Determinacy and Identification with Taylor Rules

John H. Cochrane*

June 10, 2011

Abstract

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.

*University of Chicago, Booth School of Business, and NBER. 5807 S. Woodlawn Ave. Chicago IL 60637. john.cochrane@chicagobooth.edu. 773 702 3059. I gratefully acknowledge research support from CRSP and the NSF via a grant administered by the NBER. I am grateful for many comments and suggestions in this paper's long gestation. Among many others, I thank 5 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 Inidiana University Monetary Economics Conference, Fall 2006 NBER EFG Conference and University of Chicago Money Workshop. The Online Appendix is available on the JPE website and on my webpage http://faculty.chicagobooth.edu/john.cochrane/research/Papers/.

1 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 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 non-local 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 "non-local" nominal paths. Transversality conditions can rule out *real* explosions, but not nominal explosions. Since the multiple non-local equilibria are valid, the new-Keynesian model does not determine inflation.

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

1.1 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 towards 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, money growth rule, or undertake a fiscal expansion or reform. However, *stopping an inflation* or deflation is a completely different act than *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 which precludes the price level set by the commodity standard. Still others specify inconsistent fiscal and monetary policies, introduce arbitrage opportunities, or set infinite inflation. It is these inconsistent or overdetermined policies, not the inflation or deflation stabilization, which trims equilibria.

Such assumptions seem wildly implausible, as descriptions of government behavior, or 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 which stop the inflation without blowing up the economy, allowing an equilibrium to form at each date. Therefore, "blowing up the economy" is a *choice.* 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 don't think governments *can* set policies for which private first-order conditions don't 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 don't 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 towards the equilibrium price. But there is nothing infeasible or incredible about a commodity standard. Non-Ricardian fiscal commitments work the same way.

1.2 *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 which depend on them, cannot be used as instruments.

Lags don't 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, the use of instruments, assuming the right-hand variables of policy rules are orthogonal to the disturbance, or lag-length restrictions on 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 *in*determinate, 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.

1.3 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.

1.4 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. The Appendix and online Appendix contain a much 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.

2 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 = nominal interest rate, π_t = inflation, r = 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

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

2.1 Determinacy

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

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

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, i.e. $E_t\left(\pi_{t+j}\right)$ 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),

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

Equivalently, by this criterion we select the variables $\pi_0, \{\delta_{t+1}\}$, 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 non-local equilibria? Transversality conditions can rule out real explosions, but not nominal explosions. Hyperinflations are historic realities. This condition didn't come from any economics of the model. I conclude there's nothing wrong with them, and this model does not eliminate multiple equilibria and hence does not determine inflation.

This is an example, which needs fleshing out. First, I need to write down a fully-specified model, to show there truly is nothing wrong with non-local 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 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 2008-2011 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 hyper-inflation and deflation, works just as well as $\phi > 1$ to ensure determinacy. That example is hard to describe by "stabilizing" intuition.

2.2 Identification

Now, suppose the solution (6) is in fact correct, 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)

$$\begin{aligned} E(\pi_t|\pi_{t-1}, i_{t-1}, \pi_{t-2}, i_{t-2}....) &= \rho\pi_{t-1} \\ E(i_t|\pi_{t-1}, i_{t-1}, \pi_{t-2}, i_{t-2}....) &= r + \rho^2\pi_{t-1}. \end{aligned}$$

Thus the instrumental variables regression gives exactly the same estimate

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

If the disturbance x_t were 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 can't 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 $\|\phi\| < 1$. The online Appendix 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, Galí and Gertler (2000) and other authors use, and that the many attempts at identification don't convincingly surmount them.

3 An explicit frictionless model

3.1 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

$$\frac{1}{1+i_t} = E_t\left[Q_{t,t+1}\right].$$

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. S_t denotes the real primary surplus. I follow Benhabib Schmitt-Grohé and Uribe (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

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

or, in real terms,

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

3.2 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,

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

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 using contingent claim prices from (13), equilibrium $C_t = Y$ also requires

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

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). Section 1 of the online Technical Appendix and Cochrane (2005) discuss this issue in more detail. I assume that the present value of future primary surpluses is positive and finite, $0 < \sum_{j=0}^{\infty} \frac{1}{(1+r)^j} E_t \left(S_{t+j} \right) < \infty$.

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.

3.3 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) p. 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 didn't leave anything out.

Are the non-local equilibria really globally valid? Here I follow the standard sources, in part to emphasize agreement that they are. (Woodford (2003) Ch. 2.4, starting p. 123, and Ch. 4.4 starting on p. 311, with a review; Benhabib Schmitt-Grohé and Uribe (2002).)

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

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

$\Phi(\cdot)$ 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 2.

As Benhabib Schmitt-Grohé and Uribe 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 equilibria" 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. 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

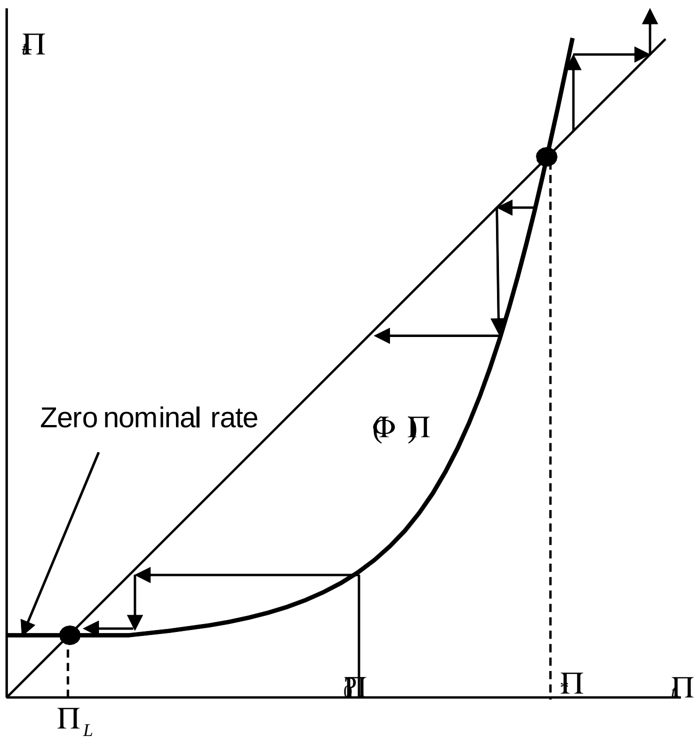

Figure 1: Dynamics in a perfect foresight Taylor-rule model.

debt valuation formula (15), i.e. that there is a set of ex-post lump-sum taxes that can validate them and hence ensure the consumer's transversality condition is satisfied. There are lots of ways the government can implement such a policy. We only need to exhibit 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 2 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 (for example, $i_t = \max(0, r + \phi\pi_t)$), the "liquidity trap" equilibrium $i_t = 0$, $\Pi_L = \beta$ (deflation at the rate r) is also a valid equilibrium.

3.4 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 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 $\frac{1}{1+r}$ and taking expectations,

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

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

$$\frac{(E_{t+1} - E_t) \left(\frac{1}{1+\pi_{t+1}} \right)}{E_t \frac{1}{1+\pi_{t+1}}} = \frac{(E_{t+1} - E_t) \sum_{j=1}^{\infty} \frac{1}{(1+r)^j} S_{t+j}.}{E_t \sum_{j=1}^{\infty} \frac{1}{(1+r)^j} S_{t+j}} \tag{22}$$

Linearizing in the style of Section 2, 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 $\phi > 1$ 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.

Without direct observations of δ_{t+1}, similar identification problems remain, discussed in Section 4.1 of the online Appendix. However, estimates of ϕ are not particularly important in this regime, as price level determinacy or the control of inflation do 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 generate spurious $\phi > 1$ as easily as new-Keynesian models can do.

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.

3.5 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 can't 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.

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're finding the unique *locally bounded* equilibrium, not the unique equilibrium itself. The economy is supposed to jump to the right equilibrium.

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 can't measure ϕ is because 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 expectation. Canzoneri, Cumby and

Diba's (2001) fiscal test has the same flaw as Clarida, Galí and Gertler's (2000) monetary test, shown in Cochrane (1998).

The regimes are also not as 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.

4 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?

4.1 Reasonable expectations?

Much of the approach is simply to think about what expectations seem reasonable. For example, Woodford (p.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, p. 58-59) writes

> "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%, inflation expectations are reasonable, is a path that starts at 1% 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: (p. 128)

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."

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 doesn't 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's a good bet 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 don't 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 hyperinflations 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 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 don't believe the model describes the world in which we live.

4.2 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's (2003) section 4.3 studies proposals to cut off inflationary equilibria to the right of Π^*. Woodford's main suggestion is (p. 138):

> ...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.

In real life governments often stop inflations by a firm peg to a foreign currency (with a fiscal reform, to make credible the 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 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 can't 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 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, $i_T = \pi_{T+1}$.)

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

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

which is a huge number, and inconsistent with $i_T = r$ demanded by first-order conditions.

We would normally say that it's impossible *both* to run a commodity standard which requires $i_T = r$ *and* to set interest rates at hyperinflationary levels which requires 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 can't 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? 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, Chari and Kehoe (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 coexist 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. The Appendix reviews their proposal in detail.

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 models in the Appendix. 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 specifying that the nominal money supply after the reform is no higher than it was before the inflation started, disallowing the recovery in real money balances which 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 (1983) do not specify a commodity standard. They propose that the government repurchase the currency, and then allow an infinite price level forever after. It turns out their proposal does not invalidate the inflation as an equilibrium path. The Appendix 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 return 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 zero.

However, though it may describe other governments and especially the UK 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 governments.

4.3 Fiscal equilibrium trimming

Benhabib, Schmitt-Grohé and Uribe (2002), mirrored in Woodford (2003) section 4.2, try to trim equilibria by adding fiscal commitments to the Taylor rule. Their ideas are aimed at trimming deflationary equilibria, but either set of 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. Ergo, the low-inflation region and all the equilibria below Π^* in Figure 1 that lead to it are ruled out. Specifically, (their equations (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% 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 (p. 132): "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."

As the above proposals are grounded in sensible policies to stabilize hyperinflations, these proposals sound like sensible and time-honored prescriptions to inflate the economy, i.e., to head back to the desired equilibrium Π^*. Benhabib, Schmitt-Grohé and Uribe describe them this way (p. 548):

> 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.

(They didn't 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 undoubtedly what real-world proponents of these policies have in mind.

But that's *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, Schmitt-Grohé and Uribe change tax policy *while also maintaining the Taylor rule* $\Phi(\Pi)$ and the dynamics of Figure 1. In Woodford's p. 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 doesn't 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's not hard to understand why the issue has become so confused. Benhabib, Schmitt-Grohé and Uribe, 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's 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, Schmitt-Grohé and Uribe'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, Schmitt-Grohé and Uribe'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're 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.

4.4 Weird Taylor rules

Woodford starts "Policies to prevent an inflationary panic" by suggesting (p.136) a stronger Taylor rule, that the graph in Figure 1 becomes *vertical* at some finite inflation Π_U above Π^*, i.e. 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 only touch the 45° line at the Π^* point. Bassetto (2004), mentioned above, likewise 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. The more unpleasant the beliefs, the more effective at ruling out equilibria, so hyperinflating away the entire monetary system ($\Phi(\Pi)$ becoming vertical), introducing an arbitrage opportunity (allowing $i < 0$ in the policy rule), and so forth certainly remove these equilibria, and perhaps more effectively than an inflation or fiscal imbalance that slowly gain steam.

However, it's not clear that all these proposals rule out equilibria. A currency can be completely inflated away in finite time. Obstfeld and Rogoff's (1983) model has this property (see the Appendix) and Zimbabwe experienced it. The unpleasant is not impossible.

And, 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.

4.5 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 the Fed directly blows up the economy with infinite interest rates, by adding a little money. This idea is also reviewed by Woodford (2003 p. 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 the Taylor rule is

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

Substituting i_t from the Taylor rule into (24), and rearranging the money vs. bonds first order condition as $M_t/P_t = L(Y, i_t)$, inflation dynamics follow

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

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 importantly, if we do not regard a belief that the Fed will directly blow up the economy (i_t rises to ∞) 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 commodity standard, money growth, or non-Ricardian regime. Second, they require belief in a deep-seated monetary non-neutrality 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 sounds again like a weak foundation for what should be a simple question, the basic determination of the price level.

5 Determinacy and identification in the three-equation model

One may well object at 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 statements, involve at least the Phillips curve and Fed control of real rates of interest: nominal rates rise, gaps appear, these gaps drive down inflation. You can't 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 only show that 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 non-identification and determinacy theorems clear. The basic model is

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

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

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

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, the point of the entire literature is that 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 vs. 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^* etc. to denote equilibrium values. y_t^* is a stochastic process, i.e. a moving average representation $y_t^*(\{x_{dt}, x_{\pi t}, ...\})$ 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, $\tilde{y}_t \equiv y_t - y_t^*$. Subtracting, deviations must follow the same model as (26)-(28), but without constants or disturbances

$$\begin{aligned} \tilde{\imath}_t &= \tilde{r}_t + E_t \tilde{\pi}_{t+1} && (29) \\ \tilde{y}_t &= E_t \tilde{y}_{t+1} - \sigma \tilde{r}_t && (30) \\ \tilde{\pi}_t &= \beta E_t \tilde{\pi}_{t+1} + \gamma \tilde{y}_t. && (31) \end{aligned}$$

5.1 Determinacy

Now, if the Fed sets $i_t = i_t^*$, i.e. $\tilde{\imath}_t = 0$, 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 $\tilde{\imath}_t = 0$ 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 only restricts 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 at least $\tilde{y} = \tilde{\pi} = 0$ 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,

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

(Both King and the online Appendix 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

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

The eigenvalues of this transition matrix are

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

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 $\tilde{\imath} = \tilde{y} = \tilde{\pi} = 0$ is explosive. Ruling out such explosions, we now have the unique locally-bounded equilibrium. (The online Appendix 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), p. 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 2. 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 a 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 non-local equilibria. I do not pursue this question here, as I find no claim in any new-Keynesian writing that this route can rule out the non-local 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 don't, 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 only useful approximations 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 2. 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 1, 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 a 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, noone 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: 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 which 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 observations are essentially observations that *equilibrium* interest rates i_t^* should, or do, vary more than one for one with 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 non-explosive, 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 $(\pi - \pi^*)$.

5.2 Identification

King's expression of the Taylor rule (34) makes the central identification point very 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 $\tilde{\pi} = \tilde{y} = \tilde{\imath} = 0$, 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, in footnote 41:

> "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."

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. The online Appendix makes this point with equations.

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

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

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

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

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's 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^*, π_t^*, are functions of shocks, $i_t^*(\{x_{dt}, x_{\pi t}, x_{it}, ...\})$, "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 failure of Section 2. The equilibrium there is, in response-to-shock form,

$$\begin{aligned} \pi_t^* &= kx_t \\ i_t^* &= r + \rho k x_t \end{aligned}$$

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^*,

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

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 2. 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 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 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 all the shocks follow $AR(1)$ processes $x_{jt} = \rho_j x_{jt-1} + \varepsilon_{jt}$. Then, substituting $y_t^* = \bar{y}_t$ 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 $\phi_y^* = -\left(1-\rho_{\bar{y}}\right)/\sigma$. 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. x_{dt} thus 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 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,

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

The term in brackets is the new monetary policy disturbance. The right hand variable is now hopelessly correlated with the error term. (The online Appendix shows the same result directly and more generally: assuming a Taylor rule without shocks, you can't 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.

5.3 Stochastic Intercept

The term in brackets in (42) ir i_t^* 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, Galí and Gertler (2000) and Woodford (2003, Ch. 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 which 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 2. 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 which 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, don't 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

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

after all, can be rearranged to read

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

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.

5.4 Clarida, Galí and Gertler

Clarida Galí and Gertler (2000) specify an empirical policy rule in partial adjustment form (in my notation),

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

where

$$\begin{aligned} \pi &= \text{inflation target, estimated} \\ \Delta y_{t+1} - \Delta\bar{y}_{t+1} &= \text{growth in output gap} \\ r &= \text{``long run equilibrium real rate'', estimated} \end{aligned}$$

(See their (4) p. 153 and Table II p. 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, Galí, and Gertler 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, Galí, and Gertler replace expected inflation and output with their ex-post realized values, writing

$$i_t = (1-\rho_1-\rho_2)\left\{r + (\phi_\pi - 1)\left[\pi_{t+1} - \pi\right] + \phi_y \Delta y_{t+1}\right\} + \rho_1 i_{t-1} + \rho_2 i_{t-2} + \varepsilon_{t+1}. \quad (44)$$

In this way, Clarida, Galí, and Gertler avoid the 100% 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, Galí and Gertler 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 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 equation 2.5, my notation)

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

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. v_t is now the monetary policy disturbance, defined as "an exogenous random shock to the interest rate." 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 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 preference shifts

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

(IS curve) or marginal cost shocks (Phillips curve shifts) should be i.i.d. But *some* other shock must not 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.

5.5 Giannoni, Rotemberg, Woodford

Rotemberg and Woodford (1997, 1998, 1999), followed by Giannoni and Woodford (2005) (see also the summary in Woodford 2003, Ch. 5) follow a different identification strategy, which allows them to estimate the parameters of the Taylor rule by OLS rather than IV regressions. Giannoni and Woodford (2005 p. 36-37) 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
>
> $$i_t = \bar{\imath} + \sum_{k=1}^{n} \phi_{ik}(i_{t-k} - \bar{\imath}) + \sum_{k=0}^{n_w} \phi_{wk}\hat{w}_{t-k} + \sum_{k=0}^{n_\pi} \phi_{\pi k}(\pi_{t-k} - \bar{\pi}) + \sum_{k=0}^{n_y} \phi_{yk}\hat{Y}_{t-k} + \varepsilon_t \quad (46)$$
>
> 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, $\bar{\imath}$ 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...(p.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 structural disturbances.

Second, they assume that the disturbance ε_t is also not correlated with *contemporaneous* values of $\hat{w}_t, \pi_t$ and $\hat{Y}_t$. This is an especially surprising result of a new-Keynesian model, because $\hat{w}_t, \pi_t$ $\hat{Y}_t$ 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 $\hat{w}_t, \pi_t$ 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 $\hat{w}_t, \pi_t$ $\hat{Y}_t$ must be predetermined by at least one quarter, so they cannot move when ε_t moves. (In the model as described in their Technical Appendix, output $\hat{Y}$ is actually fixed two quarters in advance, and the marginal utility of consumption μ_t is also fixed one quarter in advance.) It is admirable that Giannoni, Rotemberg and Woodford explain the properties of the *model* which 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 one to two quarters in advance in our economy, and therefore unable to react within the quarter to monetary policy disturbances? They certainly aren't forecastable one to two quarters in advance!

Most of all, if $\hat{w}_t, \pi_t$ $\hat{Y}_t$ do not jump when there is a monetary policy disturbance, *something else must jump*, to head off the explosive equilibria. What jumps 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, those reactions are assumed to be absent here.

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; 3) Wages, prices, and output are fixed a period in advance.

6 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 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, so can't 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 a "old-Keynesian" model (my terminology) which 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" (p. 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 (p. 663) that "it is crucial to have the interest rate response coefficient on the inflation rate ... above a critical 'stability threshold' of one," (p. 664)

> 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 that ϕ_π is greater than one.

This is exactly the *opposite* philosophy from the new-Keynesian models. In new-Keynesian models, $\phi_\pi > 1$ is the condition for a "dynamically *un*stable" 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. "$\phi_\pi > 1$" sounds superficially similar, but in fact its operation is diametrically the opposite.

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

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

(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,

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

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 $E_t\left(\pi_{t+1}\right)$. 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 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. To avoid multiple equilibria, we have to change Fed behavior to induce unstable dynamics, and then solve forward to remove these expectational errors as shocks to the economy. The difference isn't 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," so that parameters γ, σ, r may change if ϕ changes. He claims that nonetheless "these equations summarize more complex forward-looking models" (p. 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 the issue in Taylor's model. His model always has just one equilibrium. The issue is "spirals," whether that equilibrium is stable. King (2000) p. 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 doesn't mean 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 VAR literature since then.

Since it easily delivers a unique equilibrium, and thus inflation determinacy, 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 do to 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.

7 Extensions and responses

The online Appendix 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 mode to include an "IS" shock in (1). In this case, we can't even estimate ρ.

I address the question, what happens if you run Taylor rule regressions in artificial data from new-Keynesian models? This discussion generalizes the finding in section 2, 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 Galí and Gertler (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, the online Appendix 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, i.e. following better the "natural rate." This change in policy causes mis-measured 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 where 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 in the online Appendix. 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, e.g. $E_t \pi_{t+1}$ 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.

A separate online Technical Appendix collects documentation and calculation details. I explain budget constraints from Section 3.1 in more depth. I also collect analytic solutions to the standard three-equation model.

8 Conclusions and implications

8.1 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 "non-local" 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 the latter goal is not achieved.

8.2 Identification

The central empirical success of New Keynesian models are estimates such as Clarida, Galí and Gertler's (2000) that say inflation was stabilized in the U.S. 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, Galí, and Gertler's coefficients of interest rates on inflation range from 2.15 (Table IV) to as much as 3.13 (Table V). These coefficients are a *lot* greater than one. These coefficients imply that if the US returned to the 12% inflation of the late 1970s, (a 10 percentage point rise), the Federal Reserve would raise the funds rate by 21.5 to 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 which lead to explosive dynamics, and then pick only the locally-bounded equilibrium. But locally-bounded equilibrium variables are stationary, so cannot reveal the strength of the explosions, which only occur 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 persist in mis-interpreting the regressions.

8.3 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, 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 etc. – 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.

9 Appendix

This appendix reviews the parallel question of inflations with constant money supply and interest-elastic demand, and in particular Obstfeld and Rogoff (1983). I verify that standard proposals suffer 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, p. 138) and Atkeson Chari and Kehoe (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's worth examining both the general idea and their specific example.

9.1 Simple example – Cagan (1956) dynamics

I use the simplest possible example. (Minford and Srinavasan (2010) is a recent paper that uses this framework.) Suppose 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

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

or, rearranging,

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

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

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

with $E_t \left(\delta_{t+1} \right) = 0$ is possible. If $p_t > m$, then we expect a hyperinflation, and conversely. 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 an hyperinflation that emerges under a money target, 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 $p_{T+1} = p_{T+2} = ..\bar{p}$ 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 $m_0 = m_1 = m_2 = m_{T-1} = m$, $m_T = m_{T+1} = ... = \bar{p} = m + \gamma^T > m$. 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's 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 which allows the real money stock to increase is available.

As with the Taylor rule, however, a commitment ot return to a fixed price *level,* with needed fiscal backing, could rule out the hyperinflation.

9.2 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 2 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.

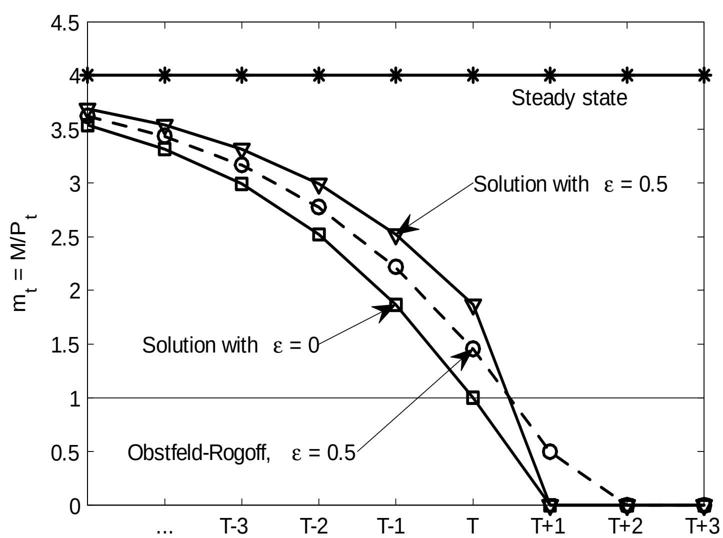

Figure 2: Hyperinfnations in the Obstfeld-Rogoff model. "Solutions with $\varepsilon = 0$" gives the hyperinflation we wish to rule out. "Obstfeld-Rogoff" gives Obstfeld and Rogoff's path when the government offers to redeem the currency for ε units of consumption good. "Solution with $\varepsilon =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

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 their economy needs a period $P_{T+1} = \bar{\bar{\bar{P}}} = 1/\varepsilon$ during which money is repurchased before going on to $P_{T+2} = \infty$ and thereafter. Figure 2 shows this path as well, marked "Obstfeld-Rogoff, $\varepsilon = 0.5$." (I use a rather large ε so the paths are distinguishable on the graph.) They claim, however, that *no matter how small* ε, first-order conditions are violated in this 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 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 2.

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 it in at the end of the period. This fact makes money slightly more valuable at the beginning of the period. Previous periods $T - 1$, etc., 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 people were holding money at T which 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,

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

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

$$\begin{aligned} \frac{u'(c_t)}{P_t} &= \frac{v'(M_t/P_t)}{P_t} + \beta \frac{u'(c_{t+1})}{P_{t+1}} \\ u'(c_t) &= \beta(1+r)u'(c_{t+1}). \end{aligned}$$

(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 equation (14))

$$\frac{u'(y)}{P_t} - \frac{v'(M_t/P_t)}{P_t} = \beta \frac{u'(y)}{P_{t+1}}. \tag{51}$$

Obstfeld and Rogoff study money targets; they assume that "..the money supply is constant at level M." (p. 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 (51). These sequences also satisfy the transversality condition, discussed in Obstfeld and Rogoff's Section II. 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 Inada 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

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

(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.)

