (138)

In real life, governments often stop inflations by a firm peg to a foreign currency (with a fiscal reform, to make credible the corresponding fiscal policy commitment), which is the modern equivalent of a commodity standard. Atkeson, Chari, and Kehoe (2010) advocate a similar idea but specify that the government switches to a money growth rule in a model

with non-interest-elastic money demand. Switching to a non-Ricardian regime to enforce a fixed price level would have the same effect.

However, this quote and the surrounding discussion do not explain how *stabilizing an inflation* serves to *rule out an equilibrium path.* First-order and market-clearing conditions can hold throughout the inflation and its stabilization, and then the path is not ruled out.

The answer is that each of these proposals implicitly pairs the stabilization with another policy specification, not needed to stop the inflation, in such a way that equilibrium cannot form. *Inconsistent policy* rules out the equilibrium path, not inflation stabilization.

The key assumption in Woodford's quote is "otherwise follows a Taylor rule." His government continues to follow the Taylor rule even after it has switched to a commodity standard! You cannot have two monetary policies at once; if you do, no equilibrium can form.

To be precise, suppose that at inflation past some level Π_U the government changes to a commodity standard (a peg), switches to a money growth rule with interest-inelastic demand, or switches to a non-Ricardian regime. At date $T-1$, $\Pi_{T-1} < \Pi_U$, so the consumer obeys his first-order conditions, the Fed follows the Taylor rule, and equilibrium inflation still follows

$$\Pi_T = \beta(1 + i_{T-1}) = \Phi(\Pi_{T-1}).$$

(In the linearized model, $\Pi_t = \phi\Pi_{t-1}$.)

Now, suppose $\Pi_T > \Pi_U$ so at date T, the government freezes this price level at P_T by one of the above policies, and $P_{T+1} = P_T$. Equilibrium at date T therefore requires $i_T = r$ from the consumer's first-order conditions

$$\Pi_{T+1} = 1 = \beta(1 + i_T).$$

(In the linearized model, $i_T = r + \Pi_{T+1}$.)

The hyperinflation has ended, but this fact does not "exclude such paths as possible equilibria." The key to "excluding equilibria" is that Woodford, Atkeson et al., and others assume that the Fed *also* continues to follow the Taylor rule,

$$1 + i_T = (1 + r)\Phi(\Pi_T).$$

(In the linearized model, $i_T = r + \phi\Pi_T$.) This is a huge number and is inconsistent with $i_T = r$ demanded by first-order conditions.

We would normally say that it is impossible both to run a commodity standard that requires $i_T = r$ *and* to set interest rates at hyperinflationary levels that require i_T to be a huge number. As Woodford explains, the government implements a commodity standard by "standing ready to exchange the commodity for money." It cannot both do that and control the quantity of money to follow an interest rate target. If the government really could commit to such a thing, there would be "no equilibrium."

But does it really make any sense that the government would try to do such a thing, that people would believe that it would try to do such a thing, that a government even can do such a thing and persist long enough to "rule out equilibrium," whatever that means? It does not have to: It can abandon the Taylor rule at T, when it starts using the commodity standard.

At a minimum, we see that *stabilizing inflation* has nothing to do with *ruling out the equilibrium path*. One period of inconsistent policy anywhere along the path is enough to accomplish the latter.

Atkeson et al. (2010) recognize the problem and carefully set up policy so that equilibrium can form on every date past T. However, they also assume that the Taylor rule requiring high interest rates coexists for one period with a money growth rule that demands low interest rates, in order to rule out equilibrium. Blowing up the world for one period is enough.

What about Obstfeld and Rogoff (1983, 1986) and the related large literature that tries to trim indeterminacies in models with fixed money supply and interest-elastic money demand? Didn't they solve all these problems years ago, as Woodford seems to suggest? Since it requires setting up a different model, I review these proposals in Appendix A. The answer is the same. Switching to a commodity standard at a very high level of inflation stops the inflation, but it allows an equilibrium at each date, so the inflationary equilibrium path is not ruled out. To rule out that equilibrium path, one must also control the money supply, for example, disallowing the recovery in real money balances that accompanies the end of hyperinflations. Again, what government would do this? How could a government do this? How could a government freely trade currency for the commodity at a given price and impose an upper limit on the money supply? I conclude that models with interest-elastic money demand $MV(i) = PY$, fixed M, and passive fiscal policies have exactly the same unsolved indeterminacies as the Taylor rule models. (In fact, Obstfeld and Rogoff's [1983] actual proposal does not even invalidate the inflation as an equilibrium path. Appendix A analyzes their case in detail.)

A variant on this policy can work, however. Suppose that if inflation exceeds some value Π_U, the government commits to instantly returning to the initial price *level*, P_0, by a commodity standard. Negative nominal rates are not a market-clearing condition, so this commitment rules out a high level of P_T as an equilibrium and, hence, the path leading up to it.

This is not a blow-up-the-world threat, as the government abandons the Taylor rule in period T. It is close to fiscal. A commodity standard must be paired with an appropriate fiscal regime. The "Ricardian" assumption will be tested by the offer to redeem the money stock at a

much higher real value. Whether one regards this as "Ricardian" or "non-Ricardian" is largely semantic. The inflation never gets going because money holders understand that money is eventually backed by real goods and by the government's ability to tax in order to provide those real goods. The future commitment leads to greater demand for money at time 0.

However, though it may describe other governments and especially the United Kingdom as it returned to the gold standard at parity after suspensions of convertibility during wars and crises, it is not a vaguely plausible description of expectations regarding current government.

C. Fiscal Equilibrium Trimming

Benhabib et al. (2002), mirrored in Woodford (2003, sec. 4.2), try to trim equilibria by adding fiscal commitments to the Taylor rule. Their ideas are aimed at trimming deflationary equilibria, but the same ideas could apply to both inflations and deflations. They specify that in low-inflation states, the government will lower taxes so much that real debt grows explosively, the consumer's transversality condition is violated, and the government debt valuation equation no longer holds. Therefore, the low-inflation region and all the equilibria below Π^* in figure 1 that lead to it are ruled out. Specifically (their eqq. [18]–[20]), they specify that in a neighborhood of Π_L, the government will commit to surpluses $S_t = \alpha(\Pi_t)(B_{t-1}/P_t)$ with $\alpha(\Pi_L) < 0$ in place of (19).

They also show that the same result can be implemented by a target for the growth rate of nominal liabilities, a "4 percent rule" for nominal debt. If deflation breaks out with such a commitment, real debt will then explode; to keep nominal debt on target, the government would need to start borrowing and spending as above. Woodford suggests this implementation as well: "let total nominal government liabilities D_t be specified to grow at a constant rate $\bar{\mu} > 1$ while monetary policy is described by the Taylor rule. . . . Thus, in the case of an appropriate fiscal policy rule, a deflationary trap is not a possible rational expectations equilibrium" (132).

As the above proposals are grounded in sensible policies to stabilize hyperinflations, these proposals sound like sensible prescriptions to inflate the economy, that is, to head back to the desired equilibrium Π^*. Benhabib et al. describe them this way: "Interestingly, this type of policy prescription is what the U.S. Treasury and a large number of academic and professional economists are advocating as a way for Japan to lift itself out of its current deflationary trap. . . . A decline in taxes increases the household's after-tax wealth, which induces an aggregate excess demand for goods. With aggregate supply fixed, price level must increase in order to reestablish equilibrium in the goods market" (2002,

548). (They did not know that zero interest rates and $1.5 trillion deficits would so soon follow!) And this is, indeed, how a coordinated fiscal-dominant regime works; it is good intuition for operation of the fiscal theory of the price level and may capture what real-world proponents of these policies have in mind.

But that is not their proposal. The proposal does not "lift the economy out of a deflationary trap" back to Π^*. Their proposal sits at Π_L with an uncoordinated policy and lets government debt explode. If their proposal did successfully steer the economy back to Π^*, then the whole path to Π_L and back would have been an equilibrium. Benhabib et al. change tax policy *while also maintaining the Taylor rule* $\Phi(\Pi)$ and the dynamics of figure 1. In Woodford's page 132 quote, "*while monetary policy is described by the Taylor rule*" is the key. We are switching to a Ricardian regime, which demands higher inflation, while simultaneously keeping the Taylor rule in place, which demands continued low inflation. The transversality condition is a consumer first-order condition. We are setting policy parameters for which consumer first-order conditions cannot hold.

Once we see that central point, we can think of many monetary-fiscal policies that preclude deflationary equilibria equivalently and more transparently. If inflation gets to an undesired level, tax everything. Burn the money stock. Introduce an arbitrage opportunity. Best of all, specify a $\Phi(\Pi)$ function that includes negative nominal interest rates—just eliminate the Π_L equilibrium in the first place! Bassetto (2004) suggests this option. Since there can be no equilibrium at negative nominal rates, such a $\Phi(\Pi)$ function works exactly the same way to rule out equilibria: In a deflationary state, the government commits to a policy that allows no equilibrium. Negative nominal rates are no more or less possible than letting debt explode or running a commodity standard with high rates or low money stock. In retrospect, it does not make sense to demand a Ramsey approach in setting up the problem—the Taylor rule must not prescribe negative nominal rates because that would violate first-order conditions—and then patch it up with policy prescriptions that do violate first-order conditions. Why not just commit to negative nominal rates in the first place?

It is not hard to understand why the issue has become so confused. Benhabib et al., Woodford, and other authors did not follow my alternative suggestions: to specify policy paths that clearly, decisively, and unrealistically, forbid equilibrium. Instead, they thought about a very reasonable-sounding response to inflation or deflation and then subtly (and doubtless unintentionally) snuck in an extra step that rules out equilibrium. It is very easy to confuse "stopping an inflation" with "ruling out this equilibrium path." The easy-to-miss little extra step matters, not the seductively sensible policy that surrounds it.

