Identification with Taylor Rules: A Critical Review

John H. Cochrane*

First draft: April 2005. Current draft: September 12, 2007

Abstract

The parameters of the Taylor rule relating interest rates to inflation and other variables are not identified in new-Keynesian models. Thus, Taylor rule regressions cannot be used to argue that the Fed conquered inflation by moving from a "passive" to an "active" policy in the early 1980s.

*Graduate School of Business, University of Chicago and NBER. 5807 S. Woodlawn Ave. Chicago IL 60637. john.cochranechicagogsb.edu. 773 702 3059. I gratefully acknowledge research support from the CRSP and the NSF via a grant administered by the NBER. I am grateful for many comments and suggestions to Fernando Alvarez, Marco Bassetto, David Backus, Mark Gertler, Eric Leeper, Lars Hansen, Peter Ireland, Edward Nelson, Christopher Sims, Michael Woodford, and participants at the Spring 2005 Inidiana University Monetary Economics Conference and Fall 2006 NBER EFG Conference. Updates and Appendices are available on my webpage
http://faculty.chicagogsb.edu/john.cochrane/research/Papers/.

1 Introduction

The new-Keynesian Taylor-rule approach to monetary economics provides the current standard model of inflation determination, for modern fiat-money economies in which the central bank follows an interest rate target, ignoring monetary aggregates.

Any good theory relies on a stylized interpretation of important historical episodes. Keynes had the *General Theory* of the great depression. Friedman and Schwartz had the *Monetary History* of the US and UK. The central story for the new-Keynesian Taylor rule is that U. S. inflation was stabilized 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 did so. Most famously, Clarida, Galí and Gertler (2000) run regressions of interest rates on inflation and output. They find inflation coefficients below one up to 1980, and above one since then.

I argue against this central interpretation of the historical record. To see the key point, we need to understand how new-Keynesian models work. They *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. Instead, new-Keynesian models say that higher inflation would lead the Fed to *raise* future inflation. For only one value of inflation today will we fail to see inflation that either explodes or, more generally, eventually leaves a local region. Ruling out non-local equilibria, new-Keynesian modelers conclude that inflation today jumps to the unique value that leads to a locally-bounded equilibrium path.

In this logic, however, interest rates and inflation that increase in response to past inflation are a threat that is never realized in the observed equilibrium. The dynamics of inflation (and real variables) in the observed equilibrium can tell us *nothing* about the dynamics of other, unobserved equilibria. The crucial coefficients, including the parameters of the interest rate policy rule are not *identified.* The change in regression coefficients pre- and post- 1980 tells us *nothing* about determinacy. If you run an interest rate regression in artificial data generated by a new-Keynesian model, the coefficient you measure can be greater or less than one, and it contains no information about the true policy-rule coefficient.

Outline and central results

I start by establishing the lack of identification in new-Keynesian models. I first study a simple classic model, featuring a constant real interest rate and a policy rule $i_t = \bar{\imath} + \phi_\pi \pi_t + x_t$ where x_t is a monetary policy disturbance and $\phi_\pi > 1$. I show that the *estimated* policy rule in this circumstance recovers the stationary dynamics of the disturbance, not the explosive dynamics represented by ϕ_π, and I show that ϕ_π does not enter anywhere in the equilibrium dynamics of observable variables; ϕ_π is not identified. The central point of the new-Keynesian model is that endogenous variables such as π_t jump in response to disturbances such as x_t, in order to head off threatened explosions. Alas, this jump makes the right hand variable π_t inescapably correlated with the error term x_t in a regression.

I then study the standard three-equation new-Keynesian model. Following King (2000), I express the policy rule as $i_t = i_t^* + \phi_\pi (\pi_t - \pi_t^*)$ where i_t^* is the "natural" or "Wicksellian" rate of interest, a function of underlying disturbances to the economy, and π_t^* is inflation in the specified equilibrium. ϕ_π appears nowhere else, and since $\pi_t = \pi_t^*$ in equilibrium, the equilibrium clearly contains no information about ϕ_π. The right-hand variable $\pi_t - \pi_t^*$that must move to identify ϕ_π is always zero in equilibrium.