With the Inada condition, then, Obstfeld and Rogoff study a special kind of hyperinflation, in which the price level increases steadily following the difference equation (51), 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 after. (Top of p. 681, their Figure 2 and shown as "solution with $\varepsilon = 0$" in Figure 2).

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." (p. 684.) They assume $\overline{\overline{P}} \equiv 1/\varepsilon > \bar{P}$, so that a price level $\overline{\overline{P}}$ is inconsistent with the first-order condition (51) and the money target, $u'(y) - v'(M/\overline{\overline{P}}) < 0$.

Here's their central claim that with this extra provision, hyperinflationary equilibrium paths are ruled out (p. 685):

> Suppose that $\{P_t\}$ is an equilibrium path with $P_0 > \overline{P}$. Let $P_T = \max\left\{P_t | P_t < \overline{\overline{P}}\right\}$. By (14) [my (51)] 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 $\overline{\overline{P}}$. But there is no $M_{T+1} \leq M$ such that $u'(y) - v'(M_{T+1}/\overline{\overline{P}}) \geq 0$. Thus there is no price level P_{T+2} satisfying (14) and $\{P_t\}$ is not an equilibrium path.

Everything follows (51) backwards from a final period in which that the government buys up all the money at $\overline{\overline{P}}$, and after which $P = \infty$. *During that final period*, the first order condition (51) cannot be not satisfied, because $\overline{\overline{P}} > \bar{P}$. 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 (51) is wrong. It does not apply* when people redeem money for a real commodity. It assumes 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 non-negative 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 $\overline{\overline{P}} = 1/\varepsilon$, and governed by the first-order condition (51).

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 redeems money, i.e. that money is redeemed by the government at the end of the period. Equivalently, we can specify an intra-day 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 non-negativity constraints are satisfied, the first order condition is the same as (51):

$$\frac{u'(c_t)}{P_t} = v'\left(\frac{M_t}{P_t}\right)\frac{1}{P_t} + \beta\frac{u'(c_{t+1})}{P_{t+1}}. \tag{52}$$

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 $\varepsilon = 1/\overline{\overline{P}}$ consumption goods, his first order condition is

$$\frac{u'(c_t)}{P_t} = v'\left(\frac{M_t}{P_t}\right)\frac{1}{P_t} + \frac{u'(c_t)}{\overline{\overline{P}}}. \tag{53}$$

Condition (52) holds if $\beta u'(c_{t+1})/P_{t+1} > u'(c_t)/\overline{\overline{P}}$ in which case the consumer sells nothing to the government. Condition (53) holds if $\beta u'(c_{t+1})/P_{t+1} < u'(c_t)/\overline{\overline{P}}$ 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 $P_{T+1} = \overline{\overline{P}}$, $P_{T+2} = \infty$ is not an equilibrium. By (53), $P_{T+1} = \overline{\overline{P}}$ 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 $P_{T+1} = \overline{\overline{P}}$ and $v'(M/\overline{\overline{P}}) > u'(y) > 0$ cannot happen, consistent with their claim.

However, it is *not* true that we must observe $P_T = \overline{\overline{P}} = 1/\varepsilon$ 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. Fundamentally it is the "arbitrage condition" (p. 685) that is wrong. 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 which 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 buy-back guarantee in place. $P_{T+1} = \infty$. Knowing this, at T, people redeem all their money at the end of the period T, so (53) is the relevant first order condition. Rearranging (53) in equilibrium ($c = y$, $M_t = M$),

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

This condition determines P_T. If $\varepsilon = 0$ so $\overline{\overline{P}} = \infty$, then $P_T = \bar{P}$, $v'(m_{T+1}) = u'(y)$ and this is the equilibrium with no buyback. A small ε means a large $\overline{\overline{P}}$, so $P_T < \bar{P}$. Periods prior to T follow the usual difference equation (51). This path *is* an equilibrium, and *is not* ruled out by the repurchase offer.

The central problem is Obstfeld and Rogoff's "arbitrage condition" (p. 685) argument that $\overline{\overline{P}} = P_T$ 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

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

the first-order condition corresponding to (51) is

$$\frac{v'(M_t/P_t)}{u'(y)} = \delta + \frac{1}{P_t}\frac{dP_t}{dt}. \tag{54}$$

Now, $v'(m)$ can rise to arbitrarily large values with a differentiable price path. $\overline{\overline{P}}$ is a valid equilibrium price. The inflationary price path described by (54), terminated by a tender when $P_T = \overline{\overline{P}}$, is a valid equilibrium.

10 References

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

Atkeson, Andrew, Chari, Varadarajan V. and Patrick J. Kehoe, 2010, "Sophisticated Monetary Policies," *Quarterly Journal of Economics* 125, 47–89.

Bassetto, Marco, 2004, "Negative Nominal Interest Rates," *American Economic Review* 94 (2), 104-108.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe, 2002, "Avoiding Liquidity Traps," *Journal of Political Economy* 110, 535-563

Canzoneri, Matthew, Robert Cumby and Behzad Diba, 2001, "Is the Price Level Determined by the Needs of Fiscal Solvency," *American Economic Review* 91, 1221 - 1238.

Clarida, Richard, Jordi Galí, and Mark Gertler, 1998, "Monetary Policy Rules in Practice: Some International Evidence,"*European Economic Review*, 42,1033-1067.

Clarida, Richard, Jordi Galí, and Mark Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147-180.

Cochrane, John H., 1998, "A Frictionless model of U.S. Inflation," in Ben S. Bernanke and Julio J. Rotemberg, Eds., *NBER Macroeconomics Annual 1998* Cambridge MA: MIT press, p. 323-384.

Cochrane, John H., 2005, "Money as Stock," *Journal of Monetary Economics* 52, 501-528.

Giannoni, Marc P., and Michael Woodford, 2005, "Optimal Inflation Targeting Rules," in Bernanke, Benjamin S. and Michael Woodford Eds., *Inflation Targeting*, Chicago: University of Chicago Press, 2005.
(pdf available at http://www.columbia.edu/~mw2230/; Technical Appendix at
http://www2.gsb.columbia.edu/faculty/mgiannoni/prog/nberit_programs/model_nberit.pdf.)

King, Robert G., 2000 "The New IS-LM Model: Language, Logic, and Limits," Federal Reserve Bank of Richmond *Economic Quarterly* 86, 45-103.

Klein, Paul, 2000, "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control* 24, 1405-1423.

Leeper, Eric , 1991, "Equilibria Under 'Active' and 'Passive' Monetary and Fiscal Policies, *Journal of Monetary Economics* 27, 129-147.

Lubik, Thomas A., and Frank Schorfheide,2004, "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review* 94,190-217.

Obstfeld, Maurice and Kenneth Rogoff, 1983, "Speculative Hyperinflations in Maximizing Models: Can we Rule them Out?," *Journal of Political Economy* 91, 675-687.

Obstfeld, Maurice and Kenneth Rogoff, 1986, "Ruling out Divergent Speculative Bubbles," *Journal of Monetary Economics* 17, 349-362.

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

Rotemberg, Julio and Michael Woodford, 1998, "An Optimization-Based Econometric Model for the Evaluation of Monetary Policy," *NBER Technical Working Paper* no. 233.

Rotemberg, Julio and Michael Woodford, 1999, "Interest-Rate Rules in an Estimated Sticky-Price Model," in Taylor, John B., Ed., *Monetary Policy Rules*, Chicago: University of Chicago Press.

Sargent, Thomas J., and Neil Wallace, 1975, "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy* 83, 241-254.

Schmitt-Grohé, Stephanie and Martín Uribe, 2000, "Price Level Determinacy and Monetary Policy Under a Balanced Budget Requirement," *Journal of Monetary Economics* 45, 211-246.

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

Sims, Christopher A., 1994, "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy," *Economic 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," *Journal of Monetary Economics* 43, 655-679.

Woodford, Michael, 1994, "Monetary Policy and Price Level Determinacy in a Cash-in-Advance Economy," *Economic Theory* 4, 345-380.

Woodford, Michael, 1995, "Price Level Determinacy Without Control of a Monetary Aggregate," *Carnegie Rochester Conference Series on Public Policy,* 43, 1-46.

Woodford, Michael, 2001, "Monetary Policy in the Information Economy," in *Economic Policy for the Information Economy*, Kansas City: Federal Reserve Bank of Kansas City.

Woodford, Michael, 2003, *Interest and Prices,* Princeton: Princeton University Press.

Online Appendix to "Determinacy and Identification with Taylor Rules"

John H. Cochrane

University of Chicago Booth School of Business

http://faculty.chicagobooth.edu/john.cochrane/research/Papers/

June 10, 2011

1 Introduction

This Appendix contains additional literature review and many extensions left out of the text for reasons of space.

Sections 2 and 3 critically review more of the literature on identification and determinacy in new-Keynesian and related models. If you want to know "what about x's approach to these issues?," you are likely to find an answer here.

In particular, Subsection 2.1 quickly reviews the learnability criterion advanced by McCallum (2003), (2009). Subsection 2.2 reviews recent proposals by Loisel (2009) and Adão, Correia and Teles (2007) which seem to solve the multiple-equilibrium question. In fact, they propose a limiting case with infinite eigenvalues, a subset of the timing issues reviewed below. Subsection 2.3 reviews Atkeson, Chari, and Kehoe (2010), and shows that they too blow up the world, though only for a day.

Subsection 3 on identification acknowledges the many papers that have made related points critical of identification in new-Keynesian models. It also tackles some attempts to overcome the identification problems. Lubik and Schorfheide (2004) try to identify the region – determinacy vs. indeterminacy – without having to measure specific parameters using likelihood based measures. I show generally that their identification comes from restrictions on the lag length of the unobservable shocks.

Section 3.3 shows that in general one sets to zero movement in eigenvectors corresponding to exploding eigenvalues, so the latter cannot be measured. Thus, the problems are not specific to the structure of the three-equation new-Keynesian model.

Section 3.4 reviews identification in full-system approaches, i.e. approaches that go beyond the single-equation approaches of the text, write down and estimate complete models. Since none of those efforts have been really focused on assessing the determinacy of full models, I do not cover them in the main text. For now, despite the appearance of identification problems, the models impose eigenvalues greater than one and study dynamics.

Section 4 collects some extensions of the frictionless model. Subsection 4.1 explores identification in all the equilibria of the frictionless model. I show that for $\phi < 1$, we still cannot identify ϕ separately from ρ. For $\phi > 1$, and with a prior that $\rho < 1$, we can identify ϕ, for every equilibrium *except* the new-Keynesian choice. If the economy does explode, we can measure the speed of that explosion.

Subsection 4.2 explores the impulse response functions of the simple model. The new-Keynesian model produces responses by supposing endogenous variables jump to a different equilibrium. The section contrasts the response function in the new-Keynesian equilibrium, which combines a (implicit) fiscal shock with a monetary policy shock, to the response function in the active-fiscal equilibrium of the same model, which has a monetary policy shock with no concurrent fiscal shock.

Section 5 addresses the question, what happens if you run Taylor rule regressions in artificial data from new-Keynesian models? As a particular example, if you run the regressions of the first half of Clarida, Galí and Gertler's (2000) paper on data from the model presented in the second half of the paper, do the regressions recover the Taylor rule parameters in the model? The answer is no. The text answered this question in the context of the simple model of section 2: We saw that if we ran the Taylor rule regression in data generated by the very New-Keynesian model, we would recover the shock autocorrelation process, not the Taylor rule parameter. This section answers the same question in the context of three-equation model.

This section also answers the question, "well, if not a change in the Taylor rule, what did Clarida Galí and Gertler measure?" The right answer is really "it's a mongrel coefficient, it doesn't matter." However, this section also gives an example of another shift in policy which could cause the measured Taylor coefficient to rise spuriously. In my example, the Taylor rule coefficient is constant at $\phi = 1.1$, but the Fed gets better at offsetting IS shocks, i.e. following better the "natural rate." This causes measured Taylor rule coefficients to rise as they do in the data. This example also answers a natural generalization of the frictionless model, by adding an "IS shock."

Section 6 addresses 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. This contrasts with the usual feeling that Taylor rules are "robust" to such timing assumptions. The problem is particularly evident on taking the continuous-time limit. This section also addresses a natural question, what if we change the timing conventions from the models of the text?

A technical Appendix follows, which collects some calculations referred to elsewhere.

Section 1 explains budget constraints from section 3.1 in more depth, in particular stressing the difference between the consumer's transversality condition and the government debt valuation equation.

Section 4 gives analytic solutions to the standard three-equation model. These models are often solved numerically, but then it's hard to know how specific parameters enter, especially for identification questions. The solutions are algebraically laborious, so I hope documenting them here is independently useful. In particular, I document analytic expressions for the eigenvalues, allowing analysis of the regions of determinacy and indeterminacy. I also fully solve the models with AR(1) shocks, which allows us to examine response functions and dynamics. In subsection 6.4 I verify the claim made in the text that zero-gap equilibrium, though achievable with a stochastic-intercept interest rate policy that directly offsets shocks, is not achievable via a Taylor rule. This verifies a claim made in the identification discussion of the text.

2 Related literature on determinacy and equilibrium selection

The question, what determines the price level, especially in an economy with fiat money and an interest rate target, naturally has a long history in economics, which I can't begin to review. Patinkin (1949) (1965) brought price-level indeterminacy questions to the fore in postwar macroeconomics. Sargent and Wallace (1975) are the standard proof that interest rate targets lead to indeterminacy. McCallum (1981) was the first to suggest that an interest rate target which varied with economic conditions might overturn Sargent and Wallace's result. I won't even try to cite the literature on fiscal foundations of price level stability. Cochrane (2005) contains one review. Sims (1994, p. 381) encapsulated the basic point well:

> "The existence and uniqueness of the equilibrium price level cannot be determined from knowledge of monetary policy alone; fiscal policy plays an equally important role."

Sims also said "these points are not new."

2.1 Learnablility

In a series of papers, summarized in McCallum (2003), McCallum argues for a "minimal state value" (MSV) criterion to pick from multiple equilibria. In my examples, this criterion obviously rules out the explosive solutions (which depend on initial inflation) or the sunspot solutions. However, this seems a philosophical rather than an economic criterion. Addressing this criticism, McCallum (2003) proposes instead that one choose equilibria by whether they are "learnable" or not, in the sense of Evans and Honkapohja (2001), and argues that one can derive the MSV criterion from this consideration. Woodford (2003b) in commenting on McCallum disagrees, and charges that the "wrong" equilibrium is often the learnable one.

McCallum (2003 p 1154) explicitly states that his proposals do not apply to selecting among nominal indeterminacies, and only apply to models with multiple real paths. Therefore, it appears, he would not apply them to the frictionless models on which I have focused. On the other hand, he does analyze the 3 equation model, presumably because in this case nominal indeterminacy spills over to real variables. Furthermore, his analysis (p. 1160) is confined to the multiple solutions that emerge when the Taylor principle is not satisfied, i.e. $\phi < 1$. He argues that none of these solutions are learnable. To him, then, the point of the Taylor rule is to make the forward-looking solution learnable: "the Taylor principle is of importance because its non-satisfaction leads to a situation in which all RE solutions fail to be learnable." He does not address the multiple explosive solutions that occur when the Taylor principle does apply ($\phi > 1$).

McCallum (2010) is a new and particularly clear example of this line. McCallum argues in the context of the simple model of Section 2 that the explosive equilibria are not "learnable" and the unique bounded equilibrium is the only "learnable" one. This argument now does explicitly apply to nominal indeterminacies. However, I (Cochrane 2010) think he got it backwards. McCallum assumed that the public can directly observe the monetary policy shock x_t; we do not have to run regressions to learn what that shock is. When we make the opposite assumption, that the policy disturbance is not directly observable, so that agents must run regressions to measure it, I obtain the opposite result: the explosive equilibria are "learnable" and the unique local equilibrium is the only one which is *not* "learnable."

This result is closely tied to identification. If an econometrician can't identify ϕ, how is the public supposed to learn it? In these models the public and econometricians have the same information sets. To measure the shock you need to know the slope coefficient, so ϕ and x_t are the same question. On the other hand, as shown in Section 4.1 ϕ *is* identified in the $\phi > 1$, explosive equilibria. When the economy does explode, you can measure how fast it does so. Econometricians can learn ϕ very quickly in this example, and so can agents. Thus, I side with Woodford's (2003b) general comment in this case: the "wrong" equilibria are the "learnable" ones.

More generally, even if true, I think this is a last gasp. Is inflation really determined at a given value because for any other value the Fed threatens to take us to a valid but "unlearnable" equilibrium? Why should we care about such a threat?

2.2 Interest rate rules that seem to work (infinite eigenvalues)

Loisel (2009) proposes a rule that responds to both current and future inflation (simplified to this setting)

$$i_t = r + E_t\pi_{t+1} + \psi\left(\pi_t - z_t\right)$$

where ψ is any nonzero constant, and z_t is any exogenous random variable. If we merge this rule with the usual Fisher equation

$$i_t = r + E_t\pi_{t+1} \tag{55}$$

we obtain a unique equilibrium

$$\pi_t = z_t.$$

This seems to be the Holy Grail: a nominal interest rate rule that delivers a unique equilibrium inflation rate in a frictionless economy. The trick, as Loisel explains, is to have the interest rate rule exactly cancel the troublesome forward-looking terms of the model.

To digest this proposal, write it as a special case of a rule that responds to current and expected future inflation,

$$i_t = r + \phi_0\pi_t + \phi_1 E_t\pi_{t+1} - \phi_0 z_t.$$

Merging this rule with the usual Fisher equation (55), we obtain

$$\begin{aligned} E_t\pi_{t+1} &= \phi_0\pi_t + \phi_1 E_t\pi_{t+1} - \phi_0 z_t \\ E_t\pi_{t+1} &= \frac{\phi_0}{1-\phi_1}\pi_t - \frac{\phi_0}{1-\phi_1}z_t \end{aligned}$$

As usual, this system displays multiple equilibria

$$\pi_{t+1} = \frac{\phi_0}{1-\phi_1}\pi_t - \frac{\phi_0}{1-\phi_1}z_t + \delta_{t+1};\quad E_t\delta_{t+1} = 0. \tag{56}$$

The eigenvalue or root is $\phi_0/(1-\phi_1)$. Thus, if

$$\frac{\phi_0}{1-\phi_1} > 1, \tag{57}$$

We have at least a unique locally-bounded equilibrium,

$$\pi_t = E_t\sum_{j=0}^{\infty}\left(\frac{1-\phi_1}{\phi_0}\right)^j z_{t+j}.$$

Condition (57) is equivalent to

$$\phi_0 + \phi_1 > 1$$

which has the familiar Taylor-rule ring to it, that overall interest rates must rise more than one-for-one with inflation.

Now, we can study not only the *point* $\phi_1 = 1$, but the *limit* $\phi_1 \to 1$. By studying that limit, we see what's going on. As $\phi_1 \to 1$, the eigenvalue or root (57) of the difference equation (56) rises to infinity. For ϕ_1 near 1, the Fed is saying "if inflation doesn't come out to the desired value, we'll hyperinflate *very fast,*" and in the limit multiple equilibria are ruled out by a threat to hyperinflate with *infinite* speed. Thus, this proposal is not really anything new and different, it is an extremely accelerated version of the usual logic. That is not a criticism: much analysis of Taylor rules finds that larger responses are better, and you can't get larger than infinite. The point is just that we can digest this proposal as a limit of the usual logic rather than have to think of it as a fundamentally new type of interest rate target.

We can also see that it is a knife-edge case of the fact, studied below, that determinacy in new-Keynesian models depends sensitively on the timing assumptions. A coefficient even ε different from $\phi_1 = 1.000$, or a time index slightly different from $+1.000$ brings us back to the usual world.

Finally, all forward-looking rules, and this one in particular, require the central bank to respond to private expectations. *In* a rational expectations equilibrium, expectations $E_t(P_{t+1})$ or $E_t\left(\pi_{t+1}\right)$ are just given by $E_t(z_{t+1})$. However, if people develop a sunspot expectation for more inflation, the Fed must increase interest rates to match it. One can easily question the informational and off-equilibrium or game-theoretic foundations of such a response.

Adão, Correia and Teles (2007) advance a similar proposal. Simplified to the linearized, constant real-rate, frictionless environment, they propose the target

$$i_t = r + E_t p_{t+1} - z_t \tag{58}$$

where z_t is any exogenous random variable. If we merge this rule with the usual Fisher equation expressed as

$$i_t = r + E_t\left(p_{t+1} - p_t\right) \tag{59}$$

we obtain a unique equilibrium

$$p_t = z_t.$$

Thus we have an interest rate target that delivers a unique, determinate, price level in a frictionless economy, with no multiple equilibria. Again, the key is that the interest rate rule exactly cancels the forward-looking term of the model, in this case the price level rather than the inflation rate. Adão, Correia and Teles' analysis is in fact conducted in the full nonlinear version of a cash in advance model with labor supply, so linearization, local approximation, and a frictionless economy are not central to the result[2].

[2] Here's the nonlinear version: Start with the consumer's first order condition

$$\frac{u_C(t)}{P_t} = (1+i_t)E_t\left[\frac{\beta u_C(t+1)}{P_{t+1}}\right]$$

Assume a constant endowment $C_t = Y$, so the u_C terms cancel. Then, we can write

$$1+i_t = \frac{1}{P_t E_t\left[\frac{\beta}{P_{t+1}}\right]}$$

Write the policy rule

$$1+i_t = \frac{1}{z_t E_t\left[\frac{\beta}{P_{t+1}}\right]}$$

The globally-unique equilibrium is $P = z_t$.

We can understand this proposal as a similar infinitely-explosive limit of the sort of "Wicksellian" price-level-stabilizing interest rate rules studied by Woodford (2003, p. 81). Generalize the rule to

$$i_t = r + \phi_1 E_t p_{t+1} + \phi_0 p_t - z_t$$

Equate to the Fisher equation (59), and we find the equilibrium condition

$$E_t p_{t+1} = \frac{1+\phi_0}{1-\phi_1} p_t + \frac{1}{1-\phi_1} z_t$$

The eigenvalue is

$$\lambda = \frac{1+\phi_0}{1-\phi_1},$$

and we have a unique locally-bounded equilibrium if $\|\lambda\| > 1$. Woodford studies the case $\phi_1 = 0$, so obtains the condition $\phi_0 > 0$. Adão, Correia and Teles specify $\phi_0 = 0$, and study the limit as $\phi_1 \to 1$. Again, you can see that this is the limit of an infinite eigenvalue, in which the threatened explosion happens infinitely fast.

Loisel's (2009) actual example of an interest-rate rule that seems to avoid multiple equilibria is given in the context of the three-equation model; in my notation

$$i_t = E_t \pi_{t+1} + \phi_{\pi,0} \pi_t + \frac{1}{\sigma} \left(E_t y_{t+1} - y_t \right)$$

If we place this rule in the standard model (26)-(28),

$$\begin{aligned} y_t &= E_t y_{t+1} - \sigma \left(i_t - E_t \pi_{t+1} \right) \\ \pi_t &= \beta E_t \pi_{t+1} + \gamma y_t \end{aligned}$$

we can quickly see that $\pi_t = 0$, $y_t = 0$ is the only equilibrium so long as $\phi_{\pi,0} \neq 0$. In this context, the variables are deviations from a desired equilibrium, so we have shown that the interest rate rule implements the desired equilibrium uniquely.

We can understand this rule as the limit of a standard rule of the form

$$i_t = \phi_{\pi,0} \pi_t + \phi_{\pi,1} E_t \pi_{t+1} + \phi_{y,0} y_t + \phi_{y,1} y_{t+1}.$$

When $\phi_{y,1} = 1/\sigma$, the (single, repeated) eigenvalue of the three-equation model, derived in the Appendix to Cochrane (2007) is

$$\lambda = \frac{1 + \sigma \left(\phi_{y,0} + \gamma \phi_{\pi,0} \right)}{\beta + \sigma \left[\gamma (1 - \phi_{\pi,1}) + \beta \phi_{y,0} \right]}$$

Taking the limit, $\phi_{\pi,1} \to 1$, $\phi_{y,1} = 1/\sigma, \phi_{y,0} = -1/\sigma$, the denominator goes to zero, so again this is a limit with an infinite eigenvalue.

2.3 Atkeson, Chari, and Kehoe; blowing up the world for a day

Atkeson, Chari and Kehoe (2010) eloquently criticize "implementation via nonexistence" or blow-up-the-world threats (p. 50). However, their "sophisticated policies" also rely on policy settings for which first-order conditions do not hold, so no equilibrium is possible. They blow up the world for only one period; *after* that period a competitive equilibrium remains possible.

Here is a simple version of the Atkeson, Chari and Kehoe model which explains these points. (Relative to their model starting on p. 53, I take $\gamma = 0$ and $y =$ constant, or $\psi = 0$.) Specify an endowment economy with constant consumption so the linearized Euler / "IS" equation is

$$i_t = E_t \pi_{t+1}. \tag{60}$$

Add an interest-inelastic money demand function with a money-demand shock v_t, so that money growth μ_t and inflation π_t follow

$$\mu_t = \pi_t + v_t. \tag{61}$$

The central bank can set either i_t or μ_t at the beginning of period t, as functions of time $t-1$ information.