There is an important difference between Benhabib et al.'s proposal and those previously mentioned, which leads to a more sympathetic reading. Their government does not *switch* to a non-Ricardian regime at low inflation when $\alpha(\Pi_t)$ turns negative. It was there all along. Since any inflation rate below Π^* leads inexorably to a state in which real government debt explodes, the valuation equation for government debt (15) does not hold for any $\Pi_0 < \Pi^*$. The fact that $\alpha(\Pi_t) > 0$ temporarily in $S_t = \alpha(\Pi_t)(B_{t-1}/P_t)$ does not, together with the Taylor rule, produce a Ricardian regime while inflation is still high. This fact gives a supply-and-demand force for raising inflation immediately, as in any non-Ricardian regime. If a consumer contemplates $\Pi_0 = \Pi^* - \varepsilon$, he sees that government bonds are worth less and tries to get rid of them, raising aggregate demand, and bringing inflation back up to Π^* immediately. The operation is the same as if the government had simply announced a non-Ricardian regime to support Π^*. The Taylor rule just makes the demand curve underlying this regime vertical.

Read this way, Benhabib et al.'s proposal is feasible, as the commitments underlying non-Ricardian fiscal regimes are feasible. History since the publication of their paper seems to have borne out their predictions for government behavior. But their proposal was supposed to *rule out* this equilibrium path, not to describe history. Their point is, with these expectations, inflation should never have fallen in the first place. So we cannot appeal to recent history in support of their analysis. Either the model is wrong—perhaps we are at Π^*—or perhaps people do not believe that the government really will let government debt explode as a response to lower-than-desired inflation. And the inflationary paths remain.

D. Weird Taylor Rules

Woodford starts "policies to prevent an inflationary panic" by suggesting (136) a stronger Taylor rule. He suggests that the graph in figure 1 becomes vertical at some finite inflation Π_U above Π^*, that is, that the Fed will set an infinite interest rate target. Similarly, Alstadheim and Henderson (2006) remove the Π_L equilibrium by introducing discontinuous policy rules, or V-shaped rules that touch the 45-degree line only at the Π^* point. Bassetto (2004), mentioned above, suggested that the Taylor rule ignore the $i \geq 0$ bound and promise negative nominal rates in a deflation.

These proposals blow up the economy directly. At one level, however, these proposals are not as extreme as they sound. After all, the Taylor principle in new-Keynesian models amounts to people believing unpleasant things about alternative equilibria. Hyperinflating away the entire monetary system ($\Phi(\Pi)$ becoming vertical), introducing an arbitrage

opportunity (allowing $i < 0$ in the policy rule), and so forth are perhaps more effective than an inflation or deflation that slowly gains steam.

However, it is not clear that all these proposals rule out equilibria. A currency can be completely inflated away in finite time. Obstfeld and Rogoff (1983) has this property, and Zimbabwe experienced it.

While they are possible commitments one might ask a future Fed to make, none of these proposals are even vaguely plausible descriptions of current beliefs about Fed behavior or current Fed statements.

E. Residual Money Demand: Letting the Economy Blow Up

Schmitt-Grohé and Uribe (2000) and Benhabib, Schmitt-Grohé, and Uribe (2001) offer a similar way to rule out hyperinflations, without assuming that the Fed directly blows up the economy with infinite interest rates, by adding a little money. This idea is also reviewed by Woodford (2003, 137) and has long roots in the literature on hyperinflations with fixed money supply and interest-elastic demand (Sims 1994).

Schmitt-Grohé and Uribe's idea is easiest to express with real balances in the utility function. With money, the Fisher equation contains monetary distortions:

$$1 + i_t = \Pi_{t+1} \frac{u_c(Y,\ M_t/P_t)}{\beta u_c(Y,\ M_{t+1}/P_{t+1})} = \Pi_{t+1}(1 + r_t), \tag{24}$$

where r_t denotes the real interest rate. (This is a perfect-foresight model, so the expectation is missing.) Suppose that the Taylor rule is

$$1 + i_t = \frac{1}{\beta}\Phi(\Pi_t).$$

When we substitute i_t from the Taylor rule into (24) and rearrange the money versus bonds first-order condition as $M_t/P_t = L(Y,\ i_t)$, inflation dynamics follow

$$\Pi_{t+1} = \Phi(\Pi_t) \frac{u_c[Y,\ L(Y,\ \Phi(\Pi_{t+1}))]}{u_c[Y,\ L(Y,\ \Phi(\Pi_t))]} \tag{25}$$

instead of (18).

The idea, then, is that this difference equation may rise to require $\Pi_{t+1} = \infty$ above some bound Π_U, even if the Taylor rule for nominal interest rates $1 + i_t = \Phi(\Pi_t)/\beta$ remains bounded for all Π_t. Woodford and Schmitt-Grohé and Uribe give examples of specifications of $u(C,\ M/P)$ for which this situation can happen.

Is this the answer? First and most important, if we do not regard a belief that the Fed will directly blow up the economy (i_t rises to infinity) as a reasonable characterization of expectations, why would people believe that the Fed will to take the economy to a state in which the economy blows up all on its own? Infinite inflation and finite interest

rates mean infinitely negative real rates, a huge monetary distortion. Surely the Fed would notice that real interest rates are approaching negative infinity!

Second, it is delicate. In general, this approach relies on particular behavior of the utility function or the cash-credit goods specification at very low real balances. Are monetary frictions really important enough to rule out inflation above a certain limit, sending real rates to negative infinity, or to rule out deflation below another limit? We have seen some astounding hyperinflations; real rates did not seem all that affected.

Sims (1994) pursues a similar idea. Perhaps there is a lower limit on nominal money demand, so that real money demand explodes in a deflation. Perhaps not; perhaps the government can print any number it wants on bills or will run periodic currency reforms; perhaps real money demand is finite for any price level.

In sum, these proposals require two things: First, they require expectations that the government will follow the Taylor rule to explosive hyperinflations and deflations, beyond anything ever observed, and despite the presence of equilibrium-preserving stabilization policies such as the switch to a commodity standard, money growth, or non-Ricardian regime. Second, they require belief in a deep-seated monetary nonneutrality sufficient to send real rates to negative infinity or real money demand to infinity, though even the beginning of such events has never been observed. At a minimum, expectations of such events sound again like a weak foundation for what should be a simple question, the basic determination of the price level.

V. Determinacy and Identification in the Three-Equation Model

One may well object to the whole idea of studying identification and determinacy in such a stripped-down model, with no monetary friction, no means by which the central bank can affect real rates, and a single disturbance. Typical verbal (old-Keynesian) explanations of Taylor rules and inflation, and typical Federal Open Market Committee (FOMC) statements, involve at least the Phillips curve and Fed control of real rates of interest: nominal rates rise, gaps appear, and these gaps drive down inflation. You cannot do that in a frictionless model. Empirical Taylor rule estimates are much more sophisticated than $i_t = \phi\pi_t + \varepsilon_t$ regressions.

It turns out that the simple model does in fact capture the relevant issues, but one can show that only by examining "real" new-Keynesian models and regressions in detail and seeing that the same points and same logic emerge.

The excellent exposition in King (2000) makes the nonidentification and determinacy theorems clear. The basic model is

$$y_t = E_t y_{t+1} - \sigma r_t + x_{dt}, \tag{26}$$

$$i_t = r_t + E_t \pi_{t+1}, \tag{27}$$

$$\pi_t = \beta E_t \pi_{t+1} + \gamma(y_t - \bar{y}_t) + x_{\pi t}, \tag{28}$$

where y denotes output, r denotes the real interest rate, i denotes the nominal interest rate, π denotes inflation, $\bar{y}_t$ is potential output, and the x are serially correlated structural disturbances. I use x, not ε, and the word "disturbance" rather than "shock" to remind us of that fact.

While seemingly ad hoc, this structure has exquisite micro foundations, which are summarized in King (2000) and Woodford (2003). The first two equations derive from consumer first-order conditions for consumption today versus consumption tomorrow. The last equation is the "new-Keynesian Phillips curve," derived from the first-order conditions of forward-looking optimizing firms that set prices subject to adjustment costs. There is an active debate on the right specification of (28), including additional inflation dynamics and the difference between output and marginal cost, but these differences do not affect my conclusions.

For both determinacy and identification questions, we can simplify the analysis by studying alternative equilibria as *deviations* from a given equilibrium, following King (2000). Use y_t^* and so forth to denote equilibrium values. The term y_t^* is a stochastic process, that is, a moving average representation $y_t^*(\{x_{dt}, x_{\pi t}, \ldots\})$ or its equivalent. There are many such equilibria. For example, given any stochastic process for $\{y_t^*\}$, you can construct the corresponding $\{\pi_t^*\}$, $\{r_t^*\}$, $\{i_t^*\}$ from (28), (26), and (27) in order.

Use tildes to denote deviations of an alternative equilibrium y_t from the * equilibrium, $\tilde{y}_t \equiv y_t - y_t^*$. After subtraction, deviations must follow the same model as (26)–(28), but without constants or disturbances:

$$\tilde{i}_t = \tilde{r}_t + E_t \tilde{\pi}_{t+1}, \tag{29}$$

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \sigma \tilde{r}_t, \tag{30}$$

$$\tilde{\pi}_t = \beta E_t \tilde{\pi}_{t+1} + \gamma \tilde{y}_t. \tag{31}$$

A. *Determinacy*

Now, if the Fed sets $i_t = i_t^*$, that is, $\tilde{i}_t = 0$, then $\tilde{\pi}_t = 0$, $\tilde{y}_t = 0$ are an equilibrium. But this is not the only equilibrium. To see this point, write (29)–(31) with $\tilde{i}_t = 0$ in a standard form as

$$\begin{bmatrix} E_t\tilde{y}_{t+1} \\ E_t\tilde{\pi}_{t+1} \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} \beta + \sigma\gamma & -\sigma \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix}. \tag{32}$$

Since the model restricts only the dynamics of expected future output and inflation, we have multiple equilibria. Any

$$\begin{bmatrix} \tilde{y}_{t+1} \\ \tilde{\pi}_{t+1} \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} \beta + \sigma\gamma & -\sigma \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix} + \begin{bmatrix} \delta_{y,t+1} \\ \delta_{\pi,t+1} \end{bmatrix} \tag{33}$$

with $E_t\delta_{y,t+1} = 0$, $E_t\delta_{\pi,t+1} = 0$ is valid, not just $\delta_{y,t+1} = \delta_{\pi,t+1} = 0$, and hence $\tilde{y}_t = \tilde{\pi}_t = 0$ for all t.