As the most general non-identification statement, I analyze a general linear model with some eigenvalues larger than one in absolute value and some less than one. I show that the eigenvalues larger than one can never be measured when we use the unique forward-looking locally-bounded equilibrium, and that equilibrium dynamics are the same as a particular equilibrium of the same model in which those eigenvalues are less than one.

This paper's title includes "review" and so does a substantial part of its body. Surely, famous modelers and empiricists have thought about these questions? They have, and carefully, so I must review their attempts to address these questions. Identification follows from assumptions, so all we can ever do is figure out what the assumptions are, and decide if they are plausible or not.

I start by examining assumptions made on the policy rules in order to identify their coefficients in regression-style analysis. These amount to standard assumptions to force right hand variables to be uncorrelated with error terms. In particular, the monetary policy disturbance, though serially correlated and correlated with right hand variables in theory, is assumed i.i.d. and independent of right hand variables, and the stochastic intercept, which in theory follows the "Wicksellian" or "natural" rate of interest, is assumed constant. Both assumptions are severe constraints on the sorts of models and equilibria specified by theory.

I next examine "full-system" estimates that estimate all of the parameters of the model, and check all of the model's predictions. These are obviously the most powerful methods with which one can address poor identification. A reading of the literature suggests many identification problems, and that no paper has even asked if parameters in the region of indeterminacy can generate observationally equivalent time series.

Since the models in the literature are complex and only analyzable by numerical methods, I study identification in a tractable fully-specified model. I find that structural parameters of the economy *can* be identified. However, policy rule parameter identification only rests on strong and implausible assumptions. One can identify a two-parameter rule – say, $i_t = \bar{\imath} + \phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + x_{it}$ – but allowing any more parameters destroys identification. If one allows responses to output as well as inflation, or if one allows the central bank to respond to shocks as well as endogenous variables, all identification is lost. In particular, there are parameters from the indeterminacy region that produce observationally equivalent dynamics to any set of parameters from the determinacy region. And even the minor identification successes are dependent on strong assumptions on the lag length (AR(1)) and number (three) of the disturbances in the model.

Lubik and Schorfheide (2004) suggest a clever way of identifying the central *determinacy* question without measuring particular parameters. I review their suggestion, and I show that it hinges crucially on the specification of the stochastic process for the unobserved disturbances. I prove that one can always construct a process for the disturbances that accounts for their observations with either a determinate ($\phi > 1$) or indeterminate ($\phi < 1$) model.

I examine Taylor's (1999) thoughts on how a Taylor rule should operate. This is an entirely old-Keynesian model, with no forward-looking terms. In this model $\phi_\pi > 1$ is a condition for *stable* roots and backward-looking solutions, not the other way around. The parameters of the policy rule can be identified under this model. Alas, the model lacks any microfoundations, and not even Taylor takes it very seriously.

Finally, I review related literature that points to related various identification problems in new-Keynesian models.

An acknowledgement

Of course, the point that behavior in other equilibria or out of equilibrium cannot be measured from data in a given equilibrium is both well known and 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." (He also states that the four principles are "not new.") Cochrane (1998) shows analogously that one cannot test the off-equilibrium government behavior that underlies the fiscal theory of the price level, as Canzoneri, Cumby, and Diba (2001) attempt; Ricardian and Non-Ricardian regimes also make observationally equivalent predictions for equilibrium time series. Identification in dynamic rational-expectations models is a huge literature, and most of the identification principles and pitfalls I apply and find here are well known. Sims (1980) is a classic statement. The devices I find are used to identify new-Keynesian models fit naturally into categories that he famously labeled as "incredible," in particular exclusion of variables and limitation on the stochastic process of unobserved disturbances. My contribution is only to apply these well-known principles to new-Keynesian models.

2 Identification

2.1 A simple model

We can see the main identification point in a very simple model consisting only of a Fisher equation 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, and x_t = random component to monetary policy. The coefficient ϕ measures how sensitive the central bank's interest rate target is to inflation. An "active" Taylor rule specifies $\phi > 1$. 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, and so forth, and it includes any Fed mismeasurement of potential output and structural disturbances. It is not a forecast error, so it is serially correlated,

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

The essential simplification of this model relative to the full new-Keynesian models that follow is that the real interest rate is constant and unrelated to other endogenous variables, output in particular.