Suppose the central bank follows an interest rate rule

$$i_t = \phi E_{t-1}(\pi_t).$$

(The model has a set of producers who announce prices x_t one period in advance, according to $x_t = E_{t-1}\pi_t$, so the Fed can see $E_{t-1}\pi_t$ directly.) Money growth will then be endogenous, satisfying (61). This is the same model as Section 2 with a money demand function, and an inconsequential change in timing convention. Its equilibria are fully described by one condition

$$E_t \pi_{t+1} = \phi E_{t-1}(\pi_t). \tag{62}$$

The equilibrium $\pi_t = 0$ is possible. But many other equilibria are possible too, indexed by alternative initial values of inflation and by "sunspot" shocks. Any process

$$\pi_t = \phi \pi_{t-1} + \delta_t;\ E_{t-1}(\delta_t) = 0$$

is an equilibrium. The Taylor rule $\|\phi\| > 1$ means that any equilibrium other than $\pi_t = 0$ eventually leads to hyperinflation or deflation, but nothing in the model as it stands rules that out. ("Pure interest rate rules," p. 63.)

By contrast, if the central bank follows a money-targeting rule $\mu_t = 0$, inflation is uniquely determined at $\pi_t = -v_t$. Interest rates are then market-determined at $i_t = E_t \pi_{t+1} = -E_t v_{t+1}$. The money demand shock v_t only serves to motivate why an interest rate rule might be desirable to avoid inflation volatility, so I'll drop it and specify $v_t = 0$ in what follows.

Here is Atkeson, Chari and Kehoe's main idea ("Reversion to a hybrid rule,"p. 65). Like them, I simplify to a perfect-foresight version of the model, though I retain the E_t notation for clarity. The Fed follows the interest rate rule $i_t = \phi E_{t-1}\pi_t$ with $\phi > 1$, so long as expected inflation is below some value $\bar{\pi}$ or above some value $\underline{\pi}$. If expected inflation $E_{t-1}\pi_t$ gets out of this range, the Fed switches to the money growth rule $\mu_t = 0$. The economy then reverts to $\pi_t = 0$, and an equilibrium forms on each date after that point.

Inflations or deflations are therefore stabilized, but how does this provision rule out inflationary paths? Let T denote the date of the switch, so $\mu_t = \pi_t = 0$ for all $t \geq T$, but π_{T-1} is still large. Everyone at time $T-1$ and before knew this would happen, so the first order condition requires $i_{T-1} = E_{T-1}(\pi_T) = 0$. However, this requirement conflicts with the policy rule $i_{T-1} = \phi E_{T-2}(\pi_{T-1})$ which is still large number. Hence, the path can't be an equilibrium. (This is their proof, bottom of p. 66 and top of p. 67.)

For one period, the Fed follows a Taylor rule demanding a high interest rate $i_{T-1} = \phi E_{T-2}(\pi_{T-1})$ *and* a money growth rule $\mu_T = 0$ which demands a low interest rate $i_{T-1} = E_{T-1}(\pi_T) = 0$, at the same time. It's a blow-up-the-world policy, just for a single period. The Fed can't do that in markets, so it's an impossible policy commitment by usual Ramsey rules. And it's a choice. The Fed can stop the inflation, allowing an equilibrium to form at each date, by just waiting one period to switch to a money growth rule, so it is not trying simultaneously to run a money growth rule and Taylor rule.

Chari, Atkeson, and Kehoe's prose is not always clear on this point, but I can find agreement here:

> ...our definition does not require that, when there is a deviation in period t, the entire sequence starting from period 0, including the deviation in period t, constitute a period-zero competitive equilibrium. Indeed, if we achieve unique implementation, then such a sequence will not constitute a period-zero equilibrium (p. 60.)

The only argument really is whether trying to run a policy that requires two different values of the interest rate for one period is as "dire" (p. 50) as a hyperinflation might be.

Minford and Srinavasan (2010) similarly rule out equilibria by having the central bank switch to a money growth rule for large values of inflation or deflation – but at the same time maintain the Taylor rule for interest rates.

3 Related literature on identification

The point that out-of-equilibrium or alternative-equilibrium behavior cannot be measured from data in a given equilibrium is well known, seemingly obvious, once stated, and applies broadly in macroeconomics. Among many others, Sims (1994, p. 384) states as one of four broad principles, "Determinacy of the price level under any policy depends on the public's beliefs about what the policy authority would do under conditions that are never observed in equilibrium." Cochrane (1998) shows analogously that one cannot test the off-equilibrium government behavior that underlies the fiscal theory of the price level: Ricardian and non-Ricardian regimes make observationally equivalent predictions for equilibrium time series without further assumptions. (The models may make very different response-function and policy predictions however.) The point that crucial estimates of many macroeconomic models hinge on "incredible" identification assumptions goes back at least to Sims (1980) My contribution is to apply these well-known principles to new-Keynesian models.

3.1 Literature on lack of identification

The papers closest to this one are Beyer and Farmer (2004, 2006). Beyer and Farmer (2007) compare an "indeterminate" AR(1) model

$$p_t = aE_t(p_{t+1})$$

with $\|a\| < 1$ to a "determinate" AR(2),

$$p_t = aE_t(p_{t+1}) + bp_{t-1} + v_t$$

where they choose a and b so that one root is stable and the other unstable. Both models have $AR(1)$ representations, so there is no way to tell them apart. They conjecture based on this result that Lubik and Schorfheide (2004) attain identification by lag length restrictions.

Beyer and Farmer (2004) compute solutions to the three equation new-Keynesian model. They note (p 24) that the equilibrium dynamics are the same for any value of the Fed's Taylor Rule coefficient on inflation, as long as that coefficient is greater than one. Thus, they see that the Taylor Rule coefficient is not identified by the equilibrium dynamics. They examine the model

$$\begin{aligned} u_t &= E_t u_{t+1} + 0.005\left(i_t - E_t \pi_{t+1}\right) - 0.0015 + v_{1t} \\ \pi_t &= 0.97 E_t \pi_{t+1} - 0.5 u_t + 0.0256 + v_{2t} \\ i_t &= 1.1 E_t \pi_{t+1} + 0.028 + v_{3t} \end{aligned}$$

where v_{it} are i.i.d. shocks. They compute the equilibrium dynamics ("reduced form") as

$$\begin{bmatrix} u_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 0.05 \\ 0.02 \\ 0.05 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0.05 \\ -0.5 & 1 & -0.25 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \tag{63}$$

They state that "all policies of the form

$$i_t = -f_{32} E_t \left[\pi_{t+1}\right] + c_3 + v_{3t},$$

for which

$$|f_{32}| > 1$$

lead to exactly the same reduced form..as long as c_3 and f_{32} are chosen to preserve the same steady state interest rate." They don't state whether this is an analytical result or simply the result of trying a lot of values; since the computation of (63) is numerical, one suspects the latter.

Davig and Leeper (2005) calculate an economy in which the Taylor rule stochastically shifts between "active" $\phi > 1$ and "passive" $\phi < 1$ states. They show that the system can display a unique locally-bounded solution even though one of the regimes is "passive." Intuitively, we can rule out a value of inflation if it will lead to a future explosion after a stochastic shift to a new regime, even if it does not lead to an explosion so long as the current regime is in place. Even if one *could* identify and measure the parameters of the Taylor rule, this model argues against the stylized history that the US moved from "passive" and hence "indeterminate" monetary policy in the 70s to an "active" and hence "determinate" policy in the 1980s. So long as agents understood some chance of moving to an "active" policy, inflation was already "determinate" in the 1970s.

Woodford (2003) notices the identification problem. On p.93, he discusses Taylor's (1999) and Clarida, Galí and Gertler's (2000) regression evidence that the Fed responded less than 1-1 to inflation before 1980 and more than 1-1 afterwards. He writes

> Of course, such an interpretation depends on an assumption that the interest-rate regressions of these authors correctly identify the character of systematic monetary policy during the period. In fact, an estimated reaction function of this kind could easily be misspecified.

An example in which the measured ϕ coefficient is 1/2 of the true value follows. However, though Woodford sees the possibility of a bias in the estimated coefficients, he does not say that the structural parameter ϕ is unidentified.

Minford, Perugini and Srinivasan (2001, 2002) address a related but different identification point: does a Taylor-rule regression of interest rates on output and inflation establish that the Fed is in fact following a Taylor rule? The answer is no: Even if the Fed targets the money stock, equilibrium nominal interest rates, output and inflation will vary, so we will see a "Taylor rule" type relation. As output rises or inflation rises with a fixed money stock, money demand rises, so equilibrium interest rates must rise. As a very simple explicit example, consider a constant money supply equal to money demand,

$$\begin{aligned} m_t^d - p_t &= \alpha y_t - \beta i_t \\ m_t^d &= m^s. \end{aligned}$$

In equilibrium, we see a Taylor-like relation between nominal interest rates, output and the price level

$$i_t = -\frac{1}{\beta} m^s + \frac{\alpha}{\beta} y_t + \frac{1}{\beta} p_t$$

This is an important point: just because the central bank *says* it is following an inflation target, and just because its short run operating *instrument* is obviously an interest rate does not by itself document that the central bank is not paying attention to a monetary aggregate, or that price level determinacy does not in the end really come from such a target.

Mavroeidis (2004, 2005) argues that Clarida, Galí, and Gertler's identification is "weak" because when inflation is well controlled, there is little variation in the right hand variable. I go further to argue that identification is absent, because there is no variation in the crucial right hand variable, which is the deviation of inflation from equilibrium.

3.2 Lubik and Schorfheide; Testing regions

Lubik and Schorfheide (2004) try to identify the region – determinacy vs. indeterminacy – without having to measure specific parameters. Alas, their identification comes from restrictions on the lag length of the unobservable shocks. Beyer and Farmer (2007) make this point with a series of examples. This discussion presents the general case.

Identification is different from approximation. We can be forgiven for running finite-order VARs when theory does not restrict lag lengths. As data increase, lag lengths can increase, and we slowly approach correct estimates. When we use a false restriction to generate identification, there is no sense in which the answer is approximately right or ever gets better as we increase sample size. (Sims 1980 p.5 footnote 5 makes this point eloquently.)

Lubik and Schorfheide explain their ideas in the same single-equation setup as in section 2, simplifying even further by assuming a white noise monetary policy disturbance, i.e. $\rho = 0$. The equilibrium is characterized again by (4) which becomes

$$E_t \pi_{t+1} = \phi \pi_t + \varepsilon_t. \tag{64}$$

The solutions are, generically,

$$\pi_{t+1} = \phi \pi_t + \varepsilon_t + \delta_{t+1}$$

where δ_{t+1} represents the inflation forecast error. If $\phi > 1$, the unique locally-bounded solution is

$$\pi_t = -\frac{\varepsilon_t}{\phi}.$$

If $\phi < 1$, then any δ_{t+1} with $E_t\delta_{t+1} = 0$ gives rise to a locally-bounded equilibrium.

Lubik and Schorfheide agree that ϕ is not identified when $\phi > 1$. For example, the likelihoods in their Figure 1 are flat functions of ϕ for the region $\phi > 1$. However, they still claim to be able to test for determinacy – to distinguish the $\phi > 1$ and $\phi < 1$ regions. The essence of their test is a claim that the model with indeterminacy $\phi < 1$ can produce time-series patterns that the model with determinacy cannot produce.

They explain the result with this simple example. Since δ_{t+1} is arbitrary, it does no harm to restrict $\delta_{t+1} = M\varepsilon_{t+1}$ with M an arbitrary parameter. In this example, then, the (local or bounded) solutions are

$$\begin{aligned} \phi &> 1\text{: } \pi_t = -\frac{\varepsilon_t}{\phi} \\ \phi &< 1\text{: } \pi_t = \phi\pi_{t-1} + \varepsilon_{t-1} + M\varepsilon_t \end{aligned} \tag{65}$$

If $\phi > 1$, the model can only produce white noise inflation π_t. If $\phi < 1$, the model produces an ARMA (1,1) in which ϕ is identified as the AR root. Thus, if you saw an ARMA(1,1), you would know you're in the region of indeterminacy. They go on to construct a likelihood ratio test for determinacy vs. indeterminacy.

Alas, this identification is achieved only by restricting the nature of the shock process x_t. If the shock process x_t is *not* white noise, then the $\phi > 1$ solution can display complex dynamics in general, and an ARMA(1,1) in particular. Since the shock process is unobserved, we cannot in fact tell even the *region* $\phi > 1$ from the region $\phi < 1$. I can sum up this point in a proposition:

Proposition: *For any stationary time-series process for $\{i_t, \pi_t\}$ that represents an equilibrium of (64), and for any ϕ, one can construct an x_t process that generates the same process for the observables $\{i_t, \pi_t\}$ as an equilibrium of (64) using the alternative ϕ. If $\phi > 1$, the observables are generated as the unique bounded forward-looking solution. Given an assumed ϕ and the process $\pi_t = a(L)\varepsilon_t$ we construct $x_t = b(L)\varepsilon_t$ with*

$$b_j = a_{j+1} - \phi a_j, \tag{66}$$

or, in lag operator notation,

$$b(L) = \left(L^{-1} - \phi\right) a(L) - a(0)L^{-1}. \tag{67}$$

In particular, any observed time series process for $\{i_t, \pi_t\}$ that is consistent with a $\phi < 1$ model is also consistent with a different $\tilde{\phi} > 1$ model. Absent restrictions on the unobserved forcing process $\{x_t\}$, there is no way to tell the regime with determinacy from the regime with indeterminacy. Equivalently, the *joint* set of parameters including ϕ and the parameters of the x_t process are unidentified; one can only identify some of these parameters, e. g. $\phi < 1$ vs. $\phi > 1$, by fixing others, e.g., the parameters of x_t.

Proof. Start with any process for inflation $\pi_t = a(L)\varepsilon_t$. Choose an arbitrary $\phi > 1$. Then, we *construct* a disturbance process $x_t = b(L)\varepsilon_t$ so that the forward-looking equilibrium with arbitrary $\phi > 1$ generates the the desired time-series process for inflation, i.e.,

$$\pi_t = a(L)\varepsilon_t = -E_t \sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} x_{t+j} = -E_t \sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} b(L)\varepsilon_{t+j}.$$

It's easy enough to check that (66) is correct:

$$-E_t \sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} b(L)\varepsilon_{t+j} = -E_t \sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} \sum_{k=0}^{\infty} \left(a_{k+1} - \phi a_k\right) \varepsilon_{t+j-k}$$

$$\begin{aligned} = \quad & -\frac{1}{\phi}\left[\left(a_1 - \phi a_0\right)\varepsilon_t + \left(a_2 - \phi a_1\right)\varepsilon_{t-1} + \left(a_3 - \phi a_2\right)\varepsilon_{t-2} + ...\right] \\ & -\frac{1}{\phi^2}\left[\left(a_2 - \phi a_1\right)\varepsilon_t + \left(a_3 - \phi a_2\right)\varepsilon_{t-1} + \left(a_4 - \phi a_5\right)\varepsilon_{t-2} + ...\right] \\ & -\frac{1}{\phi^3}\left[\left(a_3 - \phi a_2\right)\varepsilon_t + \left(a_4 - \phi a_3\right)\varepsilon_{t-1} + \left(a_5 - \phi a_4\right)\varepsilon_{t-2} + ...\right] + ... \end{aligned}$$

$$= a_0\varepsilon_t + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2} + ...$$

If we choose a $\phi < 1$, then the construction is even easier. The solutions to (4) are

$$\pi_{t+1} = \phi\pi_t + x_t + \delta_{t+1},$$

where δ_t is an arbitrary unforecastable shock. To construct an x_t we need therefore

$$\begin{aligned} (1 - \phi L)\pi_{t+1} &= x_t + \delta_{t+1} \\ (1 - \phi L)a(L)\varepsilon_{t+1} &= b(L)\varepsilon_t + \delta_{t+1}. \end{aligned}$$

Obviously, forecast errors must be equated, so we must have $\delta_{t+1} = a_0\varepsilon_{t+1}$. Then,

$$\begin{aligned} (1 - \phi L)a(L)\varepsilon_{t+1} &= b(L)\varepsilon_t + a_0\varepsilon_{t+1} \\ (1 - \phi L)a(L) &= a_0 + Lb(L), \end{aligned}$$

and (67) follows. i_t is just given by $i_t = r + E_t\left(\pi_{t+1}\right)$, and so adds nothing once we match π dynamics.

Example: Suppose we generate data from the Lubik-Schorfheide example with $\phi < 1$, i.e. $x_t = \varepsilon_t$ is i.i.d., and therefore π_t follows the ARMA(1,1) process (65),

$$\pi_t = \phi\pi_{t-1} + M\varepsilon_t + \varepsilon_{t-1} = (1 - \phi L)^{-1}(M + L)\varepsilon_t.$$

We can generate *exactly* the same solution from a model with *arbitrary* $\tilde{\phi} > 1$ if we let the policy disturbance x_t be an ARMA(1,1) rather than restrict it to be white noise. Using (67), we choose $x_t = b(L)\varepsilon_t$ with

$$b(L) = \left(L^{-1} - \tilde{\phi}\right)(1 - \phi L)^{-1}(M + L) - L^{-1}M$$

or, multiplying by $\left(1-\phi L^{-1}\right)$ and simplifying,

$$\begin{aligned}
(1-\phi L)\, x_t &= \left[\left(L^{-1}-\tilde{\phi}\right)(M+L)-(1-\phi L)\, L^{-1}M\right]\varepsilon_t \\
(1-\phi L)\, x_t &= \left[\left(1+\left(\phi-\tilde{\phi}\right)M\right)-\tilde{\phi}L\right]\varepsilon_t \\
x_t-\phi x_{t-1} &= \left[1+\left(\phi-\tilde{\phi}\right)M\right]\varepsilon_t-\tilde{\phi}\varepsilon_{t-1}
\end{aligned}$$

i.e., x_t follows an ARMA (1,1).

The technical appendix shows where formulas (66) and (67) come from.

3.3 General case; system non-identification

One might suspect that these results depend on the details of the three-equation model. What if one specifies a slightly different policy rule, or slightly different IS or Phillips curves? The bottom line is that when you estimate dynamics from stationary variables, you must find stable dynamics. You cannot measure eigenvalues greater than one. In the forward-looking bounded solution, shocks corresponding to eigenvalues greater than one are set to zero.

To study identification, I trace the standard general solution method, as in Blanchard and Kahn (1980), King and Watson (1998), and Klein (2000). The general form of the model can be written

$$\mathbf{y}_{t+1}=\mathbf{A}\mathbf{y}_t+\mathbf{C}\boldsymbol{\varepsilon}_{t+1} \tag{68}$$

where $\mathbf{y}_t$ is a vector of variables, e. g. $\mathbf{y}_t=\left[\begin{array}{ccccc} y_t & \pi_t & i_t & x_{\pi t} & x_{dt}\end{array}\right]'$. By an eigenvalue decomposition[3] of the matrix $\mathbf{A}$, write

$$\mathbf{y}_{t+1}=\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}\mathbf{y}_t+\mathbf{C}\boldsymbol{\varepsilon}_{t+1}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues λ_i, and $\mathbf{Q}$ is the corresponding matrix of eigenvectors.

Premultiplying (68) by $\mathbf{Q}^{-1}$, we can write the model in terms of orthogonalized variables as

$$\mathbf{z}_{t+1}=\boldsymbol{\Lambda}\mathbf{z}_t+\boldsymbol{\xi}_{t+1}$$

where

$$\mathbf{z}_t=\mathbf{Q}^{-1}\mathbf{y}_t;\boldsymbol{\xi}_{t+1}=\mathbf{Q}^{-1}\mathbf{C}\boldsymbol{\varepsilon}_{t+1}.$$

Since $\boldsymbol{\Lambda}$ is diagonal, we can solve for each $\mathbf{z}_t$ variable separately. We solve the unstable roots forwards and the stable roots backwards

$$\|\lambda_i\|>1: \;\; z_{it}=\sum_{j=1}^{\infty}\frac{1}{\lambda_i^j}E_t\xi_{t+j}^i=0 \tag{69}$$

$$\|\lambda_i\|<1: \;\; z_{it}=\sum_{j=0}^{\infty}\lambda_i^j\xi_{t-j}^i \tag{70}$$

$$z_{it}=\lambda_i z_{it-1}+\xi_{it}.$$

[3] King and Watson (1998) and Klein (2000) treat more general cases in which A does not have an eigenvalue decomposition. This generalization usually is just a matter of convenience, for example whether one substitutes in variable definitions or leaves them as extra relations among state variables.

Thus, we choose the unique locally-bounded equilibrium by setting the explosive z_{it} variables and their shocks to zero.

Denote by $\mathbf{z}^*$ the vector of the $\mathbf{z}$ variables corresponding to eigenvalues whose absolute value is less than one in (70), denote by $\boldsymbol{\xi}_t^*$ the corresponding shocks, denote by $\boldsymbol{\Lambda}^*$ the diagonal matrix of eigenvalues less than one in absolute value, and denote by $\mathbf{Q}^*$ the matrix consisting of columns of $\mathbf{Q}$ corresponding to those eigenvalues. Since the other z variables are all zero, we can just drop them, and characterize the dynamics of the $\mathbf{y}_t$ by

$$\begin{aligned} \mathbf{z}_t^* &= \boldsymbol{\Lambda}^* \mathbf{z}_{t-1}^* + \boldsymbol{\xi}_t^* \\ \mathbf{y}_t &= \mathbf{Q}_t^* \mathbf{z}_t^* \end{aligned}$$

The roots $\|\lambda\|$ that are greater than one do not appear anywhere in these dynamics. Thus we obtain general statements of the identification lessons that applied to ϕ in the simple example: 1) *We cannot measure eigenvalues greater than one from the equilibrium dynamics of this model.* Equation (69) shows why: 2) *There is no variation in the linear combinations of variables you need to measure* $\|\lambda\| > 1$. For this reason, 3) *The equilibrium dynamics are the same for every value of the eigenvalues supposed to be greater than one.* The latter statement includes values of those eigenvalues that are less than one. The equilibrium with $\boldsymbol{\Lambda}$ greater than one and no shocks by the new-Keynesian equilibrium selection criterion is observationally equivalent to the same no-shock equilibrium with $\boldsymbol{\Lambda}$ less than one.

This solution gives rise to more variables $\mathbf{y}$ than there are shocks, so it is stochastically singular. We have

$$\mathbf{z}_t = \mathbf{Q}^{-1} \mathbf{y}_t.$$

Since some $\mathbf{z}_t$ are zero, this relationship describes linear combinations of $\mathbf{y}$ that are always zero. However, not all elements of $\mathbf{y}$ are directly observable. The "stochastic singularity" then links endogenous observables (y, π, i) to disturbances (x_π, x_d). Similarly, the expectational errors in $\boldsymbol{\xi}_{t+1} = \mathbf{Q}^{-1}\mathbf{C}\boldsymbol{\varepsilon}_{t+1}$ jump to offset any real shocks so that $\xi_{t+1}^i = 0$ for $\|\lambda_i\| \geq 1$ at all dates.

New-Keynesian models are engineered to have "just enough" forward looking roots. In new-Keynesian models, some of the shocks are arbitrary forecast errors, because some of the structural equations involve expectations; the model stops at $E_t y_{it+1}$ = something else. In this case the backwards solution leads to indeterminacy since forecast errors can be anything. Hence, in new-Keynesian models specify that some of the roots are explosive (forward-looking) so that the forecast errors are uniquely determined and there is a unique local solution.

The only possibility to rescue identification in this context is if there are cross-equation restrictions; if we can learn the λ from the parts of the dynamics we can see. Nobody has traced this idea to see if the parameters of the Taylor rule, or more generally the parameters that control determinacy in more complex models, can be identified.

3.4 System identification

Identifying parameters by estimating the whole *system* is a promising possibility, especially if one feels uncomfortable at the strong assumptions that need to be made for single-equation methods. We write down a complete model, we find dynamics of the observable variables, and we figure out if there are or are not multiple structural parameters corresponding to each possible set of equilibrium dynamics. (Model fit can be measured by distance between impulse-response functions or by the likelihood function.) For determinacy questions, one can then address whether the whole model produces eigenvalues in the zone of determinacy, which is not just a function of Taylor-rule parameters in complex models. (Of course, full systems include specifications of the stochastic process of shocks, so one must be careful that identification does not come crucially from lag-length restrictions.)

There is now a quickly-evolving literature on estimating fully-specified new-Keynesian models and parallel investigation of identification in those estimates. Examples include Rotemberg and Woodford (1997, 1999, and especially 1998 with a focus on identification) and Giannoni and Woodford (2005); Ireland (2007), Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2003). Ontaski and Williams (2004), Iskrev (2010), and Canova and Sala (2009), Fève, Matheron, and Poilly (2007) question the identification in these estimates.

The overall conclusion is that many parameters of typical large-scale models are poorly identified – likelihood functions and other objectives are flat – so some parameters must be fixed ex-ante, or by Bayesian priors so strong that large parameter regions are excluded. Likelihood functions and other objectives often have local minima raising the global identification issue.

For example, the difference between prior and posterior is a measure of how much the data have to say about a parameter. Tellingly, in Smets and Wouter's (2003) estimates, the prior and posterior for the inflation response of monetary policy ϕ_π are nearly identical (Figure 1C p. 1147), and the estimate is 1.68 relative to a prior mean of 1.70, suggesting that the policy rule parameters are at best weakly identified, even in a local sense.

Ontaski and Williams (2004) find that changing priors affects Smets and Wouter's structural parameter estimates substantially. They also find numerous local minima. They report "although our parameter estimates differ greatly, the implied time series of the output gap that we find nearly matches that in SW and the qualitative features of many of the impulse responses are similar." Canova and Sala (2009) are a similar numerical evaluation of identification in large-scale new-Keynesian models with similar results.

Ireland (2007) shows analytically that several parameters in a large scale model are not identified. He also shows that is basically impossible to distinguish econometrically between two versions of the model that provide very different interpretations of postwar US monetary history.

None of these papers even asks whether there are equivalent parameters from the region of non-determinacy which account equally for the observed dynamics. They simply rule out parameters from the non-determinacy region ex-ante or by strong Bayesian priors. The popular DYNARE computer programs will not allow you to compute a solution from this region. In sum, nobody has tried to exploit full-system identification to surmount the difficulties posed above.