Perhaps, however, the dynamics of (32) are explosive, so $\tilde{y} = \tilde{\pi} = 0$ is the only locally bounded equilibrium, the only one in which $E_t(\tilde{y}_{t+j})$ and $E_t(\tilde{\pi}_{t+j})$ stay near zero. Alas, this hope is dashed as well: One of the eigenvalues of the transition matrix in (32), derived below, is less than one. We have just verified in this model the usual doctrine that an interest rate peg does not determine inflation.

To determine output and the inflation rate, then, new-Keynesian modelers add to the specification $i_t = i_t^*$ of what interest rates will be in *this* equilibrium, a specification of what interest rates would be like in *other* equilibria, in order to rule them out. King (2000) specifies Taylor-type rules in the form

$$i_t = i_t^* + \phi_\pi(\pi_t - \pi_t^*) + \phi_y(y_t - y_t^*) \tag{34}$$

or, more simply,

$$\tilde{i}_t = \phi_\pi\tilde{\pi}_t + \phi_y\tilde{y}_t.$$

(Both King and online App. B allow responses to expected future inflation and output as well. This generalization does not change my points.)

For example, with $\phi_y = 0$ the deviations from the * equilibrium now follow:

$$\begin{bmatrix} E_t\tilde{y}_{t+1} \\ E_t\tilde{\pi}_{t+1} \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} \beta + \sigma\gamma & -\sigma(1 - \beta\phi_\pi) \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix}. \tag{35}$$

The eigenvalues of this transition matrix are

$$\lambda = \frac{1}{2\beta}[(1 + \beta + \sigma\gamma) \pm \sqrt{(1 + \beta + \sigma\gamma)^2 - 4\beta(1 + \sigma\gamma\phi_\pi)}]. \tag{36}$$

If we impose $\sigma\gamma > 0$, then both eigenvalues are greater than one in absolute value if $\phi_\pi > 1$ or if

$$\phi_\pi < -\left(1 + 2\frac{1 + \beta}{\sigma\gamma}\right). \tag{37}$$

Thus, if the policy rule is sufficiently "active," any equilibrium other than $\tilde{i} = \tilde{y} = \tilde{\pi} = 0$ is explosive. Ruling out such explosions, we now have the unique locally bounded equilibrium. (Online App. B treats

determinacy conditions with output responses and responses to expected future inflation and output.)

As in the simple model, the point of policy is to induce *explosive* dynamics, eigenvalues greater than one, not to "stabilize" so that the economy always reverts after shocks. As pointed out by King (2000, 78), the region of negative ϕ_π described by (37), which generates oscillating explosions, works as well as the conventional $\phi_\pi > 1$ to determine inflation.

The analysis so far has exactly mirrored my analysis of the simple model of Section II. So, in fact, that model does capture the determinacy issues, despite its absence of any frictions. Conversely, determinacy in the new-Keynesian model does not fundamentally rely on frictions, the Fed's ability to control real rates, or the Phillips curve. As in the simple model, "determinacy" is a question of multiple equilibria, not inflationary or deflationary "spirals."

As in the simple model, no economic consideration rules out the explosive solutions. One might complain that I have not shown the full, nonlinear model in this case, as I did for the frictionless model. This is a valid complaint, especially since output may also explode in the linearized nonlocal equilibria. I do not pursue this question here since I find no claim in any new-Keynesian writing that this route can rule out the nonlocal equilibria. Its determinacy literature is all carried out in simpler frameworks, as I have done. And there is no reason, really, to suspect that this route will work either. Sensible economic models work in hyperinflation or deflation. If they do not, it usually reveals something wrong with the model rather than the impossibility of inflation. In particular, while linearized Phillips curve models can give large output effects of high inflations, we know that some of their simple abstractions, such as fixed intervals between price changes, are useful approximations only for low inflation. The Calvo fairy seems to visit more often in Argentina.

In one respect, this analysis is quite different from the simple model of Section II. Determinacy is a property of the entire system and depends on other parameters of the model, not just ϕ_π. Here, for $\sigma\gamma < 0$, there is a region with $\phi_\pi > 1$ in which both eigenvalues are not greater than one, so we have indeterminacy despite an "active" Taylor rule. There is another region in which both eigenvalues are greater than one despite $0 < \phi_\pi < 1$, so we have local determinacy despite a "passive" Taylor rule. The parameter configuration $\sigma\gamma < 0$ is not plausible, but as models become more complex, determinacy involves more parameters and can often involve plausible values of those parameters. The regions of determinacy are not as simple as $\phi_\pi > 1$, and testing for determinacy is not as simple as testing the parameters of the Fed reaction function. Alas, no one has tried a test for determinacy in a more complex model.

King's expression of the Taylor rule (34) is particularly useful because it clearly separates "in-equilibrium" or "natural rate" i_t^* and "alternative-equilibrium" or determinacy $\phi_\pi(\pi_t - \pi_t^*)$ issues so neatly. One can read its instructions as follows: First, the Fed should set the interest rate to the natural rate i_t^* that appropriately reflects other shocks in the economy. Then the Fed should react to inflation away from the desired equilibrium in order to induce local determinacy of the i^* equilibrium. The two issues are completely separate.

For example, many theoretical treatments find that interest rates that move more than one-for-one with inflation are desirable, for reasons other than determinacy, or one may accept empirical evidence that they do so. But both of these are observations that *equilibrium* interest rates i_t^* should, or do, vary more than one-for-one with *equilibrium* inflation, π_t^*. King's expression (34) emphasizes that these observations tell us nothing, really, about determinacy issues, whether *deviations* from equilibrium should or do follow the same patterns.

In particular, one might object that a nonexplosive, non-Ricardian regime requires $\phi < 1$ and that Taylor rule regressions give coefficients greater than one. But King's expression (34) shows us that a more than one-for-one relation between i_t^* and π_t^* is perfectly consistent with a less than one-for-one relationship $\phi < 1$ between deviations $(i - i^*)$ and $(\pi - \pi^*)$.

B. *Identification*

King's expression of the Taylor rule (34) makes the central identification point clear. In the * equilibrium, we will always see $\pi_t - \pi_t^* = 0$ and $y_t - y_t^* = 0$. Thus, a regression estimate of (34) cannot possibly estimate ϕ_π, ϕ_y. There is no movement in the necessary right-hand variables. More generally, ϕ_π and ϕ_y appear nowhere in the equilibrium dynamics characterized simply by $\tilde{\pi} = \tilde{y} = \tilde{i} = 0$, so they are not identified. Taylor determinacy depends entirely on what the Fed would do *away* from the * equilibrium, which we can never see from data *in* that equilibrium.

King recognizes the issue: "The specification of this rule leads to a subtle shift in the interpretation of the policy parameters $[\phi_\pi, \phi_y]$; these involve specifying how the monetary authority will respond to deviations of inflation from target. But if these parameters are chosen so that there is a unique equilibrium, then no deviations of inflation will ever occur" (41). He does not address the implications of this issue for empirical work.

This issue is not particular to the details of the three-equation model. In the general solution method for these sorts of models, we set to zero movements of the linear combinations of variables that correspond to

unstable eigenvalues. As a result, we cannot measure those unstable eigenvalues. Online Appendix B makes this point with equations.

So, what assumptions do people make to escape this deep problem? The prototype theoretical Taylor rule (34), repeated here,

$$i_t = i_t^* + \phi_\pi(\pi_t - \pi_t^*) + \phi_y(y_t - y_t^*), \tag{38}$$

describes how the central bank would react to potential deviations from the equilibrium π^*, y^* in order to make y_t^* and π_t^* the unique locally bounded equilibrium. To identify ϕ_π, ϕ_y, then, we have to make two assumptions.

ASSUMPTION 1. The Fed's reaction ϕ_π, ϕ_y to a *deviation* of inflation π_t and output y_t from the desired equilibrium value π_t^* and y_t^* is the same as the relation between equilibrium interest rates i_t^* and equilibrium inflation π_t^* and y_t^*; we must assume that the ϕ in (38) are the same as the ϕ^* in a relation such as

$$i_t^* = \phi_\pi^* \pi_t^* + \phi_y^* y_t^* + \cdots + x_{it}, \tag{39}$$

where x_{it} denotes a residual combination of shocks with sufficient orthogonality properties to allow some estimation. (Of course, leads and lags and other variables may appear in both [38] and [39].)

Making this assumption is (for once) relatively uncontroversial since there are no obvious observations one could make to refute it. Still, it is worth making the assumption explicit and at least worth reading Fed statements to see if they support it.

The key question is whether we are able to make assumption 1. Doing so requires restrictions on the model and equilibrium. The equilibrium quantities i_t^*, y_t^*, and π_t^* are functions of shocks, $i_t^*(\{x_{dt}, x_{\pi t}, x_{it}, \ldots\})$, "moving averages." To be able to make assumption 1, we need a second assumption.

ASSUMPTION 2. The model and Fed's choice of equilibrium (different i^* imply different y^*, π^* for a given model) must be such that equilibrium quantities can be expressed in the "autoregressive" representation (39), with parameters ϕ^* in the zone of determinacy (explosive eigenvalues), and with the error x_{it} orthogonal to something we can use as instruments.

Many models and many equilibria of a given model do not have this property. As an example, consider the identification failure of Section II. The equilibrium there is, in response to shock form,

$$\pi_t^* = kx_t,$$

$$i_t^* = r + \rho k x_t,$$

where k is a constant. We can express the interest rate equation as a relationship among endogenous variables,

$$i_t^* = r + \rho \pi_t^*.$$

However, $\|\rho\| < 1$, so we cannot use this relationship among endogenous variables as a Taylor rule for interest rate policy. This example violates the qualification "with ϕ^* in the zone of determinacy." If we try to express this equilibrium as a rule with larger ϕ_π^*,

$$i_t^* = r + \phi_\pi^* \pi_t^* + (\rho - \phi_\pi^*) k x_t,$$

we obtain an "error term"—the x_{it} in (39)—that is hopelessly correlated with the right-hand variables π_t^* and all instruments, exactly the point of Section II. This relationship violates the qualification on the error term in assumption 2.

Even "ϕ^* in the zone of determinacy" is really too loose. For example, suppose that the correlations between variables in an equilibrium require $\phi_\pi^* = 10$. This equilibrium can be supported by $\phi_\pi = 10$, but it also can be supported by a more sensible $\phi_\pi = 1.5$. We can assume that they are the same, identifying $\phi_\pi = 10$, but maybe we would not want to make the basic assumption in this case.