We can solve this model by substituting out the nominal interest rate, leaving only inflation,

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

Following standard procedure in the new-Keynesian tradition, when $\phi > 1$ we solve this difference equation forward, restricting attention to the unique locally bounded (nonexplosive) solution, giving us

$$\pi_t = -\sum_{j=0}^{\infty} \frac{1}{\phi^{j+1}} E_t \left(x_{t+j} \right) = \frac{x_t}{\rho - \phi}. \tag{5}$$

Since π_t is proportional to x_t, the dynamics of equilibrium inflation are simply those of the disturbance x_t,

$$\pi_t = \rho \pi_{t-1} + w_t. \tag{6}$$

($w_t \equiv -\varepsilon_t / (\phi - \rho)$.)

Using (1) and the solution (6), we can find the equilibrium interest rate,

$$i_t = r + E_t \pi_{t+1} = r + \rho \pi_t. \tag{7}$$

There is no error term. Thus, *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 (5) shows that the right hand variable π_t and the error term x_t are correlated – perfectly correlated in fact.

Since the issue is correlation of right hand variables with errors, 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 instruments. For example, if we use all available lagged variables as instruments, we have from (6) and (7)

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

Thus the instrumental variables regression gives exactly the same estimate

$$E(i_t | \Omega_{t-1}) = r + \rho E(\pi_t | \Omega_{t-1}).$$

Is there nothing clever we can do? No. The equilibrium dynamics of the observable variables are given by (6) and (7),

$$\begin{aligned}\pi_t &= \rho\pi_{t-1} + w_t \\ i_t &= r + \rho\pi_t\end{aligned}$$

The equilibrium dynamics *do not involve* ϕ. They are *the same for every value of* ϕ. ϕ *is not identified from data on* $\{i_t, \pi_t\}$ *in the equilibrium of this model.* The likelihood function for $\{\pi_t, i_t\}$ does not involve ϕ. The only point of ϕ is to threaten hyperinflation in order to rule out equilibria. If the threat is successful in ruling out hyperinflation equilibria, it does not matter at all *how fast* the Fed-threatened hyperinflation would have come.

This discussion assumed $\|\phi\| > 1$, but there are also equilibria with $\|\phi\| < 1$ that generate exactly the dynamics (8)-(8), so we cannot even identify that the data are generated from the region $\|\phi\| > 1$. For any ϕ, this model has multiple locally-bounded equilibria given by

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

where δ_{t+1} is any random variable with $E_t\delta_{t+1} = 0$. We can think of $\{\pi_0,\ \delta_t\}$ indexing all the possible equilibria. If $\|\phi\| > 1$, the requirement of a locally-bounded equilibrium and the resulting forward-looking solution (5) amount to the particular choice of equilibrium

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

We can make the same choice of equilibrium in the $\phi < 1$ case. If we select this equilibrium, we have

$$\pi_1 = \phi\pi_0 + x_0 + \delta_1 = \phi\frac{x_0}{\rho - \phi} + x_0 + \frac{\varepsilon_1}{\rho - \phi} = \frac{\rho x_0 + \varepsilon_1}{\rho - \phi} = \frac{x_1}{\rho - \phi}.$$

Continuing, we have (5) again, with observable implications (8)-(8).

There is a general point at work here. Conventionally, we don't even try to solve new-Keynesian models when there are not enough unstable eigenvalues. But there is no reason not to do so: one may make the same equilibrium selection choice for $\phi < 1$ that one makes for $\phi > 1$. We are not forced to do so by the locally-bounded criterion, but it is still *a* valid equilibrium, and a useful one for checking whether parameters from the indeterminacy region can replicated the observable dynamics of a determinate set of parameters.