This is not a criticism. The authors of these papers are not interested in testing for determinacy. None of them address Clarida, Galí, and Gertler's (2000) question, whether the Federal reserve moved from a 'indeterminate" regime in before 1980 to a "determinate" one after that. More broadly, they are not interested in *testing* the new-Keynesian model, asking whether some other model might equally account for the data. They are interested in matching dynamics of output, inflation, and other variables, by elaboration of the basic model, *imposing* determinacy where there is any question, and making arbitrary choices of parameters when those are weakly identified. Lack of identification, as expressed by Ontaski and Williams (2004), is almost a feature not a bug, as it means the model's ability to match dynamics is "robust" to parameter choices, though all recognize that policy analysis depends on poorly identified parameters.

Perhaps in the future the testing issue will resurface, and then we can evaluate whether the identifying assumptions are reasonable.

3.5 Identification via impulse-response functions

We can get a more concrete sense of these issues by looking at the impulse response functions of a fully specified model. Below, I find the full solution of the model

$$\begin{aligned} y_t &= E_t y_{t+1} - \sigma r_t + x_{dt} \\ i_t &= r_t + E_t \pi_{t+1} \\ \pi_t &= \beta E_t \pi_{t+1} + \gamma y_t + x_{\pi t} \\ i_t &= \phi_{\pi,0} \pi_t + \phi_{\pi,1} E_t \pi_{t+1} + x_{it} \end{aligned}$$

when each of the disturbances x follows an AR(1). That solution is

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1-\rho_d\beta & \sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) & -\sigma\left(1-\rho_i\beta\right) \\ \gamma & \left(1-\rho_\pi\right) & -\sigma\gamma \\ \gamma\left(\phi_{\pi,0}+\rho_d\phi_{\pi,1}\right) & \left(1-\rho_\pi\right)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right) & \left(1-\rho_i\right)\left(1-\rho_i\beta\right)-\sigma\gamma\rho_i \end{bmatrix} \begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix} \tag{71}$$

where the z variables are scaled versions of the disturbances;

$$\begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix} = \begin{bmatrix} \rho_d & 0 & 0 \\ 0 & \rho_\pi & 0 \\ 0 & 0 & \rho_i \end{bmatrix} \begin{bmatrix} z_{dt-1} \\ z_{\pi t-1} \\ z_{it-1} \end{bmatrix} + \begin{bmatrix} v_{dt} \\ v_{\pi t} \\ v_{it} \end{bmatrix}$$

$$\begin{aligned} x_{dt} &= \left[\left(1-\rho_d\right)\left(1-\rho_d\beta\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right)\right] z_{dt} \\ x_{\pi t} &= \left[\left(1-\rho_\pi\right)\left(1-\rho_\pi\beta\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_\pi\left(\phi_{\pi,1}-1\right)\right)\right] z_{\pi t} \\ x_{it} &= \left[\left(1-\rho_i\right)\left(1-\rho_i\beta\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right)\right] z_{it} \end{aligned}$$

These dynamics give us the impulse response function to shocks. Looking at the right-most column of (71), we see again that the Taylor-rule coefficients ϕ do not appear in the response to the monetary policy shock z_{it}. The ϕ coefficients *do* appear in the responses to the other shocks z_{dt} and $z_{\pi t}$ however, which suggests a possibility to identify these parameters. In essence, responses to *other* shocks allow you to see some movement in inflation (and output) and the Fed's response to that movement, without any intervening monetary policy shock. They offer the promise to gain the same identification of the 100% R^2 models without a monetary policy shock, but without making that unpalatable assumption. Equivalently, the other shocks (or combinations of endogenous variables that only depend on those shocks) seem to offer instruments.

Alas, this approach hinges on assumptions about the orthogonality of shocks. To identify a movement in z_{dt} (say) with no movement in z_{it}, we need to make assumptions about the correlation structure of the v shocks. As in the "stochastic intercept" discussion, we really don't have any such information.

4 Frictionless model extensions

4.1 Identification in all equilibria of the simple model

The model of section 2 has multiple equilibria. I studied identification for the New-Keynesian equilibrium choice. Here, I study identification in the other equilibria, i.e. other choices of δ_{t+1}. Perhaps other equilibria do allow identification? Perhaps when $\phi < 1$ we can identify ϕ?

The answer is, other equilibria allow us to identify the *pair* ρ, ϕ, but we can't identify ϕ separately from ρ. This result holds for all choices of δ_{t+1}, and for any value of ϕ, so long as only $\{\pi_t, i_t\}$ are observable. The new-Keynesian equilibrium is a special case in that only ρ is identified.

The equilibrium conditions are

$$
\begin{aligned}
E_t \pi_{t+1} &= \phi \pi_t + x_t && (72)\\
x_t &= \rho x_{t-1} + \varepsilon_t && (73)\\
i_t &= \phi \pi_t + x_t && (74)
\end{aligned}
$$

The equilibrium process for observables is

$$
\begin{aligned}
(i_{t+1} - \rho \pi_{t+1}) &= \phi (i_t - \rho \pi_t) + \delta_{t+1} && (75)\\
(i_{t+1} - \phi \pi_{t+1}) &= \rho (i_t - \phi \pi_t) + \varepsilon_{t+1} && (76)
\end{aligned}
$$

(This is a different δ_t than in the text, though δ still indexes equilibria.) The new-Keynesian equilibrium is the case $i_t = \rho \pi_t$ and $\delta_t = 0$. Except for this special case that the first equation is always zero, ρ and ϕ appear symmetrically in the equilibrium conditions. Therefore, we can identify the pair ρ, ϕ, but we cannot identify which is ρ and which is ϕ.

If $\phi > 1$, then we can in fact identify ρ and ϕ from the prior that $\phi > 1$ and $\rho < 1$. If the system *does* explode, then we can measure the speed of that explosion. If $\phi < 1$ then we will see both roots <1 and we cannot distinguish ϕ from ρ.

(Equation (76) follows by substituting the Taylor rule (74) into the AR(1) process for x_t, (73). To derive (75) write

$$
\begin{aligned}
E_t(i_{t+1} - \rho \pi_{t+1}) &= E_t \left((\phi - \rho) \pi_{t+1} + x_{t+1} \right) = (\phi - \rho)(\phi \pi_t + x_t) + \rho x_t \\
&= \phi (\phi \pi_t + x_t - \rho \pi_t) = \phi (i_t - \rho \pi_t).
\end{aligned}
$$

Then define δ_{t+1} as the unexpected component.)

We can also write (75) and (76) as

$$
i_{t+1} = \rho i_t + \phi (i_t - \rho \pi_t) - \frac{\phi}{\rho - \phi} \delta_{t+1} + \frac{\rho}{\rho - \phi} \varepsilon_{t+1}
$$

$$
\pi_{t+1} = \rho \pi_t + (i_t - \rho \pi_t) - \frac{1}{\rho - \phi} \delta_{t+1} + \frac{1}{\rho - \phi} \varepsilon_{t+1}
$$

Here we see in the new-Keynesian equilibrium choice $i_t = \rho \pi_t$, that only ρ remains identified since the second term becomes zero.

If we write the same dynamics in conventional VAR form,

$$
\begin{aligned}
i_{t+1} &= -\phi\rho\pi_t + (\rho+\phi)i_t - \frac{\phi}{\rho-\phi}\delta_{t+1} + \frac{\rho}{\rho-\phi}\varepsilon_{t+1} \\
\pi_{t+1} &= i_t - \frac{1}{\rho-\phi}\delta_{t+1} + \frac{1}{\rho-\phi}\varepsilon_{t+1}
\end{aligned}
$$

we see that the interest rate equation has all the information in the VAR. If we infer ϕ and ρ from the interest rate equation of the VAR, we obtain a quadratic that treats ρ and ϕ as two roots, and we can't tell which is which.

4.2 Impulse-response functions

It's useful to examine the message of the simple model in the language of impulse-response functions, since we often think about model dynamics and predictions in that framework. Again, the model is

$$
\begin{aligned}
i_t &= r + E_t\pi_{t+1} \\
i_t &= r + \phi\pi_t + x_t \\
x_{t+1} &= \rho x_t + \varepsilon_{t+1}.
\end{aligned}
$$

The equilibrium condition is

$$\pi_{t+1} = \phi\pi_t + x_t + \delta_{t+1}$$

The new-Keynesian solution, Equations (6) and (7), are

$$\pi_t = -\frac{x_t}{\phi-\rho}; \delta_{t+1} = -\frac{\varepsilon_{t+1}}{\phi-\rho} \tag{77}$$

give us the impulse-response function of inflation to a monetary policy shock. Figure 3 plots the response of x_t, i_t and π_t to the monetary policy shock.

Suppose there is a positive monetary policy shock in the Taylor rule $i_t = r + \phi\pi_t + x_t$. Equation (77) predicts that inflation jumps down immediately and then slowly recovers as the shock dynamics $x_{t+1} = \rho x_t + \varepsilon_{t+1}$ play out. This seems reasonable at first glance – monetary tightening lowers inflation. On second glance, this response seems completely counterintuitive in the context of this model. Real rates are constant, so the standard old-Keynesian intuition – higher nominal rates mean higher real rates, higher real rates lower demand, demand leads to less inflation – cannot possibly apply. The only way the Fed can possibly raise nominal rates in this model is to raise expected inflation. How can raising expected inflation lead to a sudden and persistent decline in actual inflation? In addition, notice that *actual* interest rates decline also throughout the episode. To an observer, inflation and interest rates both spontaneously move downward. The sense of a "tightening" is only that interest rates moved down less than ϕ times the downward movement in inflation.

How can this happen? The answer is, inflation "jumps" ($\delta_1 = -1/(\phi-\rho)$) to a new lower-inflation equilibrium in response to the monetary policy shock. Actual interest rates decline, because at each date the disturbance x_t in the Taylor rule $i_t = r + \phi\pi_t + x_t$ is positive, but inflation π_t has jumped down by so much that actual interest rates are lower. An observer would never see a rise in interest rates in this tightening. The observer would only see a decline in interest rates, coincident with the large decline in inflation. The "tightening" comes because interest rates don't decline as

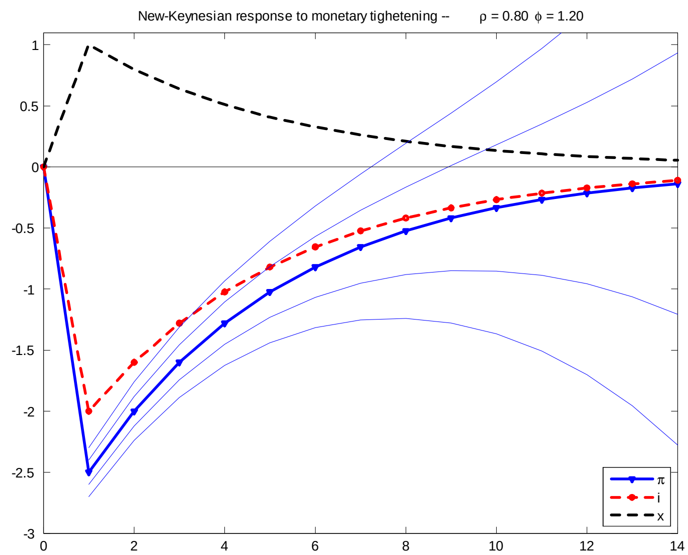

Figure 3: Response of the simple new-Keynesian model to a monetary policy shock. The model is $i_t = r + E_t \pi_{t+1}$; $i_t = \phi \pi_t + x_{t+1}$; $x_{t+1} = \rho x_t + \varepsilon_{t+1}$; the equilibrium is $\pi_t = -\frac{x_t}{\phi - \rho}$.Thin lines show explosive paths if δ_1 does not jump to the right value.

much as inflation, so if the observer knew $\phi = 1.5$, he could infer a positive shock.

There is no economic force, no supply greater than demand, that forces the δ_1 jump in inflation in response to a shock. Any other value of δ_1 and π_1 would correspond to an equilibrium. But all those other equilibria are explosive, as shown. Inflation could even not change at all, $\pi_1 = 0$. Then we would see an increase in interest rates i_t – a standard "tightening – followed by higher subsequent inflation, which is what you might have expected from a nominal interest rate rise in a frictionless economy. The response functions consist of jumps from one equilibrium to another, following the rule that we select locally bounded equilibria. That's how the model can achieve apparent magic – a positive interest rate shock lowers inflation in a completely frictionless model. It doesn't really "lower inflation," it "provokes the economy to jump to a different one of many equilibria, which has lower inflation.

Now, the jump $\delta_1 = -1/(\phi - \rho)$ means that the Treasury will raise taxes or lower spending just enough to change the present value of future surpluses as required by (22). From a fiscal point of view, we can regard (21) as fundamentally determining the price level; but with the Ricardian agreement of a passive fiscal policy, the whole point of the Taylor rule is just to induce the Treasury to embark on the contractionary *fiscal* policy which generates the required δ_1. From a fiscal point of view, the new-Keynesian response combines two shocks, x_t, the monetary policy shock, *and* δ_t, a shock to the present value of future surpluses.

The alternative "non-Ricardian" view suggests we calculate a different response function – what happens if there is a monetary policy shock x_t, but *no* fiscal response, so $\delta_t = 0$? Figure 4 presents a calculation. While we can pair "active" fiscal policy with $\phi > 1$, doing so leads to explosive solutions, so I change parameters to $\phi = 0.8$ in this example. Now, the policy shock $x_t = 1$ produces a 1% rise in interest rates and no change in inflation. Expected inflation and nominal rates are perfectly under the Fed's control with no fiscal response needed, so Taylor rule dynamics $i_t = r + \phi \pi_t + x_t$ now kick in and both interest rates and inflation take a long hump-shaped excursion. The monetary policy shock produces an *increase* in inflation, but that is what one might expect in a frictionless model.

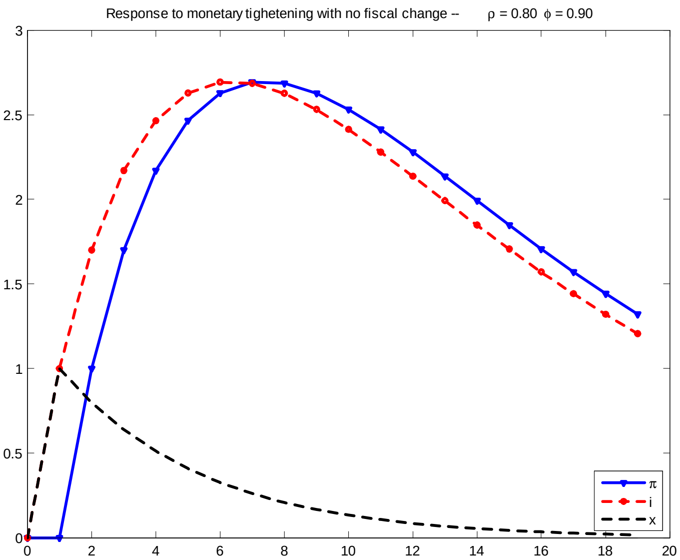

Figure 4: Impluse-response function of the simple model to a monetary policy shock, with no contemporaenous change in present value of surpluses, so $\delta_{t+1} = 0$.

Now, the new-Keynesian view illustrated by Figure 3 and the non-Ricardian view illustrated by Figure 4 are in fact observationally equivalent. In particular, the change in fiscal policy $\delta_1 = -1/(\phi - \rho)$ of figure 3 *could* have happened just by chance along with the change in monetary policy. The correlation between monetary and fiscal shocks could be exactly what the new-Keynesian model suggests they must be. One has to add identification assumptions to try to test the two models, which is beyond the scope of this paper. But the two models certainly suggest different responses to policy interventions! Any impulse response function calculated *with* the assumption that δ_{t+1} will change in a specific way to produce a particular value of inflation at π_{t+1} will be quite different than an impulse-response function calculated assuming no fiscal cooperation, and thus $\delta_{t+1} = 0$. The differences between the two responses in three-equation models are equally stark, but a subject for a different paper.

5 Regressions in model data

In the simple model of section 2, we saw that if we ran the Taylor rule regression in data generated by the very New-Keynesian model, we would recover the shock autocorrelation process, not the Taylor rule parameter. What does happen if you run Taylor rule regressions in artificial data from real New-Keynesian models, such as the standard three-equation model? The answer, is, in general, that the regressions do not recover the structural ϕ parameters. Even if the real data were drawn from the New-Keynesian model, the regressions do not measure its structural parameters. For an example, once again I turn to Clarida, Galí and Gertler (2000), since their paper is so influential and since they include a model in the same paper as an estimate: If you run the regressions of the first half of Clarida Galí and Gertler's paper on artificial data from the model in the second half of their paper, you do not recover the Taylor rule parameters of that paper. (Jensen (2002) also estimates Taylor rules in artificial data from new-Keynesian models, finding estimated coefficients far from the true ones, and often below one.)

5.1 Regressions in real data

As a backdrop, I replicate, update and slightly extend a simple version of Clarida, Galí, and Gertler's Taylor rule regressions in Table 1. I run regressions

$$i_t = a + \rho i_{t-1} + (1-\rho)\left[\phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + \phi_y\left(y_t - \bar{y}_t\right)\right] + \varepsilon_t$$

in quarterly data. I instrument expected inflation $E_t\pi_{t+1}$ on the right hand with current inflation.

Row 1 gives the basic result: a very high 3.66 coefficient on expected inflation. In the earlier period, row 9, this coefficient is only 0.85. Rows 5 and 13 show that the use of expected future inflation is a refinement, and shows a stronger result, but even a simple rule using only current inflation produces a 2.13 coefficient up from 0.78 in the earlier period. These coefficients are quite similar to Clarida, Galí, and Gertler's coefficients which range from a baseline 2.15 (Table IV) to as much as 3.13 (Table V). Coefficients of this magnitude are typical of the literature.

Row 3 and 11 present the raw regression coefficients, i.e. not divided by $(1-\rho)$. Much of the large ϕ_π estimate comes from a rise in the persistence ρ from 0.68 to 0.90, which implies larger long-run multipliers ϕ_π from the same raw coefficients. Rows 7 and 15 remind us that these coefficients require a dynamic model (i_{t-1}) and computation of the long-run multiplier $\phi = b/(1-\rho)$. If you just run a Taylor rule in levels, you recover the fact that long-run levels of inflation and interest rates must move essentially one for one.

Row		i_{t-1}	π_t	$E_t\pi_{t+1}$	$y_t - \bar{y}_t$	R^2	$R^2(i_t - \hat{\rho} i_{t-1})$
				1984:1-2010:1			
1	$\phi = b/(1-\rho)$	0.90		3.66	0.64	0.95	0.23
2	s.e.	(0.05)		(1.69)	(0.35)		
3	b	0.90		0.36	0.06		
4	s.e.	(0.05)		(0.14)	(0.04)		
5	$\phi = b/(1-\rho)$	0.92	2.13		1.00	0.96	0.19
6	s.e.	0.02	(1.09)		(0.33)		
7	b		0.99		0.48	0.36	
8	s.e. (NW 12Q)		(0.35)		(0.08)		
				1960:1-1979:2			
9	$\phi = b/(1-\rho)$	0.68		0.85	0.54	0.89	0.49
10	s.e.	(0.08)		(0.13)	(0.16)		
11	b	0.68		0.27	0.17		
12	s.e.	(0.08)		(0.08)	(0.03)		
13	$\phi = b/(1-\rho)$	0.74	0.78		0.76	0.89	0.43
14	s.e.	0.07	(0.17)		(0.21)		
15	b		0.75		0.28	0.75	
16	s.e. (NW 12Q)		(0.04)		(0.08)		

Table A1. Taylor rule regressions. i is the Federal funds rate, measured in the first month of the quarter. π is the GDP deflator. $y - \bar{y}$ is log GDP less log CBO potential GDP. b give raw regression coefficients. $b/(1-\rho)$ divide the coefficients on π and $y - \bar{y}$

by $1-\rho$ to give estimates of ϕ. "$R^2(i_t - \hat{\rho} i_{t-1})$" gives $1 - \sigma^2(\varepsilon)/\sigma^2(i_t - \hat{\rho} i_{t-1})$. In IV regressions I use π_t as an instrument for $E_t\pi_{t+1}$ and R^2 reports the variance of the instrumented right hand side divided by the variance of the left hand variable. Standard errors by GMM include ρ estimation error and correct for heteroskedasticity. NW12Q also corrects for serial correlation with a 12 quarter Newey-West weight. See section 3 for standard error calculations.

5.2 Regressions in model data

Now, suppose data are generated from a New-Keynesian model. Will regressions such as the above recover the structural coefficients? Alas, even the simple three-equation model is complex enough that the answers, though computable, are algebraically large and hence not that enlightening. Therefore, I report a numerical investigation. Start with the most straightforward three-equation model,

$$\begin{aligned} y_t &= E_t y_{t+1} - \sigma(i_t - E_t\pi_{t+1}) + x_{dt} && (78)\\ \pi_t &= \beta E_t\pi_{t+1} + \gamma y_t + x_{\pi t} \\ i_t &= \phi_{\pi 0}\pi_t + x_{it} \\ x_{jt} &= \rho_j x_{jt-1} + \varepsilon_{jt}; \quad j = d, \pi, i. && (79) \end{aligned}$$

Here I have added shocks to each equation to allow us to avoid perfect correlations between variables[4]. Following Clarida, Galí and Gertler, I use

$$\beta = 0.99; \sigma = 1.0; \gamma = 0.3.$$

I assume all the shocks are independent of each other and have a common 0.9 autoregression coefficient. I solve the model (see the online Appendix for details), simulate a very long series of artificial data, and run Taylor rule regressions in the resulting artificial data. Table 2 collects results.

The rows of Table 2, labeled "model" give the assumed value of policy parameters ρ and ϕ. We want to see if regressions can recover these values.

In row 1, I simulate data with no monetary policy shock, $\sigma_i = 0$. This regression does recover

[4] Clarida, Gali and Gertler specify a slightly different model, their (6)-(9) p. 169 and in my notation

$$\begin{aligned} y_t &= E_t y_{t+1} - \sigma(i_t - E_t\pi_{t+1}) + x_{dt} && (80)\\ \pi_t &= \beta E_t\pi_{t+1} + \gamma\,(y_t - \bar{y}_t) \\ i_t &= \rho i_{t-1} + (1-\rho)\left[\phi_{\pi 1}E_t\pi_{t+1} + \phi_y\,(y_t - \bar{y}_t)\right] \\ x_d &= \rho_d x_{dt-1} + \varepsilon_{dt}; \bar{y}_t = \rho_{\bar{y}}\bar{y}_{t-1} + \varepsilon_{\bar{y}t}. && (81) \end{aligned}$$

Potential output $\bar{y}_t$ varies, there is no additional shock to the phillips curve, and there is no monetary policy shock. I had hoped to calculate "what happens if you run Clardi Gali and Gertler's regressions in data from Clarid Gali and Gertler's model," but unfortunately this is impossible. This model produces a perfect correlation between $\pi_t, E_t\pi_{t+1}$ and $(y_t - \bar{y}_t)$. Hence the right hand variables of the Taylor-rule regression are perfectly correlated; one *cannot* run Clarida, Gali and Gertler's regressions in artificial data from their model. If one adds errors to the Phillips curve or monetary policy equation one can break that correlation, but then it's no longer really their model, and no lessons beyond what I find in the simpler model covered in the text emerge. The point of course is not to deconstruct one

paper, but to understand a whole literature using that paper as an example, so the model in the text which has a better chance of success is appropriate.

the true policy parameter. It must. With no stochastic intercept and no error term, $i_t = \phi_{\pi,0}\pi_t$, if there is any variation in inflation at all we must recover the true $\phi_{\pi 0}$. The trick and cost is easy to see – the R^2 is 100%. Thus, we could easily reject this model in any real dataset.

Row 2 adds back a monetary policy shock with 1% standard deviation, to remove the 100% R^2 prediction. Now, as in the simple model, the right hand variable is correlated with the error and we recover a coefficient of 1.44, not the true coefficient of 2.00. In row 3 we see that if there is only a monetary policy shock (i.e. if the variance of the monetary policy shock is much larger than that of the other shocks), the estimated coefficient declines to 0.86. Row 4 shows what happens if we estimate an interest-rate persistence parameter ρ where there is none in the underlying model. Interestingly we recover a rather large spurious persistence $\rho = 0.42$ estimate along with the inconsistent $\phi_{\pi,0}$ estimate.

The second group of results in Table 2 add an output response and persistence in the Taylor rule. The picture is the same, with larger inconsistencies. The estimated persistence parameter ρ, is lower, 0.65-0.55, than the true value $\rho = 0.90$. In part as a result, the estimated $\phi_{\pi,0}$, are 0.86-0.78, much lower than the true $\phi_{\pi,0} = 2$, and the ϕ_y estimate is destroyed to values near zero. The estimated values are in the zone of multiple local equilibria. If you run Taylor rule regressions in data generated from this model, you have no hope of recovering the true policy function.

Regressions in model data not only do not capture the true coefficients, they also produce far higher R^2 than in actual data, which is a revealing failure. If the model is right, regressions in artificial data should produce all features of the estimated regression, including its R^2. Yet even with all three shocks, R^2 is almost exactly one. The R^2 of a highly serially correlated variable is misleading, since so much explanatory power comes from the lagged dependent variable. Hence, I also compute in Tables 1 and 2 the R^2 of the component of interest rates not predicted by the lagged interest rate, "$R^2(i_t - \rho i_{t-1})$" which is $1 - \sigma^2(\varepsilon)/\sigma^2(i_t - \hat{\rho} i_{t-1})$. These R^2 in data generated by the model are between 0.92 and 1, much higher than the values from actual data of 0.2-0.3, shown in in table 1.