The no-gap equilibrium is a particularly good example in the three-equation context, since minimizing output gaps is a natural objective for monetary policy. To see this result most simply, suppose that all the shocks follow AR(1) processes $x_{jt} = \rho_j x_{jt-1} + \varepsilon_{jt}$. Then, when we substitute $y_t^* = \bar{y}_t$ in (26)–(28), the no-gap equilibrium is, in moving average form,

$$\pi_t^* = \frac{1}{1 - \beta\rho_\pi} x_{\pi t}, \tag{40}$$

$$i_t^* = r - \frac{1 - \rho_{\bar{y}}}{\sigma} \bar{y}_t + \frac{\rho_\pi}{1 - \beta\rho_\pi} x_{\pi t} + \frac{1}{\sigma} x_{dt}. \tag{41}$$

If the Fed sets interest rates to this i_t^*, equilibrium output can always equal potential output.

However, there is no way for the Fed to implement this policy and attain this equilibrium with a rule that does not depend explicitly on shocks and, thus, with an error term that is uncorrelated with available instruments. We could try to substitute endogenous variables for shocks in (41) as far as

$$i_t^* = r - \frac{1 - \rho_{\bar{y}}}{\sigma} y_t^* + \rho_\pi \pi_t^* + \frac{1}{\sigma} x_{dt}$$

and implement i_t^* as a Taylor rule with $\phi_\pi^* = \rho_\pi$ and $\phi_y^* = -(1 - \rho_{\bar{y}})/\sigma$. However, x_{dt} remains in the rule, and since it is not spanned by y_t^* and π_t^*, there is no way to remove it. Thus x_{dt} must become part of the monetary policy disturbance. With no reason to rule out correlation between the disturbances x_{dt} and $x_{\pi t}$, $\bar{y}_t$, nor any reason to limit serial correlation of x_{dt}, we do not have any instruments.

But even this much is false progress. The coefficient $\rho_\pi < 1$, so these attempted values of ϕ_y^* and ϕ_π^* lie outside the zone of determinacy. To try to implement i_t^* as a Taylor rule with coefficients in the zone of determinacy, we have to strengthen the inflation response in an almost silly way:

$$i_t^* = r - \left(\frac{1-\rho_{\bar{y}}}{\sigma}\right) y_t^* + \phi_\pi^* \pi_t^* + \left[(\rho_\pi - \phi_\pi^*)\pi_t^* + \frac{1}{\sigma} x_{dt}\right]. \tag{42}$$

The term in brackets is the new monetary policy disturbance. The right-hand variable is now hopelessly correlated with the error term. (Online App. B shows the same result directly and more generally: assuming a Taylor rule without shocks, you cannot produce the no-gap equilibrium with finite coefficients.) Here, the attempt to equate the correlation between i_t^* and π_t^* in the no-gap equilibrium with the Fed's response to alternative equilibria must fail.

C. *Stochastic Intercept*

The term in brackets in (42) or i_t^* in (34) are often called "stochastic intercepts." In order to attain the no-gap equilibrium in this model, the central bank must follow a policy in which the interest rate reacts directly to some of the structural shocks of the economy, as well as reacting to output and inflation. The stochastic intercept is a crucial part of new-Keynesian policy advice. Woodford (2003), for example, argues for "Wicksellian" policy in which the interest rate target varies following the "natural" rate of interest, determined by real disturbances to the economy, and then varies interest rates with inflation and output so as to produce local uniqueness. King's (2000) expression (34) offers the clearest separation between natural rate and determinacy roles.

Given this fact, it is a substantial restriction to omit the intercept from empirical work and from policy discussion surrounding empirical work. For example, Clarida et al. (2000) and Woodford (2003, chap. 4) calculate the variance of output and inflation using rules with no intercepts and discuss the merits of larger ϕ for reducing such variance. Yet all the time equilibria with zero variance of output or inflation are available, as in the no-gap equilibrium, if only we will allow the policy rule to depend on disturbances directly.

The stochastic intercept of theory is often left out of empirical work because it becomes part of the monetary policy disturbance in that context. It is inextricably correlated with the other structural shocks of the model and hence with the endogenous variables that depend on other shocks of the model. Things were bad enough with genuine monetary policy disturbances—an x_{it} unrelated to other shocks of the

model—because the new-Keynesian model predicts that right-hand variables should jump when there are shocks to this disturbance, as highlighted in Section II. The stochastic intercept makes things even worse because theory then predicts correlation between the composite monetary policy disturbance and other shocks, and other endogenous variables that depend on those shocks.

When one assumes away the stochastic intercept—or, equivalently, assumes that the monetary policy disturbance is uncorrelated with other variables—that assumption is really a restriction on the set of equilibrium paths the economy is following, and it is an assumption on Fed policy that it does not pick, by interest rate policy i_t^*, any of those equilibria. Many equilibria are left out, including the one with no gaps.

This discussion reinforces two general principles: First, do not take error term properties lightly. As Sims (1980) emphasizes, linear models are composed of identical-looking equations, distinguished only by exclusion restrictions and error-orthogonality properties. The IS curve,

$$y_t = E_t y_{t+1} - \sigma(i_t - E_t\pi_{t+1}) + x_{dt},$$

after all, can be rearranged to read

$$i_t = E_t\pi_{t+1} + \frac{1}{\sigma}(E_t y_{t+1} - y_t) + \frac{1}{\sigma}x_{dt}.$$

If we regress interest rates on output and inflation, how do we know that we are recovering the Fed's policy response, and not the parameters of the consumer's first-order condition? Only the orthogonality of the shocks (x_{dt}) with instruments distinguishes the two equations.

Second, orthogonality is a property of the *model* and a property of the *right-hand variables*, not really a property of the errors. You really have to write down a full model to understand why the endogenous right-hand variables or instruments would not respond to the shocks in the monetary policy disturbance.

With this background of possibilities and implicit assumptions, I can review the explicit assumptions in classic estimates.

D. *Clarida et al.*

Clarida et al. (2000) specify an empirical policy rule in partial adjustment form (in my notation)

$$i_t = (1 - \rho_1 - \rho_2)\{r + (\phi_\pi - 1)[E_t(\pi_{t+1}) - \pi] + \phi_y E_t(\Delta y_{t+1} - \Delta\bar{y}_{t+1})\} + \rho_1 i_{t-1} + \rho_2 i_{t-2}, \quad (43)$$

where π is the inflation target (estimated), $\Delta y_{t+1} - \Delta\bar{y}_{t+1}$ is the growth in output gap, and r is the "long-run equilibrium real rate" (estimated)

(see their [4], 153, and table 1, 157). What are the important identification assumptions?

First, *there is no error term*, no monetary policy disturbance at all. The central problem of my simple example is that any monetary policy disturbance is correlated with right-hand variables, since the latter must jump endogenously when there is a monetary policy disturbance. Clarida et al. assume this problem away. They also assume away the stochastic intercept, the component of the monetary policy disturbance that reflects adaptation to other shocks in the economy.

A regression error term appears when Clarida et al. replace expected inflation and output with their ex post realized values, writing

$$\begin{aligned} i_t = {} & (1 - \rho_1 - \rho_2)[r + (\phi_\pi - 1)(\pi_{t+1} - \pi) + \phi_y \Delta y_{t+1}] \\ & + \rho_1 i_{t-1} + \rho_2 i_{t-2} + \varepsilon_{t+1}. \end{aligned} \tag{44}$$

In this way, they avoid the 100 percent R^2 prediction, which normally results from assuming away a regression disturbance. The remaining error ε_{t+1} is a pure forecast error, so it is serially uncorrelated. This fact allows Clarida et al. to use variables observed at time t as instruments to remove correlation between the forecast error ε_{t+1} and the ex post values of the right-hand variables π_{t+1} and Δy_{t+1}. Validity for this purpose does not mean that such instruments would be valid if we were to recognize a genuine monetary policy disturbance.

Clarida, Galí, and Gertler (1998) consider a slightly more general specification that does include a monetary policy disturbance.[1] In this case, they specify (their eq. [2.5], my notation)

$$i_t = \rho i_{t-1} + (1 - \rho)[\alpha + \phi_\pi E(\pi_{t,t+n}|\Omega_t) + \phi_y E(y_t - \bar{y}_t|\Omega_t)] + v_t, \tag{45}$$

where $\bar{y}$ denotes potential output, separately measured, and Ω_t is the central bank's information set at time t, which they assume does not include current output y_t. Now v_t is the monetary policy disturbance, defined as "an exogenous random shock to the interest rate" (1039). They add, "Importantly, we assume that v_t is i.i.d." They estimate (45) by instrumental variables, using lagged output, inflation, interest rates, and commodity prices as instruments.

Obviously, the assumption of an i.i.d. disturbance is key and restrictive. For example, many commentators accuse the Fed of deviating from the Taylor rule for years at a time in the mid-2000s. This assumption also means that any other shocks in the monetary policy disturbance—stochastic intercepts, variation in the natural rate—are also i.i.d. There is no reason why preference shifts (IS curve) or marginal cost shocks (Phillips curve shifts) should be i.i.d. But some other shock must not

[1] Curiously, Clarida et al. (2000) mention the disturbance v_t below their eq. (3), p. 153, but it does not appear in the equations or the following discussion. I presume that the mention of v_t is a typo in the 2000 paper.

be i.i.d., so that there is persistent variation in the right-hand variables. Therefore, the monetary policy disturbance must not include a "stochastic intercept" that responds to the non-i.i.d. shocks.