There are two key ingredients to the lack of identification: A model that only restricts expectations, and the fact that disturbances are not observable. As a result of equation (1), the economics of this model can only determine expected inflation, not ex-post inflation. The distinguishing feature of the microfounded new-Keynesian models reviewed below is similarly that an $E_t x_{t+1}$ appears in each equation where their old-Keynesian predecessors had an x_t. This feature leads to the multiple equilibria. For an example of the second point, consider the standard model of asset pricing with a risk-neutral investor and hence a constant expected return R. To make the analogy even closer, suppose dividends are known one period ahead of time, and follow an $AR(1)$. Then we have

$$\begin{aligned} d_t &= \rho d_{t-1} + \varepsilon_t \\ E_t\left(p_{t+1} + d_t\right) &= R p_t \end{aligned}$$

Since $R > 1$, we solve the second equation forward to obtain

$$p_t = E_t \sum_{j=0}^{\infty} \frac{1}{R^{j+1}} d_{t+j} = \frac{1}{R-\rho} d_t$$

The model and solution are exactly the same as above, with $d_t = -x_t$, $p_t = \pi_t$, $R = \phi$. Yet in this model R *is* identified: We can simply regress $p_{t+1} + d_{t+1}$ on p_t or take the sample average return. The difference is that d_t is directly observable in this case, where x_t is not. Were dividends unobservable, we could not measure the average return, and in particular we could not tell if it were greater than or less than one.

Loisel (2007) and Adão, Correia and Teles (2007) study rules that exactly offset the expected future values of the economic model, in order to eliminate multiple equilibria. Loisel proposes (simplified to this setting)

$$i_t = r + E_t \pi_{t+1} + \psi\left(\pi_t - z_t\right) \tag{8}$$

where ψ is any nonzero constant, and z_t is any exogenous random variable. If we merge this rule with the Fisher equation (1), we obtain a unique equilibrium

$$\pi_t = z_t. \tag{9}$$

As I show in Cochrane (2007), this sort of rule is a limit of the usual sort of rule, in which the forward-looking eigenvalues are driven to infinity. For the current point, Equation (9) makes it clear that time-series from the equilibrium of this model cannot be used to check whether the coefficient of the rule (8) on future inflation is in fact one, or to measure the coefficient ψ on current inflation.

2.2 Identification in new-Keynesian models

One may well object at the whole idea of studying identification 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. 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 in detail and seeing, at the cost of some algebra, that the same points emerge.

The excellent exposition in King (2000) makes the non-identification theorem clear. I use that structure here. The basic model is

$$y_t = E_t y_{t+1} - \sigma\left(r_t - r\right) + x_{dt} \tag{10}$$

$$i_t = r_t + E_t \pi_{t+1} \tag{11}$$

$$\pi_t = \beta E_t \pi_{t+1} + \gamma\left(y_t - \bar{y}_t\right) + x_{\pi t} \tag{12}$$

where y denotes output, r denotes the real interest rate, i denotes the nominal interest rate, π denotes inflation, $\bar{y}$ is exogenously-varying potential output, and the x are disturbances. The disturbances x_{dt} and $x_{\pi t}$ can be serially correlated. I use a Roman letter (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 Clarida, Galí and Gertler (1999), King (2000), 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 optimizing firms that set prices subject to adjustment costs.

For the identification question, we can simplify the analysis by studying *deviations* from a given equilibrium rather than the equilibrium itself, following King (2000). For this purpose, it's convenient to solve the model backwards: find the interest rate that supports any equilibrium output process rather than (as usual) find equilibrium output for a given interest rate rule. Start with an equilibrium process for output $\{y_t^*\}$. The "neutral" or "no gap" equilibrium $y_t^* = \bar{y}_t$ is particularly nice, but I use the general case to emphasize that results do not depend on this choice. From (12) we can find the required path for equilibrium inflation π_t^*; from (10) we can find the required path for the equilibrium real rate r_t^*, and then from (11) we can find the required equilibrium nominal interest rate i_t^* :

$$\begin{aligned} \pi_t^* &= E_t \sum_{j=0}^{\infty} \beta^j \left[\gamma(y_{t+j}^* - \bar{y}_{t+j}) + x_{\pi t+j} \right] && (13) \\ r_t^* &= r + \frac{1}{\sigma} \left(E_t y_{t+1}^* - y_t^* \right) + \frac{1}{\sigma} x_{dt} && (14) \\ i_t^* &= r_t^* + E_t \pi_{t+1}^*. && (15) \end{aligned}$$