The prediction of very high R^2 in artificial data seems hard to avoid. The reason for the high R^2 is that the right hand variables all jump when there is a monetary policy shock. Thus, even if we have a very large policy shock, the right hand variables respond and incorporate its information. I read this as an indication that right hand variables in the real world are not jumping as predicted by the model.

The estimates in rows 8 and 9 of Table 2 assume and estimate a rule with expected future inflation, using instrumental variables. Now we do not recover the true values even when there are no policy shocks so that $R^2 = 1.00$. Mechanically, the most important instrument for expected inflation is past inflation, with a first-stage coefficient of about 0.5. Thus, the two-stage least squares estimate is roughly double the OLS estimate. In addition, this estimate is much higher than the true value, a counterexample to the impression one might otherwise get that estimated coefficients are always lower than true ones. Once again, adding a monetary policy shock dramatically lowers the estimate, without substantially changing the prediction of a very high R^2.

Finally, rows 10 investigates the stochastic intercept. Here, I assume the policy rule is

$$i_t = \rho i_{t-1} + \phi_\pi \pi_t + \phi_y (y_t - \bar{y}_t)$$

i.e. no error term, and policy responds to the output gap rather than output. In the model, I also write the Phillips curve as a function of the output gap,

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

Now, if we can observe potential output and estimate a policy rule with that value, we can recover the structural parameters, if we swallow the usual 100% R^2. However, the Fed typically has much more information about "potential output" than we do, and its assessment of potential is a prime source of disturbances to estimated monetary policy equations. (Of course, potential output is likely to be correlated to other shocks as well, but I ignore this fact for the calculation.) The Fed's "potential output' guess is a source of a stochastic intercept. Thus, what happens when we estimate a policy rule using only actual output? Row 10 shows the answer, again showing that we recover nothing like the true policy parameters, even without any monetary policy shock.

Row		σ_i	$\sigma_d = \sigma_\pi$	ρ	$\phi_{\pi,0}$	$\phi_{\pi,1}$	ϕ_y	R^2	$R^2(i_t - \hat{\rho} i_{t-1})$
	Model:			0	2		0		
1		0	1		2.00			1.00	
2		1	1		1.44			0.85	
3		1	0		0.86			1.00	
4		1	1	0.42	1.55			0.90	
	Model:			0.90	2		1		
5		1	1	0.65	0.86		0.03	0.99	0.92
6		1	0.5	0.59	0.78		-0.02	0.99	0.97
7		1	0.01	0.55	0.76		-0.03	1.00	1.00
	Model:			0.90		3	1		
8	IV	0	1	0.90		6.5	1.25	1.00	1.00
9	IV	1	1	0.72		1.02	0.01	0.98	0.87
	Model:			0.90	2		1		
10		0	0.5	0.85	1.30		0.28	0.99	0.90

Table A2. Regressions $i_t = \rho i_{t-1} + (1-\rho)\left(\phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + \phi_y y_t\right) + \varepsilon_t$ in long (T = 20,000) artificial data from the three-equation New Keynesian model (78)-(79). $\sigma_i, \sigma_d, \sigma_\pi$ give the standard deviation in percent of shocks. IV estimates use current inflation to instrument for expected inflation. In row 10 the policy rule and Phillips curve respond to $y_t - \bar{y}_t$, but the regression only uses only y_t. "$R^2(i_t - \hat{\rho} i_{t-1})$" gives $1 - \sigma^2(\varepsilon)/\sigma^2(i_t - \hat{\rho} i_{t-1})$.

5.3 Large estimates – A New-Keynesian interpretation

How *do* estimated Taylor rules recover large coefficients, and coefficients that change around 1980? Really, the answer is "we don't care." Once we know the coefficients are mongrels, mixing irrelevant model parameters and shock dynamics, who cares how they turn out? They are not measuring anything important in the context of the New-Keynesian model.

But many of the examples so far have all shown a downward bias. Is the fact that estimated Taylor rules show large coefficients embarrassing? Again, the answer is no. It is easy to give plausible examples in which the estimated Taylor rules give much larger than actual coefficients, and their coefficients change from below to above one depending on small other changes in specification.

For example, introduce an "IS shock" z_t, to the simple model of section 2, so the model is

$$\begin{aligned} i_t &= r + E_t\pi_{t+1} + z_t; \quad z_t = \rho_z z_{t-1} + \varepsilon_{zt}; \\ i_t &= r + \phi\pi_t + x_t; \quad x_t = \rho_x x_{t-1} + \varepsilon_{xt}. \end{aligned}$$

Now the equilibrium is

$$E_t\pi_{t+1} = \phi\pi_t + x_t - z_t$$

and the forward-looking solutions are

$$\begin{aligned} \pi_t &= -\frac{x_t}{\phi - \rho_x} + \frac{z_t}{\phi - \rho_z} \\ i_t &= -\frac{\rho_x x_t}{\phi - \rho_x} + \frac{\phi z_t}{\phi - \rho_z}. \end{aligned}$$

The x shock lowers inflation, so inflation and x are negatively correlated leading to downward bias in a Taylor rule estimate. A z shock raises inflation, however, so if z and x are positively correlated, we can have inflation positively correlated with the monetary disturbance x and a positive bias in the Taylor rule estimate.

This is also a sensible specification. Left alone, the IS shocks z_t would induce inflation variability. A central bank that wants to minimize the variance of inflation would deliberately introduce a policy disturbance $x_t = \frac{\phi-\rho_x}{\phi-\rho z} z_t$ in order to offset the IS shocks. This is an example of a "stochastic intercept," a "Wickesllian policy" in which the monetary policy disturbance offsets other structural shocks.

The bias can be large, and it is largest when the central bank is close to doing its job of offsetting z shocks. For example, let $\rho_x = \rho_z = \rho$, and let z and x be perfectly positively correlated. Then the regression coefficient of i_t on π_t is[5]

$$\hat{\phi} = \frac{cov(i_t, \pi_t)}{var(\pi_t)} = \frac{\rho - \phi\frac{\sigma_z}{\sigma_x}}{1 - \frac{\sigma_z}{\sigma_x}}.$$

A parameter σ_z just slightly higher than σ_x gives large positive coefficients. For example, $\rho = 0.7$, $\phi = 1.1$, and $\sigma_z/\sigma_x = 1.2$ produces an estimated coefficient similar to the estimated values,

$$\hat{\phi} = \frac{(0.7 - 1.1 \times 1.2)}{(1 - 1.2)} = 3.1. \tag{82}$$

On the other hand, if the IS shocks z_t and monetary policy shocks x_t are uncorrelated, the estimated coefficient is instead between ρ and ϕ,

$$\hat{\phi} = \frac{\rho + \phi\frac{\sigma_z^2}{\sigma_x^2}}{1 + \frac{\sigma_z^2}{\sigma_x^2}}.$$

[5] Algebra:

$$\frac{cov(i_t, \pi_t)}{var(\pi_t)} = \frac{cov\left(-\frac{\rho x_t}{\phi-\rho} + \frac{\phi z_t}{\phi-\rho}, -\frac{x_t}{\phi-\rho} + \frac{z_t}{\phi-\rho}\right)}{var\left(-\frac{x_t}{\phi-\rho} + \frac{z_t}{\phi-\rho}\right)} = \frac{\rho\sigma_x^2 + \phi\sigma_z^2 - (\rho+\phi)\,\sigma_x\sigma_z}{\sigma_x^2 + \sigma_z^2 - 2\sigma_x\sigma_z}$$

which simplifies to the given expression.

For example, then, $\rho = 0.7$, $\phi = 1.1$, $\sigma_z/\sigma_x = 1$ gives

$$\hat{\phi} = \frac{0.7 + 1.1}{2} = 0.9. \tag{83}$$

Thus, one possibility consistent with standard estimates is this: A New-Keynesian model with constant Taylor rule $\phi = 1.1$ operated throughout the period. Before 1980, the Fed was not very good at offsetting IS shocks, so we estimate $\hat{\phi} = 0.9$ as in (83). After 1980, the Fed got much better at offsetting IS shocks, and we estimate $\hat{\phi} = 3.1$. Inflation volatility declined because the Fed got better at offsetting IS shocks, not because it changed the Taylor rule parameter.

I do not argue that this is what happened. The point of the determinacy section of this paper is that the $\phi > 1$ passive-fiscal solutions are fundamentally flawed. It is simply one logical possibility, and a way to remind ourselves that once estimated coefficients do not measure structural parameters, changes in those mismeasured coefficients can reflect all sorts of changes in model structure.

Further pursuit of an "explanation for Taylor-Rule results" is just not interesting. Producing spurious $\hat{\phi}$ estimates consistent with the empirical findings is a necessary condition for the right model, but not sufficient. Many wrong models will also produce the observed $\hat{\phi}$ estimates. More importantly, producing this mongrel coefficient is not likely to be much help in discriminating between models focused on other effects. We don't learn much about new models by asking if they produce investment-equation regressions, consumption-function regressions, regressions of output on lags of money stocks or any other now-spurious regression run in the past when thinking about other models is that useful.

6 Rules with leads and lags

It is often claimed that the principle "raise interest rates more than one for one with inflation" is quite robust to details of model and rule specification. (Taylor (1999) and Woodford (2003) among many others.) In fact, determinacy is very sensitive to small changes in timing, whether the central bank reacts to current, lagged, or expected future inflation. This is true both in the simple model and in the standard three equation model. We have already seen a hint of this in Section 2.2; a small change in timing there changed the crucial eigenvalue to infinity. This section also documents a natural question from the text, "what happens if we change the timing?" "Robustness" may refer to model fit, impulse-responses, and policy analysis *given* determinacy. I do not address this issue.

6.1 The frictionless model

Start with our simple Fisher equation model (1),but allow the Fed to respond to expected future inflation rather than current inflation, and for simplicity ignore the disturbance x_t. Determinacy means $\pi_t = 0$. Generalize the simple model to change the timing, i.e.

$$\begin{aligned} i_t &= r + E_t \pi_{t+1} \\ i_t &= r + \phi E_t \pi_{t+j}. \end{aligned}$$

Again, we find equilibria by eliminating i_t between these two equations.

For $j = 0$ (contemporaneous inflation), the equilibrium condition is

$$E_t \pi_{t+1} = \phi \pi_t$$

as we have seen, the condition for a unique local equilibrium is $\|\phi\| > 1$.

For $j = 1$, a reaction to expected future inflation, the equilibrium condition becomes

$$E_t \pi_{t+1} = \phi E_t \pi_{t+1}.$$

If $\phi = 1$, anything is a solution. For any $\phi \neq 1$ (both $\phi > 1$ and $\phi < 1$), solutions must obey

$$E_t \pi_{t+1} = 0; \; \pi_t = \delta_{t+1}. \tag{84}$$

We conclude that inflation must be white noise – real rates are constant. But that's all we can conclude. No value of ϕ gives even local determinacy.

For $j = 2$, we have

$$E_t \pi_{t+1} = \phi E_t \pi_{t+2}$$

Now a necessary condition for "unstable" or "forward-looking" equilibrium is reversed, $\|\phi\| < 1$. Since interest rates react to inflation *two* periods ahead, and interest rates control expected inflation *one* period ahead, the interest rate and *one*-period ahead inflation must move less than *two* period ahead inflation if we want an explosive root. And even this specification is now not enough to give us a unique local equilibrium, since there is an E_t on both sides of the equation. $\pi_{t+1} = \delta_{t+1}$, $E_t(\delta_{t+1}) = 0$ is a solution for *any* value of ϕ.

In sum, in this simple model, Taylor determinacy disappears as soon as the Fed reacts to expected future rather than current inflation, and the solution is extremely sensitive to the timing convention. All the dynamics of the model, which are crucial to the idea of using forward-looking solutions to determine expectational errors, rely entirely on the assumed dynamics by which the Fed reacts to inflation.

$j = -1$ gives really weird dynamics, but preview what will happen in continuous time. Now the equilibrium condition is

$$E_t \pi_{t+1} = \phi \pi_{t-1}$$

Even and odd periods live in their own disconnected equilibria!

6.2 Continuous time and dynamics

The timing issue drives the apparently strange modifications one needs to make in order to take the continuous-time limit. Benhabib, Schmitt-Grohé and Uribe (2001) present one such model. If we eliminate money from their perfect foresight model, the Fisher equation is simply

$$i_t = r + \pi_t.$$

In continuous time (and with continuous sample paths) the distinction between past and expected future inflation vanishes. If we write a Taylor rule

$$i_t = \phi(\pi_t),$$

as they do, we see that this system behaves exactly as the $i_t = r + \phi E_t \pi_{t+1}$ or $j = 1$ case above: If $\phi(\pi) = r + \pi_t$ anything is an equilibrium; otherwise there is a unique equilibrium in perfect foresight, but the same multiplicity once we allow expectational errors. It seems there are no dynamics (or the dynamics happen infinitely quickly), so the forward-looking trick to determine expectational errors disappears.

Benhabib, Schmitt-Grohé and Uribe do have dynamics which look a lot like Figure 1 (see their Figure 1). However, these dynamics come from an entirely different source. Their model has money in the utility function. The dynamics of inflation in their model come from the standard interest-elasticity of money demand, much like Cagan (1956) hyperinflation dynamics under a money target.

Here is their argument in continuous time: With money in the utility function and a constant endowment, the first-order condition for money M_t vs. consumption C_t implies a "money demand" curve (my notation) $M_t/P_t = L(Y, i_t)$. Thus, we can write the marginal utility of consumption in equilibrium as

$$u_c(C_t, M_t/P_t) = u_c\left[Y, L(Y, i_t)\right] = \lambda(i_t) = \lambda\left[\phi(\pi_t)\right]$$

where the last equalities define the function λ. Differentiating, and using the continuous-time first order condition

$$\frac{\dot{u}_c}{u_c} = i - \pi - \rho$$

where ρ is the rate of time preference, we have

$$\dot{\pi}_t = \frac{\lambda\left[\phi(\pi_t)\right]\left[\phi(\pi_t) - \pi_t - \rho\right]}{\lambda'\left[\phi(\pi_t)\right]\phi'(\pi_t)}.$$

This differential equation in π_t turns out to look just like Figure 1.

The idea may be clearer in the discrete-time formulation. Using a Taylor rule $i_t = \phi\left(\Pi_{t+1}\right)$, Equation (25) becomes

$$\Pi_{t+1} = \beta\left[1 + \phi(\Pi_{t+1})\right]\frac{u_c(Y, M_{t+1}/P_{t+1})}{u_c(Y, M_t/P_t)}.$$

If we had no money in the utility function, you can see how once again we are stuck. There are no dynamics. However, the interest-elasticity of money demand offers hope. Substituting for money demand, we have

$$\begin{aligned}
\Pi_{t+1} &= \beta\left[1 + \phi(\Pi_{t+1})\right]\frac{u_c\left[Y, L(Y, i_{t+1})\right]}{u_c\left[Y, L(Y, i_t)\right]}. \\
\Pi_{t+1} &= \beta\left[1 + \phi(\Pi_{t+1})\right]\frac{u_c\left\{Y, L\left[Y, \phi(\Pi_{t+2})\right]\right\}}{u_c\left\{Y, L\left[Y, \phi\left(\Pi_{t+1}\right)\right]\right\}}.
\end{aligned}$$

Now we again have a difference equation that can look like Figure 1.

In sum, despite the superficially similar behavior of Benhabib, Schmitt-Grohé and Uribe's (2002) model to the frictionless models studied above, we see they are fundamentally different. Benhabib, Schmitt-Grohé and Uribe's model cannot work in a frictionless economy; it relies on the dynamics induced by interest-elastic money demand rather than dynamics induced by the Policy rule.

To mirror the sort of dynamics we have seen from $i_t = \phi(\pi_t)$ rules in continuos time, one must specify some explicit time lag between inflation and the interest rate, $i_t = \phi(\pi_{t-k})$, or $i_t = \int_{k=0}^{\infty} f(k)\pi(t-k)dk$. For example, Sims (2003) models a Taylor rule in continuous time in this way, as

$$\frac{di}{dt} = \theta_0 + \theta_1 \frac{1}{P}\frac{dP}{dt} - \theta_2 i. \tag{85}$$

Here, more inflation causes the Fed to raise the *rate of change* of interest rates. Sims also has a Fisher equation

$$i = \rho + \frac{1}{P}\frac{dP}{dt}$$

Sims solves for the interest rate,

$$\frac{di}{dt} = \theta_0 + (\theta_1 - \theta_2)\, i_t - \theta_1 \rho$$

thus wanting $\theta_1 < \theta_2$ for forward looking solutions. The specification (85) isn't an ad-hoc peculiarity, it is exactly the sort of modification we must make for Taylor-rule dynamics to work in continuous time.

6.3 Timing in the three-equation model

One might think that sensitivity to timing is a peculiarity of the frictionless model; that price stickiness will smooth things out in some sense. In particular, the forward-looking rule $i_t = r + \phi E_t \pi_{t+1}$ seems to make much more sense in the three-equation model, because it ensures that real rates $r_t = i_t - E_t \pi_{t+1}$ will rise after an increase in inflation. In the frictionless model, a rise in real rates is impossible, which could account for that model falling apart when the Taylor rule becomes forward looking. However, that could be old-Keynesian intuition sneaking in. The point of the Taylor rule is not to raise real rates, lower demand, and lower inflation; the point of the Taylor rule in a new-Keynesian model is to destabilize the system so that it explodes for all but one initial value. Let's see.

The familiar three-equation model is also sensitive to timing. The Taylor-rule parameter regions required to produce a forward-looking solution vary considerably whether the central bank reacts to current or expected future inflation, and whether the central bank reacts to output. Section 4 of the technical appendix derives the eigenvalues and provides analytical characterization of the determinacy regions. The equations aren't that revealing: since we're studying roots of quadratic equations there are multiple special cases for real roots, imaginary roots, and roots equal to one or to negative one, and one must check the smaller of two roots. Therefore, I focus here on a graphical analysis of some special cases.

Figure 5 presents the simplest case, the region of determinacy for a Taylor rule $i_t = \phi_{\pi,0}\pi_t$ in the standard three equation model.

The conventional determinacy region $\phi_{\pi,0} > 1$ is visible. A region of large negative $\phi_{\pi,0}$ with $\sigma\gamma > 0$ is also visible, corresponding to the $\phi < -1$ region studied in the simple model of Section 2. Finally, eigenvalues are a property of the whole model, not just the interest rate rule. Here, if $\sigma\gamma < 0$, all sorts of interesting ϕ regions lead to determinacy or not, including a region with $\phi_{\pi,0} > 1$ that is nonetheless indeterminate, and one with $\left\|\phi_{\pi,0}\right\| < 1$ that is nonetheless determinate. Of course $\sigma > 0$ an $\gamma > 0$ are conventional parameter restrictions, but this simple example alerts us that in more complex models eigenvalues are likely to depend on all parameters, not just the intuitively appealing (using old-Keynesian intuition!) Taylor rule parameters.

Figure 6 presents the regions of local determinacy in the three-equation model, for a policy rule that responds to expected future inflation $i_t = \phi_{\pi,1} E_t \pi_{t+1}-$ again, perhaps the most interesting

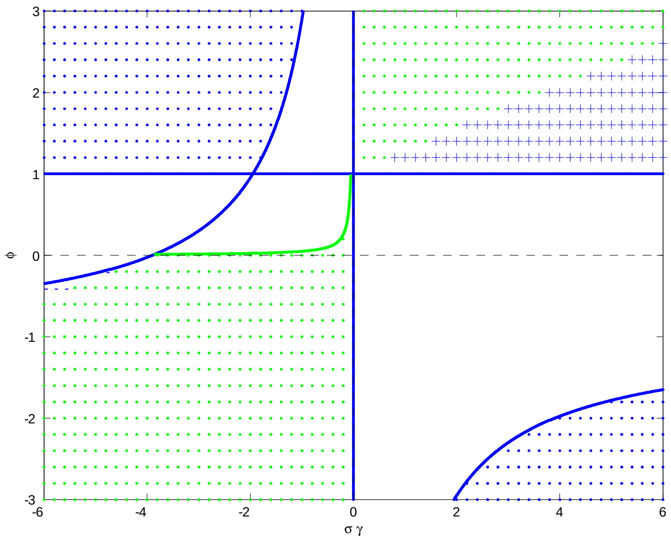

Figure 5: Regions of unique local equilibrium (both roots greater than one) for the three-equation new Keynesian model, in which the Fed follows a Taylor rule $i_t = \phi\pi_t$. $\beta = 0.95$. Blue indicates real roots, green indicates complex roots; +, - and $\cdot$ indicate positive, negative, and mixed roots respectively. Lines denote regions in which one root is equal to one.

case. In the usual parameter region $\sigma\gamma > 0$, we see a comforting region $\phi_{\pi,1} > 1$. The rest of the parameter space is quite different from the case $i_t = \phi_{\pi,0}\pi_t$ of Figure 5, however. In particular, as King (2000, Figure 3b) notices, there is a second region of *large* $\phi_{\pi,1}$ that again leads to local indeterminacy. In this region, both eigenvalues are negative, but one is less than one in absolute value. "Sunspots" that combine movements in output and inflation, essentially offsetting in the Phillips curve so that expected inflation doesn't move much, lead to stable dynamics.

Figure 6 plots regions of local determinacy combining responses to current and future inflation $\phi_{\pi,0}$ and $\phi_{\pi,1}$. We see the expected condition $\phi_{\pi,0} + \phi_{\pi,1} > 1$ in the downward-sloping line of the left hand part of the plot. However, this line disappears when $\phi_{\pi,0} < -(1-\beta)/\sigma\gamma = -0.025$. A greater $\phi_{\pi,1}$ cannot make up for an even slightly negative $\phi_{\pi,0}$. In fact, we see that the sum $\phi_{\pi,0}+\phi_{\pi,1}$ must be *less than one* for local determinacy when $\phi_{\pi,0} < 0$, an excellent counterexample to the view that $\phi_{\pi,0} + \phi_{\pi,1} > 1$ is a robust result. In both cases we see again that $\phi_{\pi,1}$ cannot get too big or again we lose determinacy for any value of $\phi_{\pi,0}$.

King concludes from experiments such as these: "Forward-looking rules, then, suggest a very different pattern of restrictions are necessary to assure that there is a neutral level of output." He also takes seriously the regions with local indeterminacy (one eigenvalue less than one) despite large $\phi_{\pi,1}$ as important policy advice, writing (p.80) "It is important, though, that it [monetary policy] not be too aggressive, since the figure shows that some larger values are also ruled out because these lead to indeterminacies."

As a slightly novel example, consider what happens if the Fed responds to expected inflation two periods ahead as well as one period ahead, i.e. consider a Taylor rule of the form

$$i_t = \phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + \phi_{\pi,2}E_t\pi_{t+2}.$$

Figure 8 presents the region of determinacy (both roots greater than one in absolute value) for this

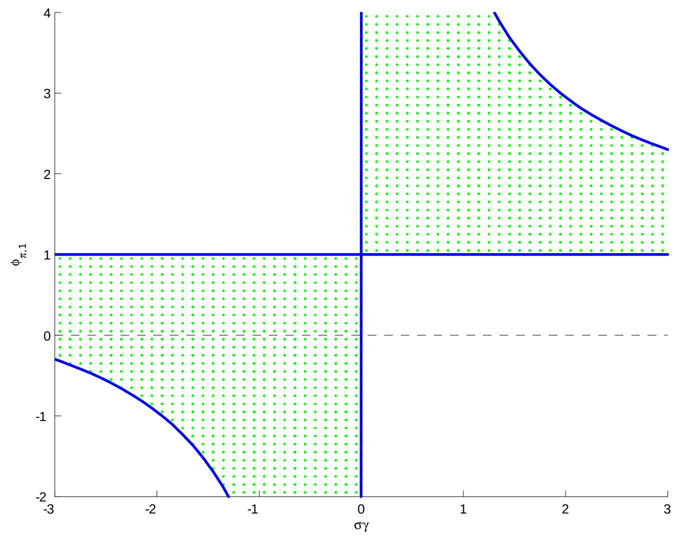

Figure 6: Zones of local determinacy – both eigenvalues greater than one in absolute value – when the policy rule responds to expected future inflation $i_t = \phi_{\pi,1} E_t \pi_{t+1}$. The plotted boundaries are $\phi_{\pi,1} = 1$, $\phi_{\pi,1} = 1 + 2(1+\beta)/(\sigma\gamma)$.

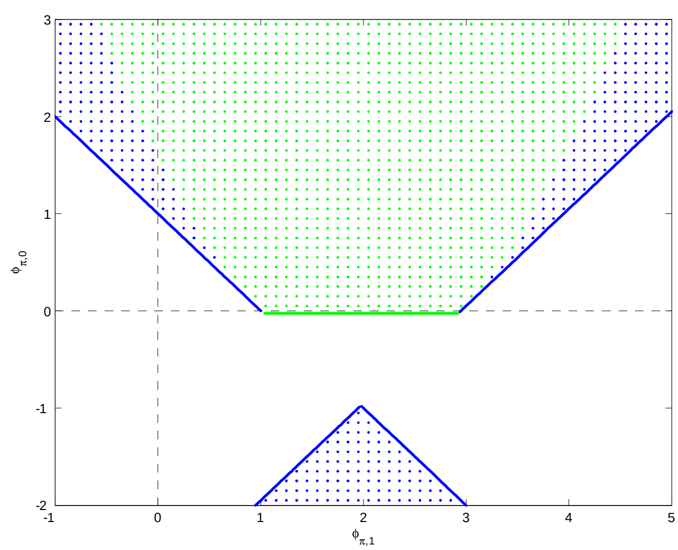

Figure 7: Zones of local determinacy when the policy rule responds to both current and expected future inflation, $i_t = \phi_{\pi,0}\pi_t + \phi_{\pi,1}\pi_{t+1}$, using $\beta = 0.95$ and $\sigma\gamma = 2$. The plotted boundaries are $\phi_{\pi,0} + \phi_{\pi,1} = 1$, $\phi_{\pi,1} - \phi_{\pi,0} = 1 + 2(1+\beta)/\sigma\gamma$ and $\phi_{\pi,0} = -(1-\beta)/\sigma\gamma$.