E. Giannoni, Rotemberg, Woodford

Rotemberg and Woodford (1997, 1998, 1999), followed by Giannoni and Woodford (2005; see also the summary in Woodford [2003, chap. 5]), follow a different identification strategy, which allows them to estimate the parameters of the Taylor rule by ordinary least squares (OLS) rather than instrumental variable (IV) regressions or maximum likelihood or other system methods. Giannoni and Woodford write the form of the Taylor rule in these papers:

> We assume that the recent U.S. monetary policy can be described by the following feedback rule for the Federal funds rate
>
> $$i_t = \bar{i} + \sum_{k=1}^{n} \phi_{ik}(i_{t-k} - \bar{i}) + \sum_{k=0}^{n_w} \phi_{wk}\hat{w}_{t-k} + \sum_{k=0}^{n_\pi} \phi_{\pi k}(\pi_{t-k} - \bar{\pi}) + \sum_{k=0}^{n_y} \phi_{yk}\hat{Y}_{t-k} + \varepsilon_t, \quad (46)$$
>
> where i_t is the Federal funds rate in period t; π_t denotes the rate of inflation between periods $t-1$ and t; $\hat{w}_t$ is the deviation of the log real wage from trend at date t, $\hat{Y}_t$ is the deviation of log real GDP from trend, $\bar{i}$ and $\bar{\pi}$ are long-run average values of the respective variables. The disturbances ε_t represent monetary policy "shocks" and are assumed to be serially uncorrelated. . . . To identify the monetary policy shocks and estimate the coefficients in [(46)], we assume . . . that a monetary policy shock at date t has no effect on inflation, output or the real wage in that period. It follows that [(46)] can be estimated by OLS. (36–37)

Since they lay out the assumptions that identify this policy rule with such clarity, we can easily examine their plausibility. First, they assume that the monetary policy disturbance ε_t is i.i.d.—uncorrelated with lags of itself and past values of the right-hand variables. This is again a strong assumption, given that ε_t is not a forecast error, but instead represents rate or structural disturbances.

Second, they assume that the disturbance ε_t is also not correlated with contemporaneous values of $\hat{w}_t$, π_t, and $\hat{Y}_t$. This is an especially surprising result of a new-Keynesian model because $\hat{w}_t$, π_t, and $\hat{Y}_t$ are endogenous variables. From the very simplest model in this paper, endogenous variables have jumped in the new-Keynesian equilibrium when there is a monetary policy (or any other) disturbance. How can $\hat{w}_t$, π_t, and $\hat{Y}_t$

not jump when there is a shock ε_t? To achieve this result, Giannoni, Rotemberg, and Woodford assume as part of their economic model that $\hat{w}_t$, π_t, and $\hat{Y}_t$ must be predetermined by at least 1 quarter, so they cannot move when ε_t moves. (In the model as described in their technical appendix, output $\hat{Y}$ is actually fixed 2 quarters in advance, and the marginal utility of consumption μ_t is also fixed 1 quarter in advance.) It is admirable that Giannoni, Rotemberg, and Woodford explain the properties of the *model* that generate the needed correlation properties of the *instruments*. But needless to say, these are strong assumptions. Are wages, prices, output, and marginal utility really fixed 1–2 quarters in advance in our economy, and therefore unable to react within the quarter to monetary policy disturbances? They certainly are not forecastable 1–2 quarters in advance!

Most of all, if $\hat{w}_t$, π_t, and $\hat{Y}_t$ do not jump when there is a monetary policy disturbance, *something else must jump* to head off the explosive equilibria. What jump in this model are expectations of future values of these variables, among others, $\hat{w}_{t+1} = E_t\hat{w}_{t+1}$, $\pi_{t+1} = E_t\pi_{t+1}$, and $\hat{Y}_{t+2} = E_t\hat{Y}_{t+2}$ as well as the state variable $E_t\mu_{t+1}$, the marginal utility of consumption. All of these variables are determined at date t. Now, we see another implicit assumption in the policy function (46): none of these expected future variables are present in the policy rule. Thus, Giannoni, Rotemberg, and Woodford achieve identification by a classic exclusion restriction. In contrast to the literature that argues for the empirical necessity and theoretical desirability of Taylor rules that react to expected future output and inflation and to other variables that the central bank can observe, they assume those reactions to be absent.

In sum, Giannoni and Woodford identify the Taylor rule in their model by two assumptions about Fed behavior and one assumption about the economy: (1) the disturbance, including "natural rate" "stochastic intercept" reactions to other shocks, is not predictable by any variables at time $t-1$; (2) the Fed does not react to expected future output or wage, price inflation, or other state variables; and (3) wages, prices, and output are fixed a period in advance.

VI. Old-Keynesian Models

Determinacy and identification are properties of specific models, not general properties of variables and parameters. Old-Keynesian models reverse many of the determinacy and identification propositions. In these models, an inflation coefficient greater than one is the key for *stable* dynamics, to produce system eigenvalues *less* than one, and to solve the model *backward*. Since the model does not have expected future terms, such a backward solution gives determinacy. The policy rules are

identified, at least up to the usual (Sims 1980) issues with simultaneous-equation macro models.

I think that much of the determinacy and identification confusion stems from misunderstanding the profound differences between new-Keynesian and old-Keynesian models. Alas, the old-Keynesian models lack economic foundations and so cannot be a serious competitor for the basic question I started with: What economic force, fundamentally, determines the price level or inflation rate?

Taylor (1999) gives us a nice explicit example of an "old-Keynesian" model (my terminology) that forms a good basis for explicit discussion of these points. (As everywhere else, this is just a good example, not a critique of a specific paper; hundreds of authors adopt old-Keynesian models.) Taylor adopts a "simple model" (662; in my notation):

$$y_t = -\sigma(i_t - \pi_t - r) + u_t, \tag{47}$$

$$\pi_t = \pi_{t-1} + \gamma y_{t-1} + e_t, \tag{48}$$

$$i_t = r + \phi_\pi \pi_t + \phi_y y_t. \tag{49}$$

We see a striking difference: *all the forward-looking terms are absent.*

Taylor states that

> it is crucial to have the interest rate response coefficient on the inflation rate . . . above a critical "stability threshold" of one. . . . The case on the left [$\phi_\pi > 1$] is the stable case. . . . The case on the right [$\phi_\pi < 1$] is unstable. . . . This relationship between the stability of inflation and the size of the interest rate coefficient in the policy rule is a basic prediction of monetary models used for policy evaluation research. In fact, because many models are dynamically unstable when ϕ_π is less than one . . . the simulations of the models usually assume thatϕ_π is greater than one. (663, 664)

This is exactly the opposite philosophy from the new-Keynesian models. In new-Keynesian models, $\phi_\pi > 1$ is the condition for a "dynamically unstable" model. New-Keynesian models want *unstable* dynamics in order to rule out multiple equilibria and force forward-looking solutions. In Taylor's model, $\phi_\pi > 1$ is the condition for *stable* dynamics, eigenvalues less than one, in which we solve for endogenous variables (including inflation) by *backward-looking* solutions. The condition "$\phi_\pi > 1$" sounds superficially similar, but in fact its operation is diametrically the opposite. Taylor is worried about "spirals," not about "determinacy."

A little more formally, and to parallel the analysis of the new-Keynesian model following (26)–(28), the standard form of Taylor's model is

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \sigma\gamma\dfrac{1-\phi_\pi}{1+\sigma\phi_y} & \sigma\dfrac{1-\phi_\pi}{1+\sigma\phi_y} \\ \gamma & 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} \dfrac{1}{1+\sigma\phi_y} & \sigma\dfrac{1-\phi_\pi}{1+\sigma\phi_y} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{50}$$

(substitute [49] into [47]). The eigenvalues of this transition matrix are

$$\lambda_1 = 1 + \sigma\gamma\frac{1-\phi_\pi}{1+\sigma\phi_y}; \quad \lambda_2 = 0.$$

Therefore, $\phi_\pi > 1$ (with the natural restrictions $\phi_y > -1/\sigma$, $\sigma\gamma > 0$) generates values of the first eigenvalue less than one. Following the usual decomposition, we can then write the unique solution of the model as a backward-looking average of its shocks:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \frac{1}{1+\sigma\phi_y+\sigma\gamma(1-\phi_\pi)} \begin{bmatrix} \lambda_1 - 1 & \lambda_1 \\ 1+\sigma\phi_y & 0 \end{bmatrix} \sum_{j=0}^{\infty} \lambda_1^j \begin{bmatrix} u_{t-j} \\ e_{t-j} \end{bmatrix}.$$

There is no multiple-equilibrium or indeterminacy issue in this backward-looking solution.

More intuitively, take $\phi_y = 0$ and assume $\phi_\pi > 1$. Then if inflation π_t rises in the policy rule (49), the Fed ends up raising the real rate (as defined here without forward-looking terms) $i_t - \pi_t$. In the IS curve (47) this lowers output y_t, and lower output in the Phillips curve (48) lowers $E_t(\pi_{t+1})$. This model thus embodies the classic concept of "stabilization" that more inflation makes the Fed raise real interest rates, which lowers demand and lowers future inflation. This is exactly the opposite of the new-Keynesian dynamics. In the new-Keynesian model, a rise in inflation π_t leads to an explosion; "stabilization" by $\phi_\pi > 1$ means that we count on π_t to have jumped to the unique value that heads off such explosions.

Why do the two models disagree so much on the desired kind of dynamics? The equations of Taylor's model have no expected future terms. Hence, there are no expectational errors. All the shocks (u_t, e_t) driving the system are exogenous economic disturbances. By contrast, the new-Keynesian model has expected future values in its "structural" equations (26)–(28), so the shocks in its standard representation such as (33) contain expectational errors. The difference is not happenstance; the whole point of the new-Keynesian enterprise is to microfound behavioral relationships, and microfounded behavior is driven by expectations of the future, not memory of the past.

Taylor regards this model as a "reduced form" in which expectations have been "solved out." He claims that nonetheless "these equations summarize more complex forward-looking models" (662). I do not think

this is true. Taylor's model is fundamentally different, not a simpler reduced-form or rough guide to give intuition formalized by a more complex new-Keynesian effort. The difference between this model and the new-Keynesian model (26)–(28) is not about policy invariance. We want to analyze dynamics for given policy parameters ϕ_π, ϕ_y. Even if γ, σ, and r change with different ϕ_π, ϕ_y, they are constant for a given ϕ_π, ϕ_y. Equations (47)–(49) are not a simpler or reduced-form version of (26)–(28). They are the same equations—with the same algebraic complexity—with different t subscripts. Different t subscripts dramatically change dynamics, including stability and determinacy.

The operation of the models is completely different. The response to shocks in an old-Keynesian model represents the means by which structural equations are brought back to balance. The response to shocks of a new-Keynesian model represents a jump to a different equilibrium, a choice among many different possibilities, in each of which the structural equations are all in balance. Determinacy is not even an issue in Taylor's model. His model always has one equilibrium. The issue is "spirals," whether that equilibrium is stable. King (2000, 72) also details a number of fundamental differences between new- and old-Keynesian models of this sort.

New-Keynesian models and results are often described with old-Keynesian intuition. This is a mistake.