Putting it all together,

$$i_t^* = r + \frac{1}{\sigma} \left(E_t y_{t+1}^* - y_t^* \right) + \frac{1}{\sigma} x_{dt} + E_t \sum_{j=0}^{\infty} \beta^j \left[\gamma(y_{t+j+1}^* - \bar{y}_{t+j+1}) + x_{\pi t+j+1} \right]. \quad (16)$$

In particular, the no-gap equilibrium $y_t = \bar{y}_t$ is achieved with

$$i_t^* = r + \frac{1}{\sigma} \left(E_t \bar{y}_{t+1} - \bar{y}_t \right) + \frac{1}{\sigma} x_{dt} + E_t \sum_{j=0}^{\infty} \beta^j x_{\pi t+j+1} \quad (17)$$

Using tildes to denote deviations from the $*$ equilibrium, $\tilde{y}_t = y_t - y_t^*$, we can subtract the values of (10)-(12) from those of the $*$ equilibrium to describe *deviations* from that equilibrium as

$$\tilde{\imath}_t = \tilde{r}_t + E_t \tilde{\pi}_{t+1} \quad (18)$$

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \sigma \tilde{r}_t \quad (19)$$

$$\tilde{\pi}_t = \beta E_t \tilde{\pi}_{t+1} + \gamma \tilde{y}_t. \tag{20}$$

This is the same model, but without constants or disturbances.

Now it is clearly true that 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 setting $i_t = i_t^*$ does not determine that this is the *only* equilibrium. Equations (10)-(12) only determine expectations $E_t\tilde{y}_{t+1}, E_t\tilde{\pi}_{t+1}$, so arbitrary shocks to output and inflation can occur so long as they are not predictable. Furthermore, the dynamics of (10)-(12) are stable, so these multiple equilibria all stay near the steady state.

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

$$i_t = i_t^* + \phi_0 \left(\pi_t - \pi_t^*\right) + \phi_1 \left(E_t \pi_{t+1} - E_t \pi_{t+1}^*\right) \tag{21}$$

or, more simply,

$$i_t = \phi_0 \tilde{\pi}_t + \phi_1 E_t \tilde{\pi}_{t+1}.$$

For example, with $\phi_1 = 0$ the deviations from the $*$ equilibrium follow[1]

$$\begin{bmatrix} E_t\tilde{y}_{t+1} \\ E_t\tilde{\pi}_{t+1} \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} \beta+\sigma\gamma & -\sigma\left(1-\beta\phi_0\right) \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix}. \tag{22}$$

The eigenvalues of this transition matrix are

$$\lambda = \frac{1}{2\beta}\left((1+\beta+\sigma\gamma) \pm \sqrt{(1+\beta+\sigma\gamma)^2 - 4\beta\left(1+\sigma\gamma\phi_0\right)}\right). \tag{23}$$

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

$$\phi_0 > 1$$

or if

$$\phi_0 < -\left(1 + 2\frac{1+\beta}{\sigma\gamma}\right). \tag{24}$$

[1] Use (18) to eliminate $\tilde{r}$, and use $\tilde{\imath}_t = \phi_0 \tilde{\pi}_t$ giving

$$\tilde{y}_t = E_t\tilde{y}_{t+1} - \sigma(\phi_0\tilde{\pi}_t - E_t\tilde{\pi}_{t+1}).$$

In matrix form, then, we have

$$\begin{bmatrix} 1 & \sigma \\ 0 & \beta \end{bmatrix}\begin{bmatrix} E_t\tilde{y}_{t+1} \\ E_t\tilde{\pi}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & \sigma\phi_0 \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix}$$

$$\begin{bmatrix} E_t\tilde{y}_{t+1} \\ E_t\tilde{\pi}_{t+1} \end{bmatrix} = \frac{1}{\beta}\begin{bmatrix} \beta & -\sigma \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \sigma\phi_0 \\ -\gamma & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_t \\ \tilde{\pi}_t \end{bmatrix}.$$

Equation (22) follows.

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 equilibrium.