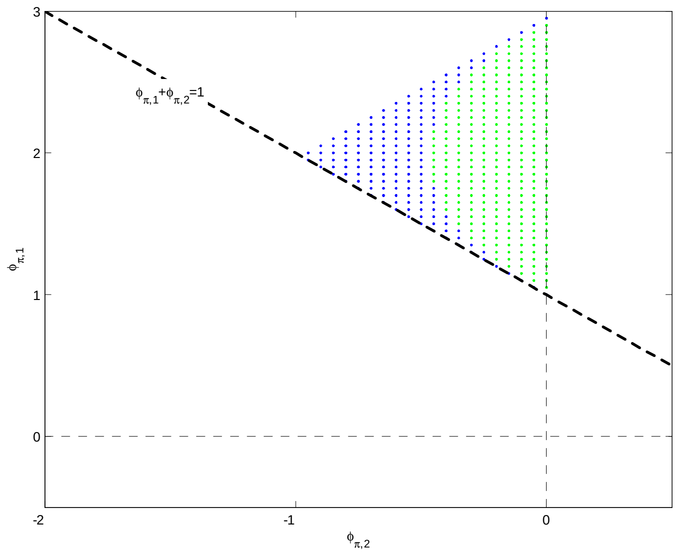

Figure 8: Region of local determinacy when the policy rule responds to expected future inflation one and two periods ahead, $i_t = \phi_1 E_t \pi_{t+1} + \phi_2 E_t \pi_{t+2}$, using $\beta = 0.95, \sigma\gamma = 2$.

case. As you can see, there is some sense in the view that $\phi_{\pi,0} + \phi_{\pi,1} + \phi_{\pi,2} > 1$ is important for determinacy. (As one raises $\phi_{\pi,0}$, the region of local determinacy descends as you would expect.) However, there's more to it than that. We *must* have $\phi_{\pi,2} \leq 0$ – the Fed must respond *negatively* if at all to expected inflation two periods out! Furthermore, we see another example in which too large $\phi_{\pi,1}$ leads to indeterminacy for any $\phi_{\pi,2}$.

Allowing responses to output adds a whole interesting new range of possibilities. Figure 9 presents the determinacy region for rules limited to current output and inflation, $i_t = \phi_{\pi,0}\pi_t + \phi_{y,0} y_t$, for $\beta = 0.95, \sigma = 2, \gamma = 1$. Figure 10 presents the same region for rules limited to expected future output and inflation, $i_t = \phi_{\pi,0} E_t \pi_{t+1} + \phi_{y,0} E_t y_{t+1}$.

In Figure 9, you can see that output responses can substitute for inflation responses. Rules are possible that use *only* output responses, ignoring inflation all together. Depending on ϕ_y and the other parameters of the model, almost any value of ϕ_π can be consistent or inconsistent with determinacy.

Figure 10 shows that the region of determinacy using expected *future* output and inflation is radically different than that which uses *current* output and inflation. In particular, almost the whole range $\phi_y < 0$ is wiped out, and there are severe constraints on how strong the inflation and output responses can be.

Two-dimensional graphs can only do so much justice to this 7-dimensional space ($\phi_{\pi,0}$, $\phi_{\pi,1}$, $\phi_{y,0}$, $\phi_{y,1}$, β, γ, σ), of course. The determinacy $\|\lambda\| = 1$ boundaries in this case are as follows. (These conditions are derived below and included in the plots. To turn them into boundaries, one has to also check that the other eigenvalue is greater than one.)

1. $\sigma\phi_{y,1} \neq 1$, real roots, $\lambda = 1$

$$\left(\phi_{\pi,0} + \phi_{\pi,1} - 1\right) + \frac{1-\beta}{\gamma}\left(\phi_{y,1} + \phi_{y,0}\right) = 0 \qquad (86)$$

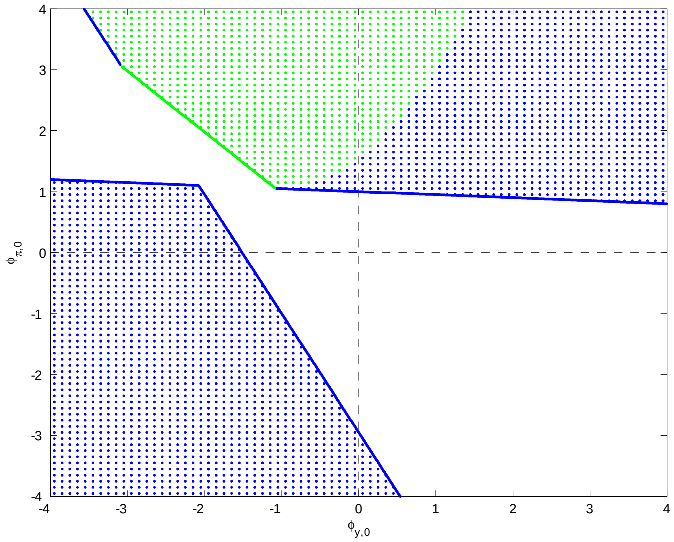

Figure 9: Region of determinacy in the three-equation model for a rule $i_t = \phi_{\pi,0}\pi_t + \phi_{y,0}y_t$, using $\beta = 0.95, \sigma = 2, \gamma = 1$.

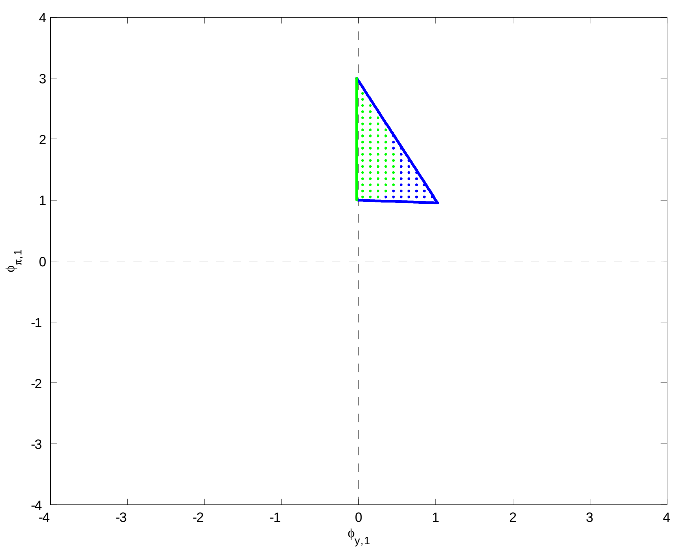

Figure 10: Region of determinacy for the three-equation model, for an interest rate rule $i_t = \phi_{\pi,1}E_t\pi_{t+1} + \phi_{y,1}E_t y_{t+1}$, using $\beta = 0.95, \sigma = 2, \gamma = 1$.

2. $\sigma\phi_{y,1} \neq 1$, real roots, $\lambda = -1$

$$\left(1 + \phi_{\pi,0} - \phi_{\pi,1}\right) - \frac{1+\beta}{\gamma}\left(\phi_{y,1} - \phi_{y,0}\right) = -2\frac{(1+\beta)}{\sigma\gamma}$$

3. $\sigma\phi_{y,1} \neq 1$, Complex roots

$$\gamma\phi_{\pi,0} + \phi_{y,0} + \beta\phi_{y,1} = \frac{\beta - 1}{\sigma}.$$

4. $\sigma\phi_{y,1} = 1$, $\lambda = 1$,

$$\left(\phi_{\pi,0} + \phi_{\pi,1}\right) + \frac{(1-\beta)}{\sigma\gamma}\left(1 + \sigma\phi_{y,0}\right) = 1$$

5. $\sigma\phi_{y,1} = 1$, $\lambda = -1$:

$$\phi_{\pi,0} - \phi_{\pi,1} + \frac{(1+\beta)}{\sigma\gamma}\left(1 + \sigma\phi_{y,0}\right) = -1 \tag{87}$$

Equations (86)-(87) show a variety of interesting additional interactions. Four of the five conditions do not involve $\phi_{\pi,0} + \phi_{\pi,1}$ and $\phi_{y,0} + \phi_{y,1}$ – we see analytically that responding to current vs. future output and inflation are not the same thing. In fact the second and last conditions includes $\phi_{\pi,0} - \phi_{\pi,1}$: The two responses have opposite effects. (Unsurprisingly, these are conditions in which the eigenvalue is equal to negative one.) The other conditions trade off responses with coefficients that depend on structural parameters.

The point of all these examples is to emphasize that the Taylor rule does not "stabilize inflation" in new-Keynesian models; rather it threatens explosive equilibria. The examples emphasize that the regions of determinacy depend on the entire system, not just the policy rule. The regions are sensitive functions of policy rule parameters and timing, as well as economic parameters. The "robustness" may be a feature of old-Keynesian models and stories, but not of these models.

6.4 Zero-inflation and zero-gap equilibria

With the solution (71) in hand we can easily check that the $\pi_t = 0$ and $y_t = 0$ (there is no $\bar{y}$ here) are not achievable, so to achieve those equilibria a "stochastic intercept" policy is necessary. The Fed could set x_{it} and hence z_{it} to zero, eliminating this source of variance. Given that the variance of the z_t is not zero, there is nothing to do about any of the loadings of y_t or π_t except the loading of y_t on $z_{\pi t}$. (Top, center.) That loading can be set to zero with $\phi_{\pi,0} = \rho_\pi\left(1 - \phi_{\pi,1}\right)$. But the other loadings remain. Hence, the only hope is to set to zero the variances of the z_t, by altering the coefficients that relate z_t to x_t. To obtain this result, we need to explode the denominators,

$$\begin{aligned} \phi_{\pi,0} + \left(\phi_{\pi,1} - 1\right)\rho_d &\rightarrow \infty \\ \phi_{\pi,0} + \left(\phi_{\pi,1} - 1\right)\rho_\pi &\rightarrow \infty \end{aligned}$$

The only way to do this is to send the individual ϕ_π responses to ∞. Dong so sets $\pi_t = 0$. The loading of y on z_π offsets this operation, though, so even setting $\phi_\pi = \infty$ does not give $\sigma(y) = 0$.

7 References

Atkeson, Andrew, Chari, Varadarajan V. and Patrick J. Kehoe, 2010, "Sophisticated Monetary Policies," *Quarterly Journal of Economics* 125, 47–89.

Adão, Bernardino, Isabel Correia and Pedro Teles, 2007, "Unique Monetary Equilibria with Interest Rate Rules," Manuscript, Banco de Portugal, Universidade Catolica Portuguesa and CEPR.

Benhabib, Jess; Stephanie Schmitt-Grohé, Martín Uribe, 2001, "Monetary Policy and Multiple Equilibria," *American Economic Review* 91, 167–86.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe, 2002, "Avoiding Liquidity Traps," *Journal of Political Economy* 110, 535-563

Beyer, Andreas, and Roger Farmer 2004, "On the Indeterminacy of new-Keynesian Economics," European Central Bank Working Paper 323.

Beyer, Andreas, and Roger Farmer 2007, "Testing for Indeterminacy: An Application to US Monetary Policy: A Comment," *American Economic Review,* 97, 524-529.

Blanchard, Olivier J., and Charles M. Kahn, 1980, "The Solution of Linear Difference Models Under Rational Expectations," *Econometrica* 48, 1305-1310.

Cagan, Phillip, 1956, "The Monetary Dynamics of Hyperinflation," in *Studies in the Quantity Theory of Money*, Milton Friedman, ed., Chicago: University of Chicago Press, 25- 117.

Canova, Fabio, and Luca Sala, 2009, "Back to Square One: Identification Issues in DSGE Models", Journal of Monetary Economics, 56(4), 2009, 431-449.

Christiano, Lawrence., Eichenbaum, Martin and Charles Evans, 2005, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy,* 113, 1-45.

Cochrane, John H., 2009, Can Learnability Save New-Keynesian Models? *Journal of Monetary Economics* 56 1109–1113.

Evans, George W. and Seppo Honkapohja, 2001, *Learning and Expectations in Macroeconomics*, Princeton: Princeton University Press.

Fève, Patrick, Julien Matheron, and Céline Poilly, 2007, "Monetary Policy Dynamics in the Euro Area," *Economics Letters* 96 97–102.

Iskrev, Nikolay, 2010, "Local identification in DSGE models," *Journal of Monetary Economics* 57,189-202.

Jensen, Henrik, 2002,"Monetary Policy Frameworks and Real Equilibrium Determinacy," Manuscript, University of Copenhagen.

Hansen, Lars Peter, and Thomas J. Sargent, 1980, "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2 (2), 7-46.

Ireland, Peter, 2007, "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," *Journal of Money, Credit and Banking,* 39, 1851-1882.

King, Robert G., and Mark W. Watson, 1998, "The Solution of Singular Linear Difference Systems Under Rational Expectations," *International Economic Review* 39, 1015-1026.

Loisel, Olivier, 2009, "Bubble-Free Policy Feedback Rules." *Journal of Economic Theory* 144: 1521-1559.

Mavroeidis, Sophocles, 2004, "Weak identification of forward-looking models in monetary economics," *Oxford Bulletin of Economics and Statistics*, 66 Supplement, 609-635.

Mavroeidis, Sophocles, 2005, Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve," *Journal of Money Credit and Banking* 37, 421-448.

McCallum, Bennett T., 1981, "Price Level Determinacy With an Interest Rate Policy Rule and Rational Expectations," *Journal of Monetary Economics* 8, 319-329.

McCallum, Bennett T., 2003, "Multiple-Solution Indeterminacies in Monetary Policy Analysis," *Journal of Monetary Economics*, 50, 1153-75.

McCallum, Bennett T., 2009, "Inflation Determination With Taylor Rules: Is the new-Keynesian Analysis Critically Flawed? *Journal of Monetary Economics* 56, 1101-1108.

Minford, Patrick, Francesco Perugini and Naveen Srinivasan, 2001,"The Observational Equivalence of Taylor Rule and Taylor-Type Rules" Forthcoming in F. Columbus ed., *Progress in Macroeconomics Research,* New York: Nova Science Publishers, also Centre for Economic Policy Research Discussion Paper No. 2959, September 2001.

Minford, Patrick Francesco Perugini and Naveen Srinivasan, 2002, "Are Interest Rate Regressions Evidence for a Taylor Rule?," *Economics Letters,* 76 (1), 145-150.

Minord, Patrick, and Naveen Srinivasan, 2010, "Determinacy in New Keynesian models: a role for money after all?, Manuscript, Cardiff Business School.

Obstfeld, Maurice and Kenneth Rogoff, 1983, Speculative Hyperinflations in Maximizing Models: Can we Rule them Out?" *Journal of Political Economy 91, 675-687.*

Obstfeld, Maurice and Kenneth Rogoff, 1986,"Ruling out Divergent Speculative Bubbles," *Journal of Monetary Economics* 17, 349-362.

Ontaski, Alexei, and Noah Williams, 2004, "Empirical and Policy Performance of a Forward-Looking Monetary Model," Manuscript, Columbia University and Princeton University. http://www.columbia.edu/~ao2027/forward16.pdf

Patinkin, Don, 1949, "The Indeterminacy of Absolute Prices in Classical Economic Theory," *Econometrica* 17, 1-27.

Patinkin, Don, 1965, *Money, Interest, and Prices*, 2nd Edition. New York: Harper and Row.

Sargent, Thomas J., and Neil Wallace, 1975, "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy* 83, 241-254.

Schmitt-Grohé, Stephanie and Martín Uribe, 2000, "Price Level Determinacy and Monetary Policy Under a Balanced Budget Requirement," *Journal of Monetary Economics* 45, 211-246.

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

Sims, Christopher A., 1994, "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy," *Economic Theory* 4, 381-99.

Sims, Christopher, 2003, "Limits to Inflation Targeting," in Bernanke, Ben S., and Michael Woodford, eds., *The Inflation-Targeting Debate*, NBER Studies in Business Cycles Volume 32, Ch. 7, p. 283-310. Chicago: University of Chicago Press. http://sims.princeton.edu/yftp/Targeting/TargetingFiscalPaper.pdf.

Smets, Frank, and Raf Wouters, 2003, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1(5), 1123-1175.

Technical Appendix to "Determinacy and Identification with Taylor Rules"

This technical appendix documents calculations made in the text and in the online Appendix.

1 Budget constraints in the frictionless model

This section presents a somewhat more careful discussion of budget constraints in the model of Section 3.1.

The household flow budget constraint states that bonds sold + income + transfers = consumption + (taxes-transfers) + bonds bought,

$$B_{t-1} + P_t Y_t = P_t \left(C_t + S_t\right) + Q_{t,t+1} B_t.$$

The household also faces a constraint that the present value of terminal wealth is zero.

$$\lim_{j\to\infty} E_t \left[Q_{t,t+j} B_{t+j-1}\right] = 0.$$

This latter "transversality condition" can be weakened to a bound on borrowing, since the consumer will never choose to overaccumulate wealth. In finite economies, you can't die with debts, and this is the limit of that condition for infinite-period economies. The condition also prevents consumers from arbitrage between a sequence of spot markets and markets for infinitely-lived contingent claims. Iterating forward, these two conditions are equivalent to the present value budget constraint,

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

In real terms, these constraints are

$$\frac{B_{t-1}}{P_t} + Y_t + G_t = C_t + S_t + E_t \left(m_{t,t+1} \frac{B_t}{P_{t+1}}\right)$$

$$\lim_{j\to\infty} E_t \left[m_{t,t+j} \frac{B_{t+j-1}(t+j)}{P_{t+j}}\right] = 0.$$

and hence

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

The government faces a flow identity, taxes plus bonds sold = spending plus bonds redeemed,

$$Q_{t,t+1} B_t + P_t S_t = B_{t-1}$$

or, in real terms

$$E_t \left(m_{t,t+1} \frac{B_t}{P_{t+1}}\right) + S_t = \frac{B_{t-1}}{P_t}.$$

It is tempting to iterate this forward as well and derive a government "intertemporal budget constraint." However, the government does *not* face a transversality condition. This fact is easiest to see in a finite economy. As a matter of budget constraint, we do not let agents die with debts. However, suppose agents developed a utility for government debt; they decide it makes nice wallpaper and are willing to hold it in the last period rather than cash it in and consume the proceeds. They can do this, and nothing in the government budget constraint should rule this out. The statement that government debt is zero at the end of time is an equilibrium condition, deriving from the fact that consumers without such utility will choose not to hold it. Thus, the government can, as a matter of constraint, make policy plans that, at off-equilibrium prices, would violate the consumer's budget constraint.

Thus, Equation (21) is an equilibrium condition that derives from the consumer's present value budget constraint, equilibrium $C = Y$, and the transversality condition for the consumer's choice to be an optimum. It is not a "government budget constraint." A "budget constraint" limits the demands an agent can announce at off-equilibrium prices, and there is nothing that stops the government from announcing plans that violate this equilibrium condition at off-equilibrium prices. Cochrane (2005) gives an extended discussion of this point.

2 Expected values

In Appendix section 3.2, I exhibited a construction of the unobserved monetary policy shock process $x_t = b(L)\varepsilon_t$ that rationalizes any equilibrium $\{i_t, \pi_t\}$ for hypothesized value of the Taylor parameter ϕ. While my proof checked that the construction of $b(L)$ in (66) and (67) works, it does not show how I derived those formulas. At the cost of greater algebra, this section shows where they come from.

We often generate one series (price, consumption, inflation) as an expected discounted sum of another (dividends, income, policy disturbances)

$$\begin{aligned} y_t &= E_t \sum_{j=0}^{\infty} \theta^j x_{t+j} \\ x_t &= b(L)\varepsilon_t \end{aligned}$$

Task 1 Find a representation for $y_t = a(L)\varepsilon_t$. The answer is (Hansen and Sargent (1980))

$$y_t = \left(\frac{Lb(L) - \theta b(\theta)}{L - \theta} \right) \varepsilon_t.$$

Here's why. Start by writing out

$$y_t^* = \sum_{j=0}^{\infty} \theta^j x_{t+j} = \frac{1}{1 - \theta L^{-1}} x_t = \frac{1}{1 - \theta L^{-1}} b(L)\varepsilon_t.$$

$$y_t^* = \begin{array}{ccccccc} & & & b_0\varepsilon_t & +b_1\varepsilon_{t-1} & +b_2\varepsilon_{t-2} & +... \\ & & +(\theta b_0\varepsilon_{t+1}) & +\theta b_1\varepsilon_t & +\theta b_2\varepsilon_{t-1} & +\theta b_3\varepsilon_{t-2} & +... \\ & +(\theta^2 b_0\varepsilon_{t+2}) & +(\theta^2 b_1\varepsilon_{t+1}) & +\theta^2 b_2\varepsilon_t & +\theta^2 b_3\varepsilon_{t-1} & +\theta^2 b_4\varepsilon_{t-2} & +... \\ +(\theta^3 b_0\varepsilon_{t+3}) & +(\theta^3 b_1\varepsilon_{t+2}) & +(\theta^3 b_2\varepsilon_{t+1}) & +\theta^3 b_3\varepsilon_t & +\theta^3 b_4\varepsilon_{t-1} & +\theta^3 b_5\varepsilon_{t-2} & +... \end{array}$$

Now, y_t is formed by simply getting rid of all the terms involving future ε_{t+j}, which I put in parentheses. Next sum the columns. For example, the ε_{t+1} term is

$$\theta b_0 + \theta^2 b_1 + \theta^3 b_2 + ... = \theta b(\theta)$$

Thus, we can write

$$\begin{aligned}
y_t &= \left\{ \frac{b(L)}{1-\theta L^{-1}} - \left[\theta b(\theta) L^{-1} + \theta^2 b(\theta) L^{-2} + \theta^3 b(\theta) L^{-3} + ... \right] \right\} \varepsilon_t \\
&= \left\{ \frac{b(L)}{1-\theta L^{-1}} - b(\theta) \left[\theta L^{-1} + \theta^2 L^{-2} + \theta^3 L^{-3} + ... \right] \right\} \varepsilon_t \\
&= \left\{ \frac{b(L)}{1-\theta L^{-1}} - \frac{b(\theta)\theta L^{-1}}{1-\theta L^{-1}} \right\} \varepsilon_t \\
&= \left\{ \frac{Lb(L) - b(\theta)\theta}{L-\theta} \right\} \varepsilon_t
\end{aligned}$$

Example. Suppose

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

It's easy to work out by hand that

$$E_t \sum_{j=0} \theta^j x_{t+j} = \sum \theta^j \rho^j x_t = \frac{1}{1-\rho\theta} x_t = \frac{1}{1-\rho\theta} \frac{1}{1-\rho L} \varepsilon_t.$$

Our formula gives

$$\begin{aligned}
E_t \sum_{j=0} \theta^j x_{t+j} &= \left\{ \frac{\frac{L}{1-\rho L} - \frac{\theta}{1-\rho\theta}}{L-\theta} \right\} \varepsilon_t \\
&= \left\{ \frac{\frac{L(1-\rho\theta) - \theta(1-\rho L)}{(1-\rho L)(1-\rho\theta)}}{L-\theta} \right\} \varepsilon_t \\
&= \left\{ \frac{\frac{L-\theta}{(1-\rho L)(1-\rho\theta)}}{L-\theta} \right\} \varepsilon_t \\
&= \frac{1}{(1-\rho L)(1-\rho\theta)} \varepsilon_t
\end{aligned}$$

just as it should.

Task 2, reverse engineering Suppose you have a representation for $y_t = a(L)\varepsilon_t$. Construct an $x_t = b(L)\varepsilon_t$ that justifies it by $y_t = E_t \sum_{j=0}^{\infty} \theta^j x_{t+j}$. We want to end up with

$$a(L) = \frac{Lb(L) - \theta b(\theta)}{L-\theta}.$$

Solving,

$$a(L)(L-\theta) = Lb(L) - \theta b(\theta).$$

Evaluate at $L=0$ to find $b(\theta)$

$$\begin{aligned} a(0)(-\theta) &= -b(\theta)\theta \\ a(0) &= b(\theta) \end{aligned}$$

Then substitute

$$\begin{aligned} a(L)(L-\theta) &= Lb(L) - a(0)\theta \\ b(L) &= \frac{a(L)(L-\theta)+a(0)\theta}{L} \\ b(L) &= a(L)(1-\theta L^{-1}) + a(0)\theta L^{-1} \\ b(L) &= a(L) - \theta L^{-1}\left(a(L)-a(0)\right) \end{aligned} \tag{88}$$

That's the answer.