Identification in Taylor's model does not suffer the central problem of identification in new-Keynesian models. The behavior we are assessing is not how the Fed would respond to the emergence of alternative equilibrium paths; it is completely revealed by the Fed's behavior in equilibrium. The parameters ϕ_π, ϕ_y appear in the equilibrium dynamics (50) and hence the likelihood function. That does not mean that identification is easy; it means we "only" have to face the standard issues in simultaneous-equation models as reviewed by Sims (1980) and studied extensively by the vector autoregression (VAR) literature since then.

Since it easily delivers a unique equilibrium, why not conclude that Taylor's model is the right one to use? Alas, our quest is for *economic* models of price determinacy. This model fails on the crucial qualification—as Taylor's (p. 662) discussion makes very clear. If in fact inflation has nothing to do with expected future inflation, so inflation is mechanistically *caused* by output gaps, and if in fact the Fed controls the output gap by changing interest rates, then, yes, the Taylor rule does lead to inflation determinacy. But despite a half century of looking for them, economic models do not deliver the "if" part of these statements. If we follow this model, we are giving up on an economic understanding of price-level determination in favor of (at best) a mechanistic description.

VII. Extensions and Responses

Online Appendix B contains an extensive critical review of the literature, responses to many objections, and extensions left out of the text for reasons of space. If you want to know "What about *x*'s approach to determinacy or identification?" you are likely to find an answer there.

I investigate identification in the other equilibria of the simple model. For $\|\phi\| < 1$, ϕ is not identified in any equilibrium. For $\|\phi\| > 1$, however, you can identify ϕ for every equilibrium *except* the new-Keynesian local equilibrium. If explosions occur, you can measure their rate.

I explore the impulse-response functions of the simple model and contrast new-Keynesian and non-Ricardian choices in terms of impulse-response functions. I generalize the simple model to include an IS shock in (1). In this case, we cannot even estimate ρ.

I address the question, what happens if you run Taylor rule regressions in artificial data from fuller (three-equation) new-Keynesian models? This discussion generalizes the finding in Section II in which regressions recovered the shock autocorrelation process rather than the Taylor rule parameter to the three-equation model. Unsurprisingly, Taylor rule regressions do not recover Taylor rule parameters in artificial data from typical models.

One may ask, "well, if not a change in the Taylor rule, what did Clarida et al. (2000) measure?" The right answer is really "it doesn't matter": once a coefficient loses its structural interpretation, who cares how it comes out? Or perhaps, "you need a different model to interpret the coefficient." However, online Appendix B gives an example of how other changes in behavior can show up as spurious Taylor rule changes. In the example, the Taylor rule coefficient is constant at $\phi = 1.1$, but the Fed gets better at offsetting IS shocks, that is, following better the natural rate. This change in policy causes mismeasured Taylor rule coefficients to rise as they do in the data.

An obvious question is whether full likelihood approaches, involving dynamics of the entire model, might be able to identify parameters in which the single-equation methods I surveyed here are faltering. Equivalently, perhaps the impulse-response function to other shocks can identify the Taylor rule parameters (or, more generally, system eigenvalues). I survey these issues. While identification in full systems has been studied and criticized, nobody has tried to use full-system methods to test for determinacy. This literature *imposes* determinacy and explores model specification to better fit second moments. This is not a criticism; "fitting the data" rather than "testing the model" is a worthy goal. But it means I have no useful results to report or literature to review on whether this approach can overcome identification problems in order to test for determinacy.

I explore leads and lags in Taylor rules in the context of the simple frictionless model, continuous-time models, and the three-equation model. It turns out that determinacy questions depend quite sensitively on the timing assumptions in the Taylor rule. The problem is particularly evident on taking the continuous-time limit. Given that changing a time index, for example, $E_t\pi_{t+1}$ in place of π_{t-1}, can reverse stability properties, this finding is not surprising, but it does counter the impression that new-Keynesian Taylor rule determinacy is robust to changes in specification.

VIII. Conclusions and Implications

A. *Determinacy*

Practically all verbal explanations for the wisdom of the Taylor principle—the Fed should increase interest rates more than one-for-one with inflation—use old-Keynesian, stabilizing, logic: This action will raise real interest rates, which will dampen demand, which will lower future inflation. New-Keynesian models operate in an entirely different manner: By raising interest rates in response to inflation, the Fed induces accelerating inflation or deflation, or at a minimum a large "nonlocal" movement, unless inflation today jumps to one particular value.

Alas, there is no economic reason why the economy should pick this unique initial value, as inflation and deflation are valid economic equilibria. No supply/demand force acts to move inflation to this value. The attempts to rule out multiple equilibria basically state that the government will blow up the economy should accelerating inflation or deflation occur. This is not a reasonable characterization of anyone's expectations. Such policies also violate the usual criterion that the government must operate in markets just like agents. I conclude that inflation is just as indeterminate, in microfounded new-Keynesian models, when the central bank follows a Taylor rule with a Ricardian fiscal regime, as it is under fixed interest rate targets.

The literature—understandably, I think—confused "stopping an inflation" with "ruling out an equilibrium path." Alas, now that confusion is lifted, we can see that the latter goal is not achieved.

B. *Identification*

The central empirical success of new-Keynesian models is estimates such as Clarida et al.'s (2000) that say inflation was stabilized in the United States by a switch from an "indeterminate" to a "determinate" regime. The crucial Taylor rule parameter is not identified in the new-Keynesian model, so we cannot interpret regressions in this way. The new-Keynesian

model has nothing to say about inflation in an indeterminate regime, so Taylor rule regressions in the 1970s are doubly uninterpretable in the new-Keynesian context.

Clarida et al.'s coefficients of interest rates on inflation range from 2.15 (table 4) to as much as 3.13 (table 5). These coefficients are a lot greater than one. These coefficients imply that if the United States returned to the 12 percent inflation of the late 1970s (a 10-percentage-point rise), the Federal Reserve would raise the funds rate by 21.5–31.3 percentage points. If these predictions seem implausibly large, digesting the estimates as something less than structural helps a great deal.

The identification issue stems from the heart of all new-Keynesian models with Ricardian fiscal regimes. The models have multiple equilibria. The modelers specify policy rules that lead to explosive dynamics and then pick only the locally bounded equilibrium. But locally bounded equilibrium variables are stationary and so cannot reveal the strength of the explosions, which occur only in the equilibria we do not observe.

Endogenous variables are supposed to jump in response to disturbances, to head off explosions. Such jumps induce correlation between right-hand variables of the policy rule and its error, so that rule will be exquisitely hard to estimate. One can only begin to get around these central problems by strong assumptions, in particular that the central bank does not respond to many variables, and to natural rate shocks in particular, in ways that would help it to stabilize the economy.

The literature—understandably, I think—did not appreciate that "determinacy" and "desirable rate in equilibrium" are separate issues; that new-Keynesian models, unlike their old-Keynesian counterparts, achieve determinacy by responses to alternative equilibria, which are not measurable, not by responses to equilibrium variation in inflation, which are; that "achieving determinacy" is a different reading of history than "raising rates to lower inflation"; and that "determinacy"—eliminating multiple equilibria—is different from "stability"—avoiding inflationary or deflationary "spirals." Again, however, now that the distinction is clear, we need not continue to misinterpret the regressions.

C. If Not This, Then What?

The contribution of this paper is negative, establishing that one popular theory does not, in the end, determine the price level or the inflation rate. So what theory can determine the price level, in an economy like ours? Commodity standards and $MV = PY$ can work in theory but do not apply to our economy, with fiat money, interest-elastic money demand, and no attempt by the central bank to target quantities.

The price level can be determined for economies like ours in models that adopt—or, perhaps, recognize—that governments follow a fiscal

regime that is at least partially non-Ricardian. Such models solve all the determinacy and uniqueness problems in one fell swoop. And the change is not really so radical. Though the deep question of where the price level comes from changes, the vast majority of the new-Keynesian ingredients can be maintained. Whether the results are the same is an open question.

"Economic" is an important qualifier. Most of the case for Taylor rules in popular and central bank writing, in FOMC statements, and too often in academic contexts emphasizes the old-Keynesian stabilizing story. This is a pleasant and intuitively pleasing story to many. However, it throws out the edifice of theoretical coherence—explicit underpinnings of optimizing agents, budget constraints, clearing markets, and so forth—that is the hallmark achievement of the new-Keynesian effort. If inflation is, in fact, stabilized in modern economies by interest rate targets interacted with backward-looking IS and Phillips curves, economists really have no idea why this is so.

Appendix A

This appendix reviews the parallel question of inflations with constant money supply and interest-elastic demand. I verify that standard proposals suffer the same problems described in the text. I conclude that models with fixed money, interest-elastic demand, and Ricardian fiscal policies have the same indeterminacies as new-Keynesian models.

Obstfeld and Rogoff (1983) are often cited as the standard way to eliminate hyperinflationary equilibria in such models, for example by Woodford (2003, 138) and Atkeson et al. (2010). The main idea for which they are cited is that the government switches to a commodity standard when inflation gets out of hand. Their actual idea is different, but it is worth examining both the general idea and their specific example.

A. *Simple Example: Cagan Dynamics*

I use the simplest possible example. (Minford and Srinivasan [2010] is a recent paper that uses this framework.) Suppose that money supply m is constant, money demand is interest elastic, and the real rate is constant and zero. Then the log price-level path must satisfy

$$m = m_t = p_t - \alpha(E_t p_{t+1} - p_t)$$

or, rearranging,

$$(E_t p_{t+1} - m) = \gamma(p_t - m); \quad \gamma \equiv \left(\frac{1+\alpha}{\alpha}\right).$$

The term $p_t = m$ is an equilibrium, but there are many others. Any path

$$(p_{t+1} - m) = \gamma(p_t - m) + \delta_{t+1}$$

with $E_t(\delta_{t+1}) = 0$ is possible. If $p_t > m$, then we expect a hyperinflation, and con-

versely. To conclude $p_t = m$, we need some device to disallow the inflationary or deflationary equilibria.

With a commodity standard (and sufficient fiscal backing) in place of the money target, the price level is nailed at whatever value p^* the government chooses, but money is endogenous in the quantity $m^* = p^*$.

To stop a hyperinflation, the central bank can switch to a commodity standard. For example, the price level path $p_0 = m + 1$, $p_1 = m + \gamma$, $p_2 = m + \gamma^2$, ..., $p_T = m + \gamma^T = \bar{p}$ followed by $p_{T+1} = p_{T+2} = \cdots = \bar{p}$ is an equilibrium if the central bank switches to a commodity standard at the level $\bar{p}$. Of course, the government then must allow the money supply to expand passively to $\bar{m} = \bar{p} = m + \gamma^T$. The money stock on this equilibrium path is $m_0 = m_1 = m_2 = m_{T-1} = m$, $m_T = m_{T+1} = \ldots = \bar{p} = m + \gamma^T > m$. As in real hyperinflations, both real and nominal money balances expand when the hyperinflation is stopped.

That switch stops the inflation, but the inflation and its end still represent an equilibrium path since first-order conditions are satisfied at every date. To rule out such paths as equilibria, we have to add something else. New-Keynesian models ruled out such equilibrium paths by insisting on a Taylor rule and commodity standard at incompatible values. The analogue here is to assume that the government also keeps intact the money stock target m while it nails the price level to $\bar{p}$ with a commodity standard. Once again, that is a blow-up-the-world policy, impossible by Ramsey rules, since a commodity standard requires the government to freely buy and sell currency. Once again, it is a choice, since the standard policy that allows the real money stock to increase is available. As with the Taylor rule, however, a commitment to return to a fixed price *level*, with needed fiscal backing, could rule out the inflation.

B. *Obstfeld and Rogoff*

Obstfeld and Rogoff's (1983) actual analysis is quite different. Their government does not attempt to stabilize inflation by introducing a commodity standard or other means. Instead, their government buys up all the money stock and leaves the economy to barter—zero money, infinite price level—thereafter.

Obstfeld and Rogoff start with an economy that can hyperinflate to an infinite price level in finite time, jumping from $P_T = \bar{P}$ (defined below) to $P_{T+1} = \infty$ in one step. Figure A1 plots this path, labeled "solution with $\varepsilon = 0$." The figure plots $m_t = M/P_t$ with $M = 1$ for clarity, so a jump to $P_{T+1} = \infty$ is a jump of m_{T+1} to zero.

Obstfeld and Rogoff claim to remove this equilibrium by a small change: The government offers to buy back the money stock in return for ε consumption goods per dollar. With this guarantee, they claim that their economy needs a period $P_{T+1} = \bar{\bar{P}} = 1/\varepsilon$ during which money is repurchased before going on to $P_{T+2} = \infty$ and thereafter. Figure A1 shows this path as well, marked "Obstfeld-Rogoff, $\varepsilon = 0.5$." (I use a rather large ε so that the paths are distinguishable on the graph.) They claim, however, that no matter how small ε, first-order conditions are violated in period $T + 1$, so this equilibrium path is ruled out.

Alas, this result is wrong. In period $T + 1$, when consumers sell all their money back to the government, the first-order condition studied by Obstfeld and Rogoff

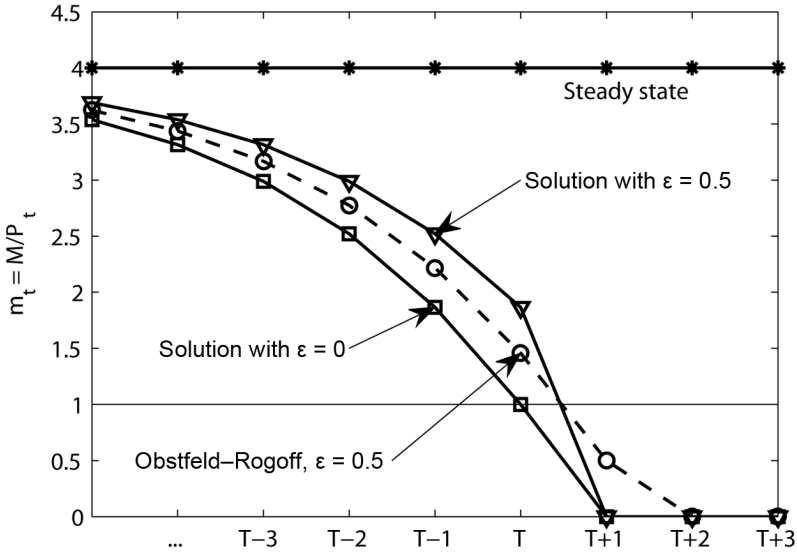

FIG. A1.—Hyperinflations in the Obstfeld-Rogoff model. "Solution with $\varepsilon = 0$" gives the hyperinflation we wish to rule out. "Obstfeld-Rogoff" gives Obstfeld and Rogoff's path when the government offers to redeem the currency for ε units of consumption good. "Solution with $\varepsilon = 0.5$" gives the actual path in that case. The lower horizontal line indicates $M/\bar{P}$; $u'(y) = 1$, $M = 1$, $\beta = 1/2$, and $v(m) = m^{-1/2}$.

no longer applies. In this regular first-order condition, the consumer thinks about holding a bit more money, enjoying its transactions services, and then getting rid of it the next day. When the consumer sells all his money back to the government for ε consumption goods per dollar, however, the "next-day" margin is absent. Instead, he enjoys the marginal utility of the ε consumption goods tendered by the government.

This correct first-order condition does hold in this period, and equilibrium still holds at every date under Obstfeld and Rogoff's repurchase offer. An equilibrium exists for every offer ε, and price paths are continuous in ε. There is no discontinuity in the existence of equilibrium at $\varepsilon = 0$. This equilibrium is labeled "solution with $\varepsilon = 0.5$" in figure A1.

Here is how the equilibrium with repurchase offer works: We still have $P_{T+1} = \infty$. Then P_T is slightly higher than it was without the repurchase offer. Previously, at T, the consumer was happy to hold money despite the fact that it would be worth nothing at the beginning of the next period $T + 1$, because the marginal transactions value was so high. Now, he gets a slight extra benefit of holding money that he can redeem at the end of the period. This fact makes money slightly more valuable at the beginning of the period. Previous periods $T - 1$ and so on follow the usual difference equation with inflationary dynamics.

The solution is more intuitive in retrospect. How could offering one kernel of corn for a billion dollars destroy an equilibrium? Given that people were

holding money at T that they knew would be worthless at $T+1$, why would a tiny residual value make any difference? It doesn't.

Here is the analysis in detail. Obstfeld and Rogoff assume that consumers maximize a standard utility function defined over consumption and real money balances,

$$\sum_{t=0}^{\infty} \beta^t[u(c_t) + v(m_t)]; \quad m_t \equiv M_t/P_t.$$

They introduce capital distinct from consumption, but this bit of realism is not relevant here, so I specialize to a capital price $q_t = 1$.

The consumer's first-order conditions are

$$\frac{u'(c_t)}{P_t} = \frac{v'(M_t/P_t)}{P_t} + \beta \frac{u'(c_{t+1})}{P_{t+1}},$$

$$u'(c_t) = \beta(1 + r)u'(c_{t+1}).$$

(Obstfeld and Rogoff also study carefully the transversality conditions, but those are not at issue here.) There is a constant endowment, so equilibrium requires $c_t = y$, and the equilibrium conditions become (their eq. [14])

$$\frac{u'(y)}{P_t} - \frac{v'(M_t/P_t)}{P_t} = \beta \frac{u'(y)}{P_{t+1}}. \tag{A1}$$

Obstfeld and Rogoff study money targets; they assume that "the money supply is constant at level M" (678). The corresponding steady-state price level $\overline{P}$ satisfies

$$u'(y) - v'(M/\overline{P}) = \beta u'(y).$$

However, many other sequences $\{P_t\}$ satisfy the equilibrium condition (A1). These sequences also satisfy the transversality condition, discussed in Obstfeld and Rogoff's section 2. As in the new-Keynesian model, the hyperinflations are economically viable equilibria without further policy specification.

Obstfeld and Rogoff study a special case of this model, in which $v(m)$ satisfies the Inada condition $\lim_{m \to 0} v'(m) = \infty$. (They also assume $\lim_{m \to 0} mv'(m) = 0$.) As a result of this assumption, at very high but finite price levels P_t, we have $v'(M/P_t) > u'(y)$. Here, real money balances are so marginally valuable, people are willing to hold money for a period even if it will be valueless the next day. Define the cutoff point for this behavior $\bar{P}$ where

$$\bar{P} : u'(y) - v'(M/\bar{P}) = 0.$$

(This cutoff point $\bar{P}$ uses a small bar; the steady state $\overline{P}$ uses a big bar. This is Obstfeld and Rogoff's notation.) Given $P_{T+1} = \infty$, we will observe $P_T = \bar{P}$. Since money cannot be worth negative amounts in the future, we never observe a price level higher than $\bar{P}$. (Think about this equilibrium as the way in which expectations of P_{T+1} determine equilibrium prices at T.)

Obstfeld and Rogoff study a special kind of hyperinflation in which the price level increases steadily following the difference equation (A1), attains a value $P_T = \bar{P}$ where money is so scarce people hold it only for one period's transactions value, and then jumps to $P_{T+1} = \infty$ forever after (top of p. 681, their fig. 2 and shown as "solution with $\varepsilon = 0$" in fig. A1).

To trim these equilibria, Obstfeld and Rogoff assume that "the government

promises to redeem each dollar bill for ε units of capital [equal to consumption in my simplification], but does not offer to sell money for capital" (684). They assume $\underline{\overline{P}} \equiv 1/\varepsilon > \bar{P}$, so that a price level $\underline{\overline{P}}$ is inconsistent with the first-order condition (A1) and the money target, $u'(y) - v'(M/\underline{\overline{P}}) < 0$.

Here is their central claim that with this extra provision, hyperinflationary equilibrium paths are ruled out: "Suppose that $\{P_t\}$ is an equilibrium path with $P_0 > \underline{\overline{P}}$. Let $P_T = \max\{P_t | P_t < \underline{\overline{P}}\}$. By (14) [my (A1)] P_T must be below $\bar{P}$, so that $u'(y) - v'(M/P_T) > 0$ while P_{T+1} must exceed P_T and therefore equal $\underline{\overline{P}}$. But there is no $M_{T+1} \leq M$ such that $u'(y) - v'(M_{T+1}/\underline{\overline{P}}) \geq 0$. Thus there is no price level P_{T+2} satisfying (14) and $\{P_t\}$ is not an equilibrium path" (685).