We can also write the answer out explicitly:

$$\begin{aligned} b(L) &= a_0 + a_1 L + a_2 L^2 + a_3 L^3 + ... - \theta L^{-1}\left(a_1 L + a_2 L^2 + ...\right) \\ &= \left(a_0 - \theta a_1\right) + \left(a_1 - \theta a_2\right) L + \left(a_2 - \theta a_3\right) L^2 + ... \end{aligned}$$

i.e.

$$b_j = a_j - \theta a_{j+1} \tag{89}$$

Our application. We have $y_t = -\phi^{-1} E_t \sum_{j=0}^{\infty} \phi^{-j} x_{t+j}$, i.e. $\theta = \phi^{-1}$ and we need to multiply x_t by an additional $-\phi$ after we're done. Equation (88) becomes

$$\begin{aligned} b(L) &= -\phi\left[a(L) - \phi^{-1}L^{-1}\left(a(L)-a(0)\right)\right] \\ b(L) &= -\phi a(L) + L^{-1}\left(a(L)-a(0)\right) \\ b(L) &= \left(L^{-1}-\phi\right)a(L) - L^{-1}a(0) \end{aligned}$$

3 Standard error algebra

This section documents the standard errors calculated in Table A1. For linear instrumental variables, the coefficients are

$$\beta = \left(E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t')\right)^{-1} E(x_t z_t')E(z_t z_t')^{-1}E(z_t y_t)$$

This is GMM with a specific a matrix,

$$a = E(x_t z_t')E(z_t z_t')^{-1}$$

i.e.

$$g_T = E(z_t \varepsilon_t) = E(z_t(y_t - x_t'\beta))$$

$$\begin{aligned} E(x_t z_t')E(z_t z_t')^{-1}E(z_t(y_t - x_t'\beta)) &= 0 \\ E(x_t z_t')E(z_t z_t')^{-1}E(z_t y_t) - E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t')\beta &= 0 \end{aligned}$$

The GMM formula for standard errors is then

$$Tcov(\hat{b}) = (ad)^{-1} aSa'(ad)^{-1\prime}$$

$$d = \frac{\partial g_T}{\partial b'} = E(z_t x_t')$$

The standard formula assuming conditionally homoskedastic errors and no lags is

$$S = \sigma_\varepsilon^2 E(z_t z_t')$$

hence

$$\begin{aligned} Tcov(\hat{b}) &= \sigma_\varepsilon^2 (E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t'))^{-1}E(x_t z_t')E(z_t z_t')^{-1} \times \\ &\quad \times E(z_t x_t')((E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t'))^{-1} \\ &= \sigma_\varepsilon^2 (E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t'))^{-1} \end{aligned}$$

For heteroskedasticity-consistent standard errors

$$S = E(z_t z_t' \varepsilon_t^2)$$

so

$$\begin{aligned} Tcov(\hat{b}) &= \sigma_\varepsilon^2 (E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t'))^{-1}E(x_t z_t')E(z_t z_t')^{-1} \times \\ &\quad \times E(z_t z_t' \varepsilon_t^2)E(z_t z_t')^{-1}E(z_t x_t')((E(x_t z_t')E(z_t z_t')^{-1}E(z_t x_t'))^{-1} \end{aligned}$$

To develop standard errors for ϕ when it's estimated as $b/(1-\rho)$, I map into GMM. The moment conditions and estimates are unaffected. The standard errors are based on.

$$\begin{aligned} g_T &= E\left\{ z_t \left[i_t - \left(a + \rho i_{t-1} + (1-\rho)\tilde{x}_t'\phi \right) \right] \right\} \\ d &= \frac{\partial g_T}{\partial \left[a \ \rho \ \phi' \right]} = \left[\begin{array}{ccc} E(z_t) & E(z_t i_{t-1}) - E\left(z_t \tilde{x}_t' \right) \phi & (1-\rho)E(z_t \tilde{x}_t') \end{array} \right] \end{aligned}$$

Here $\tilde{x}$ is only the subset of variables that are used for the Taylor rule, i.e. $[\pi_t \ y_t]$.

In the case of OLS,

$$d = \frac{\partial g_T}{\partial \left[a \ \rho \ \phi' \right]} = \left[\begin{array}{ccc} E \begin{bmatrix} 1 \\ i_{t-1} \\ \tilde{x}_t \end{bmatrix} & E\left(\begin{bmatrix} 1 \\ i_{t-1} \\ \tilde{x}_t \end{bmatrix} i_{t-1} \right) - E\left(\begin{bmatrix} 1 \\ i_{t-1} \\ \tilde{x}_t \end{bmatrix} \tilde{x}_t' \right) \phi & (1-\rho) E\left(\begin{bmatrix} 1 \\ i_{t-1} \\ \tilde{x}_t \end{bmatrix} \tilde{x}_t' \right) \end{array} \right]$$

We can also form standard errors for $b/(1-\rho) = \phi$ by the delta method.

$$\begin{aligned} b/(1-\rho) &\approx 1/(1-\bar{\rho}) * \left(b - \bar{b} \right) + \bar{b}/(1-\bar{\rho})^2 (\rho - \bar{\rho}) \\ var(b/(1-\rho)) &= 1/(1-\bar{\rho})^2 var(b) + \bar{b}^2/(1-\bar{\rho})^4 var(\rho) + 2\bar{b}/(1-\bar{\rho})^3 cov(b,\rho) \end{aligned}$$

Comparing delta method and GMM,

$$g_T = E(x_t \left(y_t - x_t' b\right)) = 0$$

$$\begin{aligned}
\hat{b} &= E(x_t x_t')^{-1} E(x_t y_t) \\
d &= E(x_t x_t') \\
a &= I \\
var(b) &= d^{-1} S d^{-1\prime}/T
\end{aligned}$$

They give the same result,

$$\begin{aligned}
\phi &= f(b); b = h(\phi) \\
g_T &= E(x_t \left(y_t - x_t' h(\phi)\right)) = 0 \\
\frac{dg_T}{\partial \phi_i} &= E(x_t x_t') \frac{\partial h}{\partial \phi_i} \\
\frac{dg_T}{\partial \phi'} &= E(x_t x_t') \frac{\partial h}{\partial \phi'} \\
var(\phi) &= \left(\partial h / \partial \phi'\right)^{-1} E(x_t x_t')^{-1} S .../T \\
&= \left(\partial f / \partial \phi'\right) E(x_t x_t')^{-1} S .../T
\end{aligned}$$

4 Three-equation model algebra

This section summarizes analytic solutions of the standard three-equation model. The model is usually solved numerically, but using computer algebra I am able to find analytic expressions for the eigenvectors which allow an analytic expression for the model solution. I follow the standard methodology, as outlined in Text Section 3.3.

The standard three-equation model with Taylor responses to current and future output and inflation is, in deviation form

$$\begin{aligned}
y_t &= E_t y_{t+1} - \sigma r_t \\
i_t &= r_t + E_t \pi_{t+1} \\
\pi_t &= \beta E_t \pi_{t+1} + \gamma y_t \\
i_t &= \phi_{\pi,0} \pi_t + \phi_{\pi,1} E_t \pi_{t+1} + \phi_{y,0} y_t + \phi_{y,1} E_t y_{t+1}
\end{aligned}$$

The steps to solution are

1. Express the model in standard form
2. Find eigenvalues; characterize the region of local determinacy
3. Find eigenvectors; express the solution of the model as endogenous variable = function of shocks.

4.1 Express the model in standard form.

The standard form of the model is

$$\begin{bmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} \frac{\beta+\sigma\beta\phi_{y,0}-\sigma\gamma\left(\phi_{\pi,1}-1\right)}{1-\sigma\phi_{y,1}} & \sigma\frac{\beta\phi_{\pi,0}+\left(\phi_{\pi,1}-1\right)}{1-\sigma\phi_{y,1}} \\ -\gamma & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}. \tag{90}$$

Derivation: Eliminating i and r,

$$y_t = E_t y_{t+1} - \sigma\left(\phi_{\pi,0}\pi_t + \left(\phi_{\pi,1}-1\right)E_t\pi_{t+1} + \phi_{y,0}y_t + \phi_{y,1}E_t y_{t+1}\right).$$

$$\begin{aligned} \left(1-\sigma\phi_{y,1}\right)E_t y_{t+1} - \sigma\left(\phi_{\pi,1}-1\right)E_t\pi_{t+1} &= \left(1+\sigma\phi_{y,0}\right)y_t + \sigma\phi_{\pi,0}\pi_t \\ \beta E_t\pi_{t+1} &= -\gamma y_t + \pi_t \end{aligned}$$

$$\begin{bmatrix} 1-\sigma\phi_{y,1} & -\sigma\left(\phi_{\pi,1}-1\right) \\ 0 & \beta \end{bmatrix} \begin{bmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} 1+\sigma\phi_{y,0} & \sigma\phi_{\pi,0} \\ -\gamma & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$$

$$\begin{bmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1-\sigma\phi_{y,1}} & \frac{\sigma\left(\phi_{\pi,1}-1\right)}{\beta\left(1-\sigma\phi_{y,1}\right)} \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} 1+\sigma\phi_{y,0} & \sigma\phi_{\pi,0} \\ -\gamma & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$$

$$\begin{bmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1+\sigma\phi_{y,0}-\sigma\gamma\left(\phi_{\pi,1}-1\right)/\beta}{1-\sigma\phi_{y,1}} & \sigma\frac{\phi_{\pi,0}+\left(\phi_{\pi,1}-1\right)/\beta}{1-\sigma\phi_{y,1}} \\ -\frac{\gamma}{\beta} & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$$

$$\begin{bmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = \frac{1}{\beta} \begin{bmatrix} \frac{\beta+\sigma\beta\phi_{y,0}-\sigma\gamma\left(\phi_{\pi,1}-1\right)}{1-\sigma\phi_{y,1}} & \sigma\frac{\beta\phi_{\pi,0}+\left(\phi_{\pi,1}-1\right)}{1-\sigma\phi_{y,1}} \\ -\gamma & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$$

4.2 Find Eigenvalues

The eigenvalues of the transition matrix solve the following quadratic equation:

$$\beta\left(1-\sigma\phi_{y,1}\right)\lambda^2 - \left[1+\beta+\sigma\gamma(1-\phi_{\pi,1})+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\right]\lambda + 1 + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0} = 0$$

Derivation:

$$\begin{Vmatrix} \frac{1+\sigma\phi_{y,0}+\sigma\gamma(1-\phi_{\pi,1})/\beta}{1-\sigma\phi_{y,1}} - \lambda & \sigma\frac{\phi_{\pi,0}-(1-\phi_{\pi,1})/\beta}{1-\sigma\phi_{y,1}} \\ -\gamma/\beta & 1/\beta-\lambda \end{Vmatrix} = 0$$

$$\begin{aligned} \left(\frac{1+\sigma\phi_{y,0}+\sigma\gamma(1-\phi_{\pi,1})/\beta}{1-\sigma\phi_{y,1}} - \lambda\right)(1/\beta-\lambda) + \sigma\gamma\left(\frac{\phi_{\pi,0}-(1-\phi_{\pi,1})/\beta}{1-\sigma\phi_{y,1}}\right)/\beta &= 0 \\ \left[1+\sigma\phi_{y,0}+\sigma\gamma(1-\phi_{\pi,1})/\beta - \lambda\left(1-\sigma\phi_{y,1}\right)\right](1-\lambda\beta) + \sigma\gamma\left(\phi_{\pi,0}-(1-\phi_{\pi,1})/\beta\right) &= 0 \end{aligned}$$

$$\begin{aligned} 0 &= \beta\left(1-\sigma\phi_{y,1}\right)\lambda^2 - \left[\left(1+\sigma\phi_{y,0}+\sigma\gamma(1-\phi_{\pi,1})/\beta\right)\beta + \left(1-\sigma\phi_{y,1}\right)\right]\lambda \\ &\quad +1+\sigma\phi_{y,0}+\sigma\gamma(1-\phi_{\pi,1})/\beta + \sigma\gamma\left(\phi_{\pi,0}-(1-\phi_{\pi,1})/\beta\right) \end{aligned}$$

$$\beta\left(1-\sigma\phi_{y,1}\right)\lambda^2-\left[1+\beta+\sigma\gamma(1-\phi_{\pi,1})+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\right]\lambda+1+\sigma\phi_{y,0}+\sigma\gamma\phi_{\pi,0}=0.$$

There are two solutions of the quadratic equation defining eigenvalues, depending on parameter values:

1. If $\sigma\phi_{y,1}\neq 1$,

$$\begin{aligned}\lambda &= \frac{1}{2\beta\left(1-\sigma\phi_{y,1}\right)}\left\{1+\beta+\sigma\gamma\left(1-\phi_{\pi,1}\right)+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\right.\\ &\left.\pm\sqrt{\left(1+\beta+\sigma\gamma\left(1-\phi_{\pi,1}\right)+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\right)^2-4\beta\left(1-\sigma\phi_{y,1}\right)\left(1+\sigma\phi_{y,0}+\sigma\gamma\phi_{\pi,0}\right)}\right\}\end{aligned}$$

2. If $\sigma\phi_{y,1}=1$,

$$\beta\left(1-\sigma\phi_{y,1}\right)\lambda^2-\left[1+\beta+\sigma\gamma(1-\phi_{\pi,1})+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\right]\lambda+1+\sigma\phi_{y,0}+\sigma\gamma\phi_{\pi,0}=0$$

$$-\left[\beta+\sigma\gamma(1-\phi_{\pi,1})+\sigma\beta\phi_{y,0}\right]\lambda+1+\sigma\phi_{y,0}+\sigma\gamma\phi_{\pi,0}=0$$

$$\lambda=\frac{1+\sigma\left(\phi_{y,0}+\gamma\phi_{\pi,0}\right)}{\beta+\sigma\left[\gamma(1-\phi_{\pi,1})+\beta\phi_{y,0}\right]}$$

4.3 Characterize the region of local determinacy

The region of local determinacy is the part of the parameter space that produces two eigenvalues greater than one. That region is *not* as simple as "the sum of the Taylor coefficients on inflation is greater than one."

To find the regions of determinacy, write

$$\begin{aligned}\lambda &= \frac{1}{2a}\left(b\pm\sqrt{b^2-4ac}\right)\\ a &\equiv \beta\left(1-\sigma\phi_{y,1}\right)\\ b &\equiv 1+\beta+\sigma\gamma\left(1-\phi_{\pi,1}\right)+\sigma\beta\phi_{y,0}-\sigma\phi_{y,1}\\ c &\equiv 1+\sigma\phi_{y,0}+\sigma\gamma\phi_{\pi,0}\end{aligned}$$

The boundaries $\|\lambda\|=1$ are as follows. (See "Detailed Algebra" below.)

1. $\sigma\phi_{y,1}\neq 1$, real roots $b^2-4ac>0$, $\lambda=1$:

$$\left(\phi_{\pi,0}+\phi_{\pi,1}-1\right)+\frac{1-\beta}{\gamma}\left(\phi_{y,1}+\phi_{y,0}\right)=0$$

2. $\sigma\phi_{y,1}\neq 1$, real roots $b^2-4ac>0$, $\lambda=-1$:

$$\left(1+\phi_{\pi,0}-\phi_{\pi,1}\right)-\frac{1+\beta}{\gamma}\left(\phi_{y,1}-\phi_{y,0}\right)=-2\frac{(1+\beta)}{\sigma\gamma}$$

3. $\sigma\phi_{y,1} \neq 1$, Complex roots $b^2 - 4ac < 0$,

$$\gamma\phi_{\pi,0} + \phi_{y,0} + \beta\phi_{y,1} = \frac{\beta - 1}{\sigma}.$$

4. $\sigma\phi_{y,1} = 1$, $\lambda = 1$,

$$\phi_{\pi,0} + \phi_{\pi,1} + \frac{(1-\beta)}{\sigma\gamma}\left(1 + \sigma\phi_{y,0}\right) = 1$$

5. $\sigma\phi_{y,1} = 1$, $\lambda = -1$:

$$\phi_{\pi,0} - \phi_{\pi,1} + \frac{(1+\beta)}{\sigma\gamma}\left(1 + \sigma\phi_{y,0}\right) = -1$$

4.4 Special cases

a. Only $\phi_{\pi,0} \neq 0$.

$$\lambda = \frac{1}{2\beta}\left((1+\beta+\sigma\gamma) \pm \sqrt{(1+\beta+\sigma\gamma)^2 - 4\beta\left(1+\sigma\gamma\phi_{\pi,0}\right)}\right)$$

The condition for real roots is $(1+\beta+\sigma\gamma)^2 - 4\beta\left(1+\sigma\gamma\phi_{\pi,0}\right) > 0$. The $\|\lambda\| = 1$ regions are then

$$\phi_{\pi,0} = 1$$

$$\phi_{\pi,0} = -\left(1 + 2\frac{(1+\beta)}{\sigma\gamma}\right)$$

for complex roots, we have

$$\phi_{\pi,0} = \frac{\beta - 1}{\sigma\gamma}$$

b. No output response, both inflation responses. $\phi_{\pi,0}, \phi_{\pi,1} \neq 0$.

$$\lambda = \frac{1}{2\beta}\left(1+\beta+\sigma\gamma\left(1-\phi_{\pi,1}\right) \pm \sqrt{\left(1+\beta+\sigma\gamma\left(1-\phi_{\pi,1}\right)\right)^2 - 4\beta\left(1+\sigma\gamma\phi_{\pi,0}\right)}\right)$$

The boundaries $\|\lambda\| = 1$ are as follows.

1. Real roots, $\lambda = 1$

$$\phi_{\pi,0} + \phi_{\pi,1} = 1$$

2. Real roots, $\lambda = -1$:

$$\phi_{\pi,0} - \phi_{\pi,1} = -\left(1 + 2\frac{(1+\beta)}{\sigma\gamma}\right)$$

3. Complex roots

$$\phi_{\pi,0} = \frac{\beta - 1}{\sigma\gamma}.$$

In the case $\phi_{\pi,0} = 0$, we have

$$\|\lambda\|^2 = \frac{1}{4\beta^2} 4\beta = \frac{1}{\beta} > 1$$

in the entire complex root region. (The complex root region in the plot with $\phi_{\pi,0} = 0$ has a very small band of real roots surrounding the plotted complex roots, and these decline quickly to one at the plotted boundary.)

c. *Contemporaneous output and inflation responses.* $\phi_{\pi,0}, \phi_{y,0}$

$$\begin{aligned}
\lambda &= \frac{1}{2a}\left(b \pm \sqrt{b^2 - 4ac}\right) \\
a &\equiv \beta \\
b &\equiv 1 + \beta + \sigma\gamma + \sigma\beta\phi_{y,0} \\
c &\equiv 1 + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0}
\end{aligned}$$

The boundaries $\|\lambda\| = 1$ in this case are as follows.

1. $\sigma\phi_{y,1} \neq 1$, real roots $b^2 - 4ac > 0$, $\lambda = 1$:

$$\phi_{\pi,0} + \frac{1-\beta}{\gamma}\phi_{y,0} = 1$$

2. $\sigma\phi_{y,1} \neq 1$, real roots $b^2 - 4ac > 0$, $\lambda = -1:$

$$\phi_{\pi,0} + \frac{1+\beta}{\gamma}\phi_{y,0} = -\left(1 + 2\frac{(1+\beta)}{\sigma\gamma}\right)$$

3. $\sigma\phi_{y,1} \neq 1$, Complex roots $b^2 - 4ac < 0$,

$$\gamma\phi_{\pi,0} + \phi_{y,0} = \frac{\beta - 1}{\sigma}.$$

Detailed algebra for determinacy regions:

1. $\sigma\phi_{y,1} \neq 1$, real roots, $\lambda = 1$.

$$\frac{1}{2a}\left(b \pm \sqrt{b^2 - 4ac}\right) = 1$$

$$\begin{aligned}
b \pm \sqrt{b^2 - 4ac} &= 2a \\
b^2 - 4ac &= (2a - b)^2
\end{aligned}$$

$$\begin{aligned}
0 = {} & \left(1 + \beta + \sigma\gamma\left(1 - \phi_{\pi,1}\right) + \sigma\beta\phi_{y,0} - \sigma\phi_{y,1}\right)^2 - 4\beta\left(1 - \sigma\phi_{y,1}\right)\left(1 + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0}\right) \\
& - \left(2\beta - 2\beta\sigma\phi_{y,1} - \left(1 + \beta + \sigma\gamma\left(1 - \phi_{\pi,1}\right) + \sigma\beta\phi_{y,0} - \sigma\phi_{y,1}\right)\right)^2
\end{aligned}$$

$$
\begin{aligned}
0 \;=\; & -4\sigma\beta\phi_{y,0} - 4\beta\sigma\phi_{y,1} + 4\sigma^2\beta\phi_{y,0}\phi_{y,1} - 4\beta\sigma\gamma\phi_{\pi,1} - 4\beta\sigma\gamma\phi_{\pi,0} \\
& -4\beta\sigma^2\phi_{y,1}\gamma - 4\beta^2\sigma^2\phi_{y,1}\phi_{y,0} + 4\sigma\beta^2\phi_{y,0} + 4\beta\sigma\gamma \\
& +4\beta^2\sigma\phi_{y,1} - 4\beta^2\sigma^2\phi_{y,1}^2 + 4\beta\sigma^2\phi_{y,1}^2 + 4\beta\sigma^2\phi_{y,1}\gamma\phi_{\pi,0} + 4\beta\sigma^2\phi_{y,1}\gamma\phi_{\pi,1}
\end{aligned}
$$

$$
\begin{aligned}
0 \;=\; & -\phi_{y,0} - \phi_{y,1} + \sigma\phi_{y,0}\phi_{y,1} - \gamma\phi_{\pi,1} - \gamma\phi_{\pi,0} - \sigma\phi_{y,1}\gamma - \beta\sigma\phi_{y,1}\phi_{y,0} \\
& +\beta\phi_{y,0} + \gamma + \beta\phi_{y,1} - \beta\sigma\phi_{y,1}^2 + \sigma\phi_{y,1}^2 + \sigma\phi_{y,1}\gamma\phi_{\pi,0} + \sigma\phi_{y,1}\gamma\phi_{\pi,1}
\end{aligned}
$$

$$
\left(1 - \sigma\phi_{y,1}\right)\left(\beta\phi_{y,1} - \phi_{y,1} - \phi_{y,0} + \gamma + \beta\phi_{y,0} - \gamma\phi_{\pi,0} - \gamma\phi_{\pi,1}\right) = 0
$$

$$
\left(1 - \sigma\phi_{y,1}\right)\left((\beta - 1)\left(\phi_{y,1} + \phi_{y,0}\right) - \gamma\left(\phi_{\pi,0} + \phi_{\pi,1} - 1\right)\right) = 0
$$

We have already assumed $\sigma\phi_{y,1} \neq 1$,so

$$
\frac{(\beta - 1)}{\gamma}\left(\phi_{y,1} + \phi_{y,0}\right) - \left(\phi_{\pi,0} + \phi_{\pi,1} - 1\right) = 0
$$

$$
\left(\phi_{\pi,0} + \phi_{\pi,1} - 1\right) + \frac{1 - \beta}{\gamma}\left(\phi_{y,1} + \phi_{y,0}\right) = 0
$$

This identifies parameters at which *one* eigenvalue is equal to one. We also have to check that the other one is greater than one. I do this numerically to make the plots.