Everything follows (A1) backward from a final period in which the government buys up all the money at $\underline{\overline{P}}$ and after which $P = \infty$. During that final period, the first-order condition (A1) cannot be not satisfied because $\underline{\overline{P}} > \bar{P}$. We can see equilibria at $P_T = \infty$ and $P_T = \bar{P}$, but not in between.

The trouble with this analysis is that the first-order condition (A1) is wrong. It does not apply when people redeem money for a real commodity. It assumes that the consumer holds all his money from time $T+1$ to time $T+2$. Obstfeld and Rogoff left the option to tender money to the government out of their budget constraint, along with the constraint that money held overnight and money tendered to the government must each be nonnegative and the latter less than money holdings. It is not true that in a period in which the government buys back money, the equilibrium price level in that period must be $\underline{\overline{P}} = 1/\varepsilon$ and be governed by the first-order condition (A1).

To get it right, we have to be specific about timing. I assume that the consumer receives the benefit of money holding $v(M/P)$ in the period in which he redeems money, that is, that money is redeemed by the government at the end of the period. Equivalently, we can specify an intraday timing. The offer to buy back money is good at any time during the day, so it will always be optimal to redeem money at the end of the day, after receiving $v(M/P)$ and before money loses value overnight. The opposite assumption, that consumers do not get $v(M/P)$ or must redeem at the beginning of the day, just changes the dating convention, not the basic argument.

If the consumer consumes one unit less, holds a bit more money this period, and then reduces money holdings next period and the nonnegativity constraints are satisfied, the first-order condition is the same as (A1):

$$\frac{u'(c_t)}{P_t} = v'\left(\frac{M_t}{P_t}\right)\frac{1}{P_t} + \beta\frac{u'(c_{t+1})}{P_{t+1}}. \tag{A2}$$

However, if the consumer consumes one unit less, holds a bit more money this period, and then sells it to the government at the end of the period for $\varepsilon = 1/\underline{\overline{P}}$ consumption goods, his first-order condition is

$$\frac{u'(c_t)}{P_t} = v'\left(\frac{M_t}{P_t}\right)\frac{1}{P_t} + \frac{u'(c_t)}{\underline{\overline{P}}}. \tag{A3}$$

Condition (A2) holds if $\beta u'(c_{t+1})/P_{t+1} > u'(c_t)/\underline{\overline{P}}$, in which case the consumer sells nothing to the government. Condition (A3) holds if

$$\beta u'(c_{t+1})/P_{t+1} < u'(c_t)/\underline{\overline{P}}$$

and in particular if $P_{t+1} = \infty$, in which case the consumer holds nothing overnight and sells everything to the government.

It is still true that $P_{T+1} = \bar{\bar{P}}$, $P_{T+2} = \infty$ is not an equilibrium. By (A3), $P_{T+1} = \bar{\bar{P}}$ implies $v'(m) = 0$. If people know they can put their money back to the government for consumption goods at the same rate they can acquire money by reducing consumption, it is as if there is no interest cost to holding money. Only complete satiation can be an equilibrium in this circumstance. Thus, Obstfeld and Rogoff's period with $P_{T+1} = \bar{\bar{P}}$ and $v'(M/\bar{\bar{P}}) > u'(y) > 0$ cannot happen, consistent with their claim.

However, it is not true that we must observe $P_T = \bar{\bar{P}} = 1/\varepsilon$ in the repurchase period, followed by $P_{T+1} = \infty$. Money can trade during a period at a higher value than that which the government offers in redemption at the end of the period. At $P_T = \bar{P}$, people were willing to hold money despite zero value the following day, and this did not violate "arbitrage." Hence, they are willing to hold money during the day that has greater value than it will have when the government repurchases the money at the end of the day. When the buyback is in place with $\varepsilon > 0$, we observe an equilibrium with $P_T < \bar{P}$, not equal to or above $\bar{P}$.

Here, then, is how the hyperinflationary equilibrium actually ends, with the buyback guarantee in place: $P_{T+1} = \infty$. Knowing this, at T, people redeem all their money at the end of the period T, so (A3) is the relevant first-order condition. Rearranging (A3) in equilibrium ($c = y$, $M_t = M$), we get

$$u'(y)\left(1 - \frac{P_T}{\bar{\bar{P}}}\right) = v'\left(\frac{M}{P_T}\right).$$

This condition determines P_T. If $\varepsilon = 0$ so $\bar{\bar{P}} = \infty$, then $P_T = \bar{P}$, $v'(m_{T+1}) = \underline{u}'(y)$, and this is the equilibrium with no buyback. A small ε means a large $\bar{\bar{P}}$, so $P_T < \bar{P}$. Periods prior to T follow the usual difference equation (A1). This path *is* an equilibrium and *is not* ruled out by the repurchase offer.

The central problem is Obstfeld and Rogoff's "arbitrage" condition (685) that $\bar{\bar{P}} = P_T$ in any period that people are tendering money. That argument is not valid in this discrete-time model because people can get $v(m)$ plus the redemption value. This arbitrage argument would be valid in a continuous-time version of the model, and perhaps the error comes from mixing correct continuous-time intuition with a discrete-time model. However, a continuous-time version of the same proof does not work because the first-order conditions are different. If utility is

$$\int_{t=0}^{\infty} e^{-\delta t}[u(c_t) + v(M_t/P_t)]dt,$$

the first-order condition corresponding to (A1) is

$$\frac{v'(M_t/P_t)}{u'(y)} = \delta + \frac{1}{P_t}\frac{dP_t}{dt}. \tag{A4}$$

Now, $\underline{v}'(m)$ can rise to arbitrarily large values with a differentiable price path. Then $\bar{\bar{P}}$ is a valid equilibrium price. The inflationary price path described by (A4), terminated by a tender when $P_T = \bar{\bar{P}}$, is a valid equilibrium.

References

Alstadheim, Ragna, and Dale W. Henderson. 2006. "Price-Level Determinacy, Lower Bounds on the Nominal Interest Rate, and Liquidity Traps." *B.E. J. Macroeconomics* (Contributions) 6 (1): art. 12. http://www.bepress.com/bejm/contributions/vol6/iss1/art12.

Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe. 2010. "Sophisticated Monetary Policies." *Q.J.E.* 125:47–89.

Bassetto, Marco. 2004. "Negative Nominal Interest Rates." *A.E.R. Papers and Proc.* 94 (2): 104–8.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2001. "Monetary Policy and Multiple Equilibria." *A.E.R.* 91:167–86.

———. 2002. "Avoiding Liquidity Traps." *J.P.E.* 110:535–63.

Canzoneri, Matthew, Robert Cumby, and Behzad Diba. 2001. "Is the Price Level Determined by the Needs of Fiscal Solvency?" *A.E.R.* 91:1221–38.

Clarida, Richard, Jordi Galí, and Mark Gertler. 1998. "Monetary Policy Rules in Practice: Some International Evidence." *European Econ. Rev.* 42:1033–67.

———. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Q.J.E.* 115:147–80.

Cochrane, John H. 1998. "A Frictionless Model of U.S. Inflation." *NBER Macroeconomics Annual* 13:323–84.

———. 2005. "Money as Stock." *J. Monetary Econ.* 52:501–28.

Friedman, Milton. 1968. "The Role of Monetary Policy." *A.E.R.* 58:1–17.

Giannoni, Marc P., and Michael Woodford. 2005. "Optimal Inflation Targeting Rules." In *Inflation Targeting*, edited by Ben S. Bernanke and Michael Woodford. Chicago: Univ. Chicago Press. http://www.columbia.edu/126mw2230/(pdf); http: / / www2.gsb.columbia.edu / faculty / mgiannoni / prog / nberit _programs/model_nberit.pdf (tech. app.).

King, Robert G. 2000. "The New IS-LM Model: Language, Logic, and Limits." *Fed. Reserve Bank Richmond Econ. Q.* 86:45–103.

Klein, Paul. 2000. "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model." *J. Econ. Dynamics and Control* 24:1405–23.

Leeper, Eric. 1991. "Equilibria under 'Active' and 'Passive' Monetary and Fiscal Policies." *J. Monetary Econ.* 27:129–47.

Lubik, Thomas A., and Frank Schorfheide. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *A.E.R.* 94:190–217.

McCallum, Bennett T. 1981. "Price Level Determinacy with an Interest Rate Policy Rule and Rational Expectations." *J. Monetary Econ.* 8:319–29.

Minford, Patrick, and Naveen Srinivasan. 2010. "Determinacy in New Keynesian Models: A Role for Money after All?" Manuscript, Cardiff Bus. School.

Obstfeld, Maurice, and Kenneth Rogoff. 1983. "Speculative Hyperinflations in Maximizing Models: Can We Rule Them Out?" *J.P.E.* 91:675–87.

———. 1986. "Ruling Out Divergent Speculative Bubbles." *J. Monetary Econ.* 17: 349–62.

Rotemberg, Julio, and Michael Woodford. 1997. "An Optimization-Based Econometric Model for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual* 12:297–346.

———. 1998. "An Optimization-Based Econometric Model for the Evaluation of Monetary Policy." Technical Working Paper no. 233, NBER, Cambridge, MA.

———. 1999. "Interest-Rate Rules in an Estimated Sticky-Price Model." In *Monetary Policy Rules*, edited by John B. Taylor. Chicago: Univ. Chicago Press.

Sargent, Thomas J., and Neil Wallace. 1975. "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule." *J.P.E.* 83: 241–54.

Schmitt-Grohé, Stephanie, and Martín Uribe. 2000. "Price Level Determinacy and Monetary Policy under a Balanced Budget Requirement." *J. Monetary Econ.* 45:211–46.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica* 48:1–48.

———. 1994. "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy." *Econ. Theory* 4:381–99.

Taylor, John B. 1999. "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank." *J. Monetary Econ.* 43:655–79.

Woodford, Michael. 1995. "Price Level Determinacy without Control of a Monetary Aggregate." *Carnegie-Rochester Conf. Ser. Public Policy* 43:1–46.

———. 2001. "Monetary Policy in the Information Economy." In *Economic Policy for the Information Economy*. Kansas City: Fed. Reserve Bank Kansas City.

———. 2003. *Interest and Prices*. Princeton, NJ: Princeton Univ. Press.