2. $\sigma\phi_{y,1} \neq 1$, real roots, $\lambda = -1$

$$
\frac{1}{2a}\left(b + \sqrt{b^2 - 4ac}\right) = -1
$$

$$
b^2 - 4ac = (2a + b)^2
$$

$$
\begin{aligned}
0 \;=\; & \left(1 + \beta + \sigma\gamma\left(1 - \phi_{\pi,1}\right) + \sigma\beta\phi_{y,0} - \sigma\phi_{y,1}\right)^2 - 4\beta\left(1 - \sigma\phi_{y,1}\right)\left(1 + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0}\right) \\
& - \left(2\beta - 2\beta\sigma\phi_{y,1} + \left(1 + \beta + \sigma\gamma\left(1 - \phi_{\pi,1}\right) + \sigma\beta\phi_{y,0} - \sigma\phi_{y,1}\right)\right)^2
\end{aligned}
$$

$$
\begin{aligned}
0 \;=\; & -4\sigma\beta\phi_{y,0} + 12\beta\sigma\phi_{y,1} - 8\beta^2 - 8\beta + 4\sigma^2\beta\phi_{y,0}\phi_{y,1} + 4\beta\sigma\gamma\phi_{\pi,1} - 4\beta\sigma\gamma\phi_{\pi,0} \\
& +4\beta\sigma^2\phi_{y,1}\gamma + 4\beta^2\sigma^2\phi_{y,1}\phi_{y,0} - 4\sigma\beta^2\phi_{y,0} - 4\beta\sigma\gamma \\
& +12\beta^2\sigma\phi_{y,1} - 4\beta^2\sigma^2\phi_{y,1}^2 - 4\beta\sigma^2\phi_{y,1}^2 + 4\beta\sigma^2\phi_{y,1}\gamma\phi_{\pi,0} - 4\beta\sigma^2\phi_{y,1}\gamma\phi_{\pi,1}
\end{aligned}
$$

$$
\left(\sigma\phi_{y,1} - 1\right)\left(-\beta\sigma\phi_{y,1} - \sigma\phi_{y,1} + \sigma\beta\phi_{y,0} + \sigma\gamma + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0} - \sigma\gamma\phi_{\pi,1} + 2 + 2\beta\right) = 0
$$

$$
-\beta\sigma\phi_{y,1} - \sigma\phi_{y,1} + \sigma\beta\phi_{y,0} + \sigma\gamma + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0} - \sigma\gamma\phi_{\pi,1} + 2 + 2\beta = 0
$$

$$
(1 + \beta)\left(\phi_{y,1} - \phi_{y,0}\right) + \gamma\left(\phi_{\pi,1} - \phi_{\pi,0} - 1\right) = 2\frac{(1 + \beta)}{\sigma}
$$

$$
\left(\phi_{\pi,1} - \phi_{\pi,0} - 1\right) + \frac{(1 + \beta)}{\gamma}\left(\phi_{y,1} - \phi_{y,0}\right) = 2\frac{(1 + \beta)}{\sigma\gamma}
$$

$$
\left(1 - \phi_{\pi,1} + \phi_{\pi,0}\right) - \frac{(1 + \beta)}{\gamma}\left(\phi_{y,1} - \phi_{y,0}\right) = -2\frac{(1 + \beta)}{\sigma\gamma}
$$

3. $\sigma\phi_{y,1} \neq 1$, Complex roots,

$$\left\| \frac{1}{2a} \left(b \pm \sqrt{b^2 - 4ac} \right) \right\| = 1$$

$$\begin{aligned}
\left(b - i\sqrt{\|(b^2 - 4ac)\|} \right) \left(b + i\sqrt{\|(b^2 - 4ac)\|} \right) &= \|2a\| \\
\left(b^2 + \left\| \left(b^2 - 4ac \right) \right\| \right) &= (2a)^2
\end{aligned}$$

the roots are complex because $b^2 - 4ac < 0$

$$\left(b^2 - \left(b^2 - 4ac \right) \right) = (2a)^2$$

$$4ac = 4a^2$$

$$\begin{aligned}
c &= a \\
1 + \sigma\phi_{y,0} + \sigma\gamma\phi_{\pi,0} &= \beta \left(1 - \sigma\phi_{y,1} \right) \\
\gamma\phi_{\pi,0} + \phi_{y,0} + \beta\phi_{y,1} &= \frac{\beta - 1}{\sigma}.
\end{aligned}$$

In the special case $\sigma\phi_{y,1} = 1$, we have

$$\lambda = \frac{1 + \sigma \left(\phi_{y,0} + \gamma\phi_{\pi,0} \right)}{\beta + \sigma \left[\gamma(1 - \phi_{\pi,1}) + \beta\phi_{y,0} \right]}$$

4. $\lambda = 1$:

$$1 + \sigma \left(\phi_{y,0} + \gamma\phi_{\pi,0} \right) = \beta + \sigma \left[\gamma(1 - \phi_{\pi,1}) + \beta\phi_{y,0} \right]$$

$$\begin{aligned}
\frac{1 - \beta}{\sigma} &= \gamma(1 - \phi_{\pi,1}) + \beta\phi_{y,0} - \left(\phi_{y,0} + \gamma\phi_{\pi,0} \right) \\
\frac{1 - \beta}{\sigma} &= -\gamma \left(\phi_{\pi,1} + \phi_{\pi,0} - 1 \right) + (\beta - 1)\,\phi_{y,0} \\
-\frac{1 - \beta}{\sigma\gamma} &= \left(\phi_{\pi,0} + \phi_{\pi,1} - 1 \right) + \frac{(1 - \beta)}{\gamma}\phi_{y,0} \\
\left(\phi_{\pi,0} + \phi_{\pi,1} - 1 \right) &= -\frac{(1 - \beta)}{\sigma\gamma} \left(1 + \sigma\phi_{y,0} \right) \\
\phi_{\pi,0} + \phi_{\pi,1} + \frac{(1 - \beta)}{\sigma\gamma} \left(1 + \sigma\phi_{y,0} \right) &= 1
\end{aligned}$$

5. $\lambda = -1$:

$$1 + \sigma \left(\phi_{y,0} + \gamma\phi_{\pi,0} \right) = -\beta - \sigma \left[\gamma(1 - \phi_{\pi,1}) + \beta\phi_{y,0} \right]$$

$$(1 + \phi_{\pi,0} - \phi_{\pi,1}) = -\frac{(1 + \beta)}{\sigma\gamma}\sigma\phi_{y,0} - \frac{1 + \beta}{\sigma\gamma}$$

$$\begin{aligned}
(1 + \phi_{\pi,0} - \phi_{\pi,1}) &= -\frac{(1 + \beta)}{\sigma\gamma} \left(1 + \sigma\phi_{y,0} \right) \\
\phi_{\pi,0} - \phi_{\pi,1} + \frac{(1 + \beta)}{\sigma\gamma} \left(1 + \sigma\phi_{y,0} \right) &= -1
\end{aligned}$$

4.5 Dynamics, responses, and estimated coefficients.

Next, we want to exhibit artificial data, compute impulse responses, and evaluate model dynamics. To that end, I add shocks to the IS, Phillips curves, and monetary policy rule. With shocks, the system is

$$\begin{aligned}
y_t &= E_t y_{t+1} - \sigma r_t + x_{dt} \\
i_t &= r_t + E_t \pi_{t+1} \\
\pi_t &= \beta E_t \pi_{t+1} + \gamma y_t + x_{\pi t} \\
i_t &= \phi_{\pi,0}\pi_t + \phi_{\pi,1} E_t \pi_{t+1} + \phi_{y,0} y_t + \phi_{y,1} E_t y_{t+1} + x_{it}
\end{aligned}$$

Standard form. I eliminate i, r to express the model in standard form, as above. Adding the responses of each variable to the x shocks, and including AR(1) dynamics for the disturbances, and including expectational errors δ_{t+1}, the standard form is

$$\begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ x_{dt+1} \\ x_{\pi t+1} \\ x_{it+1} \end{bmatrix} = \begin{bmatrix} \frac{1+\sigma\phi_{y,0}+\frac{\sigma\gamma}{\beta}\left(1-\phi_{\pi,1}\right)}{1-\sigma\phi_{y,1}} & \sigma\frac{\phi_{\pi,0}-\frac{\left(1-\phi_{\pi,1}\right)}{\beta}}{1-\sigma\phi_{y,1}} & -\frac{1}{1-\sigma\phi_{y,1}} & \frac{\sigma}{\beta}\frac{\left(1-\phi_{\pi,1}\right)}{\left(1-\sigma\phi_{y,1}\right)} & \frac{\sigma}{1-\sigma\phi_{y,1}} \\ -\frac{\gamma}{\beta} & \frac{1}{\beta} & 0 & -\frac{1}{\beta} & 0 \\ 0 & 0 & \rho_d & 0 & 0 \\ 0 & 0 & 0 & \rho_\pi & 0 \\ 0 & 0 & 0 & 0 & \rho_i \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \\ x_{dt} \\ x_{\pi t} \\ x_{it} \end{bmatrix} + \begin{bmatrix} \delta_{yt+1} \\ \delta_{\pi t+1} \\ \varepsilon_{dt+1} \\ \varepsilon_{\pi t+1} \\ \varepsilon_{it+1} \end{bmatrix}$$

Derivation:

$$E_t\pi_{t+1} = \frac{1}{\beta}\left(\pi_t - \gamma y_t - x_{\pi t}\right)$$

$$r_t = \phi_{\pi,0}\pi_t + \left(\phi_{\pi,1}-1\right) E_t\pi_{t+1} + \phi_{y,0}y_t + \phi_{y,1}E_t y_{t+1} + x_{it}$$

$$\begin{aligned}
E_t y_{t+1} &= y_t + \sigma r_t - x_{dt} \\
E_t y_{t+1} &= y_t + \sigma\left(\phi_{\pi,0}\pi_t + \left(\phi_{\pi,1}-1\right) E_t\pi_{t+1} + \phi_{y,0}y_t + \phi_{y,1}E_t y_{t+1} + x_{it}\right) - x_{dt} \\
(1-\sigma\phi_{y,1})E_t y_{t+1} &= (1+\sigma\phi_{y,0})y_t + \sigma\phi_{\pi,0}\pi_t + \frac{\sigma}{\beta}\left(\phi_{\pi,1}-1\right)\left(\pi_t - \gamma y_t - x_{\pi t}\right) + \sigma x_{it} - x_{dt}
\end{aligned}$$

$$\begin{aligned}
(1-\sigma\phi_{y,1})E_t y_{t+1} &= \left(1+\sigma\phi_{y,0}+\frac{\sigma\gamma}{\beta}\left(1-\phi_{\pi,1}\right)\right) y_t + \sigma\left(\phi_{\pi,0} - \frac{\left(1-\phi_{\pi,1}\right)}{\beta}\right)\pi_t + \\
&\quad + \frac{\sigma}{\beta}\left(1-\phi_{\pi,1}\right) x_{\pi t} + \sigma x_{it} - x_{dt}
\end{aligned}$$

Eigenvectors: The eigenvectors corresponding to the stable eigenvalues are (thanks to Scientific Workplace, but verified)

$$\begin{bmatrix} 1-\beta\rho_d \\ \gamma \\ \left(1-\rho_d\right)\left(1-\rho_d\beta\right) + \sigma\left(1-\beta\rho_d\right)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right) + \sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right) \\ 0 \\ 0 \end{bmatrix} \leftrightarrow \rho_d,$$

$$\begin{bmatrix} -\sigma\left(\phi_{\pi,0}+\rho_{\pi}\left(\phi_{\pi,1}-1\right)\right) \\ (1-\rho_{\pi})+\sigma\left(\phi_{y,0}+\rho_{\pi}\phi_{y,1}\right) \\ 0 \\ (1-\rho_{\pi})(1-\rho_{\pi}\beta)+\sigma(1-\beta\rho_{\pi})\left(\phi_{y,0}+\rho_{\pi}\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_{\pi}\left(\phi_{\pi,1}-1\right)\right) \\ 0 \end{bmatrix} \leftrightarrow \rho_{\pi},$$

$$\begin{bmatrix} -\sigma(1-\beta\rho_i) \\ -\sigma\gamma \\ 0 \\ 0 \\ (1-\rho_i)(1-\rho_i\beta)+\sigma(1-\beta\rho_i)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right) \end{bmatrix} \leftrightarrow \rho_i$$

Solution for $\{y,\pi\}$. Thus, the solution is

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 1-\rho_d\beta & \sigma\left(\rho_{\pi}\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) & -\sigma(1-\rho_i\beta) \\ \gamma & (1-\rho_{\pi})+\sigma(\phi_{y,0}+\rho_{\pi}\phi_{y,1}) & -\sigma\gamma \end{bmatrix} \begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix}$$

$$\begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix} = \begin{bmatrix} \rho_d & 0 & 0 \\ 0 & \rho_{\pi} & 0 \\ 0 & 0 & \rho_i \end{bmatrix} \begin{bmatrix} z_{dt-1} \\ z_{\pi t-1} \\ z_{it-1} \end{bmatrix} + \begin{matrix} v_{dt} \\ v_{\pi t} \\ v_{it} \end{matrix}$$

$$\begin{aligned} x_{dt} &= \left[(1-\rho_d)(1-\rho_d\beta)+\sigma(1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right)\right] z_{dt} \\ x_{\pi t} &= \left[(1-\rho_{\pi})(1-\rho_{\pi}\beta)+\sigma(1-\beta\rho_{\pi})\left(\phi_{y,0}+\rho_{\pi}\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_{\pi}\left(\phi_{\pi,1}-1\right)\right)\right] z_{\pi t} \\ x_{it} &= \left[(1-\rho_i)(1-\rho_i\beta)+\sigma(1-\beta\rho_i)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right)\right] z_{it} \end{aligned}$$

Here I express the solution in terms of z variables. Since the z are simply scalar multiples of the x one can also express the solution in terms of x.

Include interest rates. Adding interest rates back to the system, the full solution is

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = A \begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix} \tag{91}$$

where the columns of A are

$$
\begin{aligned}
A_{:,1} &= \begin{bmatrix} 1-\rho_d\beta \\ \gamma \\ (1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\gamma\left(\phi_{\pi,0}+\rho_d\phi_{\pi,1}\right) \end{bmatrix} \\
A_{:,2:} &= \begin{bmatrix} \sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) \\ (1-\rho_\pi)+\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1}) \\ (1-\rho_\pi)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)+\rho_\pi\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1}) \end{bmatrix} \\
A_{:,3} &= \begin{bmatrix} -\sigma\left(1-\rho_i\beta\right) \\ -\sigma\gamma \\ (1-\rho_i)\left(1-\rho_i\beta\right)-\sigma\gamma\rho_i \end{bmatrix};
\end{aligned}
$$

and

$$
\begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \end{bmatrix} = \begin{bmatrix} \rho_d & 0 & 0 \\ 0 & \rho_\pi & 0 \\ 0 & 0 & \rho_i \end{bmatrix} \begin{bmatrix} z_{dt-1} \\ z_{\pi t-1} \\ z_{it-1} \end{bmatrix} + \begin{matrix} v_{dt} \\ v_{\pi t} \\ v_{it} \end{matrix} \tag{92}
$$

$$
\begin{aligned}
x_{dt} &= \left[(1-\rho_d)(1-\rho_d\beta)+\sigma(1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right)\right]z_{dt} \\
x_{\pi t} &= \left[(1-\rho_\pi)(1-\rho_\pi\beta)+\sigma(1-\beta\rho_\pi)\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_\pi\left(\phi_{\pi,1}-1\right)\right)\right]z_{\pi t} \\
x_{it} &= \left[(1-\rho_i)(1-\rho_i\beta)+\sigma(1-\beta\rho_i)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right)\right]\ z_{it}
\end{aligned}
$$

Since the z are just multiples of the x, one can eliminate them, at the cost of larger formulas. The formula seems daunting, but it simplifies usefully in the obvious special cases.

Derivation

$$
\begin{aligned}
y_t &= E_t y_{t+1}-\sigma(i_t-E_t\pi_{t+1})+x_{dt} \\
i_t &= \frac{1}{\sigma}\left(E_t y_{t+1}-y_t\right)+\frac{1}{\sigma}x_{dt}+E_t\pi_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
i_t \quad = & \begin{bmatrix} -\frac{1}{\sigma}(1-\rho_d)(1-\rho_d\beta)+\rho_d\gamma \\ +\frac{1}{\sigma}\left(((1-\rho_d)(1-\rho_d\beta)+\sigma(1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right))\right) \end{bmatrix} z_{dt} \\
& +\left[\frac{1}{\sigma}(\rho_\pi-1)\sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right)+\rho_\pi\left((1-\rho_\pi)+\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1})\right)\right]z_{\pi t} \\
& +\left[\frac{1}{\sigma}(\rho_i-1)\left(-\sigma(1-\rho_i\beta)\right)-\sigma\gamma\rho_i\right]z_{it}
\end{aligned}
$$

$$
\begin{aligned}
i_t \quad = & \left[\ (1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\gamma\left(\phi_{\pi,0}+\rho_d\phi_{\pi,1}\right)\ \right]z_{dt} \\
& +\left[(1-\rho_\pi)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)+\rho_\pi\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1})\right]z_{\pi t} \\
& +\left[(1-\rho_i)(1-\rho_i\beta)-\sigma\gamma\rho_i\right]z_{it}
\end{aligned}
$$

4.6 Dynamics with potential output

Adding potential output as another AR(1) shock, and assuming the Taylor rule responds to potential output, we have

$$
\begin{aligned}
y_t &= E_t y_{t+1} - \sigma r_t + x_{dt} \\
i_t &= r_t + E_t \pi_{t+1} \\
\pi_t &= \beta E_t \pi_{t+1} + \gamma \left(y_t - \bar{y}_t \right) + x_{\pi t} \\
i_t &= \phi_{\pi,0} \pi_t + \phi_{\pi,1} E_t \pi_{t+1} + \phi_{y,0} \left(y_t - \bar{y}_t \right) + \phi_{y,1} E_t \left(y_{t+1} - \bar{y}_t \right) + x_{it}
\end{aligned}
$$

Following the same steps, the standard form is

$$
\begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ x_{dt+1} \\ x_{\pi t+1} \\ x_{it+1} \end{bmatrix} = A \begin{bmatrix} y_t \\ \pi_t \\ x_{dt} \\ x_{\pi t} \\ x_{it} \\ \bar{y}_t \end{bmatrix} + \begin{bmatrix} \delta_{yt+1} \\ \delta_{\pi t+1} \\ \varepsilon_{dt+1} \\ \varepsilon_{\pi t+1} \\ \varepsilon_{it+1} \\ \varepsilon_{\bar{y}t+1} \end{bmatrix}
$$

where the columns of A are

$$
\begin{aligned}
A_{:,1:2} &= \begin{bmatrix} \frac{1+\sigma\phi_{y,0}+\frac{\sigma\gamma}{\beta}\left(1-\phi_{\pi,1}\right)}{1-\sigma\phi_{y,1}} & \frac{\sigma\left(\phi_{\pi,0}-\frac{\left(1-\phi_{\pi,1}\right)}{\beta}\right)}{1-\sigma\phi_{y,1}} \\ -\frac{\gamma}{\beta} & \frac{1}{\beta} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \\
A_{:,3:6} &= \begin{bmatrix} -\frac{1}{1-\sigma\phi_{y,1}} & \frac{\sigma}{\beta}\frac{\left(1-\phi_{\pi,1}\right)}{\left(1-\sigma\phi_{y,1}\right)} & \frac{\sigma}{\left(1-\sigma\phi_{y,1}\right)} & \frac{\sigma\left(\frac{\gamma}{\beta}\left(\phi_{\pi,1}-1\right)-\phi_{y,0}-\phi_{y,1}\rho_{\bar{y}}\right)}{\left(1-\sigma\phi_{y,1}\right)} \\ 0 & -\frac{1}{\beta} & 0 & \frac{\gamma}{\beta} \\ \rho_d & 0 & 0 & 0 \\ 0 & \rho_\pi & 0 & 0 \\ 0 & 0 & \rho_i & 0 \\ 0 & 0 & 0 & \rho_{\bar{y}} \end{bmatrix}
\end{aligned}
$$

The eigenvectors corresponding to the stable eigenvalues are

$$
\begin{bmatrix} -\sigma\left(\phi_{\pi,0}-\rho_\pi\left(\phi_{\pi,1}-1\right)\right) \\ \left(1-\rho_\pi\right)+\sigma\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right) \\ 0 \\ \left(1-\rho_\pi\right)\left(1-\rho_\pi\beta\right)+\sigma\left(1-\beta\rho_\pi\right)\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_\pi\left(\phi_{\pi,1}-1\right)\right) \\ 0 \\ 0 \end{bmatrix} \leftrightarrow \rho_\pi
$$

$$\begin{bmatrix} 1-\beta\rho_d \\ \gamma \\ (1-\rho_d)(1-\rho_d\beta)+\sigma(1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right) \\ 0 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow \rho_d$$

$$\begin{bmatrix} -\sigma(1-\beta\rho_i) \\ -\sigma\gamma \\ 0 \\ 0 \\ (1-\rho_i)(1-\rho_i\beta)+\sigma(1-\beta\rho_i)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right) \\ 0 \end{bmatrix} \leftrightarrow \rho_i$$

$$\begin{bmatrix} \sigma\left(-\gamma\rho_y+\left(1-\beta\rho_y\right)\phi_{y,0}+\rho_y\left(1-\beta\rho_y\right)\phi_{y,1}+\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)\right) \\ -\gamma\left(1-\rho_y\right) \\ 0 \\ 0 \\ 0 \\ \left(1-\rho_y\right)\left(1-\rho_y\beta\right)+\sigma\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_y\left(\phi_{\pi,1}-1\right)\right) \end{bmatrix} \leftrightarrow \rho_y$$

Thus, the solution is

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = A \begin{bmatrix} z_{dt} \\ z_{\pi t} \\ z_{it} \\ z_{yt} \end{bmatrix}$$

where the columns of A are

$$\begin{aligned} A_{:,1:2} &= \begin{bmatrix} 1-\rho_d\beta & \sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) \\ \gamma & (1-\rho_\pi)+\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1}) \end{bmatrix} \\ A_{:,3:4} &= \begin{bmatrix} -\sigma(1-\rho_i\beta) & \sigma\left(-\gamma\rho_y+\left(1-\beta\rho_y\right)\phi_{y,0}+\rho_y\left(1-\beta\rho_y\right)\phi_{y,1}+\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)\right) \\ -\sigma\gamma & -\gamma\left(1-\rho_y\right) \end{bmatrix} \end{aligned}$$

$$[z_t] = \rho\,[z_{t-1}] + [v_t]$$

$$\begin{aligned} x_{dt} &= \left((1-\rho_d)(1-\rho_d\beta)+\sigma(1-\beta\rho_d)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right)\right) z_{dt} \\ x_{\pi t} &= \left((1-\rho_\pi)(1-\rho_\pi\beta)+\sigma(1-\beta\rho_\pi)\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_\pi\left(\phi_{\pi,1}-1\right)\right)\right) z_{\pi t} \\ x_{it} &= \left((1-\rho_i)(1-\rho_i\beta)+\sigma(1-\beta\rho_i)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right)\right)\ z_{it} \\ x_{yt} &= \left(\left(1-\rho_y\right)\left(1-\rho_y\beta\right)+\sigma\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_y\left(\phi_{\pi,1}-1\right)\right)\right) z_{yt} \end{aligned}$$

Adding interest rates

$$
\begin{aligned}
y_t &= E_t y_{t+1} - \sigma(i_t - E_t\pi_{t+1}) + x_{dt} \\
i_t &= \frac{1}{\sigma}\left(E_t y_{t+1} - y_t\right) + \frac{1}{\sigma}x_{dt} + E_t\pi_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
\left[\begin{array}{c} i_t \end{array}\right] &= \left[\begin{array}{c} -\frac{1}{\sigma}\left(1-\rho_d\right)\left(1-\rho_d\beta\right) + \rho_d\gamma + \frac{1}{\sigma}\left(1-\rho_d\right)\left(1-\rho_d\beta\right) \\ +\left(1-\beta\rho_d\right)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right) + \gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right) \end{array}\right] z_{dt} \\
&+ \left[\begin{array}{c} \frac{1}{\sigma}\left(\rho_\pi - 1\right)\sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) \\ +\rho_\pi\left(\left(1-\rho_\pi\right)+\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1})\right) \end{array}\right] z_{\pi t} \\
&+ \left[\frac{1}{\sigma}\left(\rho_i - 1\right)\left(-\sigma\left(1-\rho_i\beta\right)\right) - \sigma\gamma\rho_i\right] z_{it} \\
&+ \left[\begin{array}{c} \left(\rho_y - 1\right)\left(-\gamma\rho_y + \left(1-\beta\rho_y\right)\phi_{y,0} + \rho_y\left(1-\beta\rho_y\right)\phi_{y,1} + \gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)\right) \\ -\gamma\rho_y\left(1-\rho_y\right) \end{array}\right] z_{yt}
\end{aligned}
$$

$$
\begin{aligned}
\left[\begin{array}{c} i_t \end{array}\right] &= \left[\begin{array}{c} \left(1-\beta\rho_d\right)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right) + \gamma\left(\phi_{\pi,0}+\rho_d\phi_{\pi,1}\right) \end{array}\right] z_{dt} \\
&+ \left[\left(1-\rho_\pi\right)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right) + \rho_\pi\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1})\right] z_{\pi t} \\
&+ \left[\left(1-\rho_i\right)\left(1-\rho_i\beta\right) - \sigma\gamma\rho_i\right] z_{it} \\
&- \left[\left(1-\rho_y\right)\left(\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right) + \left(1-\beta\rho_y\right)\phi_{y,0} + \rho_y\left(1-\beta\rho_y\right)\phi_{y,1}\right)\right] z_{yt}
\end{aligned}
$$

The overall solution is then

$$
\left[\begin{array}{c} y_t \\ \pi_t \\ i_t \end{array}\right] = A \left[\begin{array}{c} z_{dt} \\ z_{\pi t} \\ z_{it} \\ z_{yt} \end{array}\right]
$$

$$
\begin{aligned}
A_{:,1:2} &= \left[\begin{array}{cc} 1-\rho_d\beta & \sigma\left(\rho_\pi\left(1-\phi_{\pi,1}\right)-\phi_{\pi,0}\right) \\ \gamma & \left(1-\rho_\pi\right)+\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1}) \\ \left(1-\beta\rho_d\right)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\gamma\left(\phi_{\pi,0}+\rho_d\phi_{\pi,1}\right) & \left(1-\rho_\pi\right)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)+\rho_\pi\sigma(\phi_{y,0}+\rho_\pi\phi_{y,1}) \end{array}\right] \\
A_{:,3:4} &= \left[\begin{array}{cc} -\sigma\left(1-\rho_i\beta\right) & \sigma\left(-\gamma\rho_y+\left(1-\beta\rho_y\right)\phi_{y,0}+\rho_y\left(1-\beta\rho_y\right)\phi_{y,1}+\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)\right) \\ -\sigma\gamma & -\gamma\left(1-\rho_y\right) \\ \left(1-\rho_i\right)\left(1-\rho_i\beta\right)-\sigma\gamma\rho_i & -\left[\left(1-\rho_y\right)\left(\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)+\left(1-\beta\rho_y\right)\phi_{y,0}+\rho_y\left(1-\beta\rho_y\right)\phi_{y,1}\right)\right] \end{array}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
x_{dt} &= \left[\left(1-\rho_d\right)\left(1-\rho_d\beta\right)+\sigma\left(1-\beta\rho_d\right)\left(\phi_{y,0}+\rho_d\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_d\left(\phi_{\pi,1}-1\right)\right)\right] z_{dt} \\
x_{\pi t} &= \left[\left(1-\rho_\pi\right)\left(1-\rho_\pi\beta\right)+\sigma\left(1-\beta\rho_\pi\right)\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_\pi\left(\phi_{\pi,1}-1\right)\right)\right] z_{\pi t} \\
x_{it} &= \left[\left(1-\rho_i\right)\left(1-\rho_i\beta\right)+\sigma\left(1-\beta\rho_i\right)\left(\phi_{y,0}+\rho_i\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_i\left(\phi_{\pi,1}-1\right)\right)\right] z_{it} \\
\bar{y}_t &= \left[\left(1-\rho_y\right)\left(1-\rho_y\beta\right)+\sigma\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right)+\sigma\gamma\left(\phi_{\pi,0}+\rho_y\left(\phi_{\pi,1}-1\right)\right)\right] z_{yt}
\end{aligned}
$$