Now, I can use King's expression of the Taylor rule (21) to make the central identification point. *In the * equilibrium*, we will always see $\pi_t - \pi_t^* = 0$. Thus, a regression estimate of (21) cannot possibly estimate ϕ_0, ϕ_1. There is no movement in the necessary right hand variables. More generally, ϕ_0 and ϕ_1 appear nowhere in the *equilibrium* dynamics characterized simply by $\tilde{\pi} = \tilde{y} = \tilde{\imath} = 0$, so they are *not identified*. The same dynamics of the * equilibrium hold for *any* values of ϕ_0 and ϕ_1, which is another way of saying "not identified." Taylor determinacy depends entirely on what the Fed would do *out* of the * equilibrium, which we can never see from data *in* that equilibrium.

King recognizes the problem, writing in footnote 41,

> "The specification of this rule leads to a subtle shift in the interpretation of the policy parameters $[\phi_0, \phi_1]$; 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."

King does not address the implications of this non-identification for empirical work.

The desired, * equilibrium of the new-Keynesian model does contain relations between interest rates, output, and inflation. For example, we can write from (13)-(15)

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

and I investigate below further substitutions to eliminate x_{dt}. The problem is, this relation, together with the other equations of the model, implies stable dynamics and hence multiple equilibria. To force unstable dynamics on the system, in order to rule out the multiple equilibria, we have to imagine that the central bank would respond more strongly to *deviations* of inflation or output from this equilibrium than it does to movements in inflation and output *in* the equilibrium, as expressed in (21).

It is this final step that causes all the trouble. Since we never observe $\pi_t \neq \pi_t^*$, it raises my econometric point that such deviations are not identified. It raises the corresponding theoretical points that the Fed could not signal its ϕ behavior by any visible action, and that agents in the economy would be hard-pressed to learn such coefficients as well.

2.3 Regressions in new-Keynesian model output

What happens if you run interest-rate regressions in artificial data from a new-Keynesian model? We know the answer for the simple model given above, and we have a theorem that the result will not measure determinacy. Still, it would be interesting to know the answer in the standard three-equation new-Keynesian model. If not ϕ, and viewed through the model, what *did* Clarida, Galí and Gertler (2000) measure?

In general, the answer is a) not ϕ and b) a huge mountain of algebra. While easy enough to evaluate numerically, such answers don't give much intuition. For some special cases, though, I can find intelligible and interesting algebraic formulas. I present the algebra in the Appendix.

Suppose the central bank follows, and we estimate, a rule of the form

$$\begin{aligned} i_t &= \phi_0 \pi_t + x_{it} \\ x_{it} &= \rho_i x_{it-1}, \end{aligned}$$

and the economy follows the standard three-equation model (10)-(12). When this is the only disturbance to the system (no disturbances x_{dt}, $x_{\pi t}$ and no variation in the natural rate $\bar{y}_t$), the *estimated* coefficient in a regression of i_t on π_t is

$$\hat{\phi}_0 = \rho_i + \frac{(1-\rho_i)(1-\rho_i\beta)}{\sigma\gamma} \tag{25}$$

First, note that ϕ_0 appears nowhere in the right hand side. Second, we start with the autocorrelation of the monetary policy disturbance ρ_i, as in the simple case studied above. Third, there are now extra terms, involving the parameters of the other equations of the model, in particular the intertemporal substitution elasticity σ and the Phillips coefficient γ. The parameters σ and γ can take on small values, so $\hat{\phi}_0$ can be greater than one.

Suppose instead that the central bank follows, and we estimate, a rule of the form

$$i_t = \phi_1 E_t \pi_{t+1} + x_{it},$$

i.e. reacting to expected future inflation. In the same case of the three-equation model, the *estimated* coefficient is

$$\hat{\phi}_1 = 1 + \frac{(1-\rho_i)(1-\rho_i\beta)}{\sigma\gamma\rho_i}. \tag{26}$$

Once again, ϕ_1 is absent from the right hand side. In this case, we *generically* find a coefficient greater than one, so long as parameters obey their usual signs, though this finding has *nothing* to do with the actual Taylor rule.

This observation solves a puzzle of the simple example: How can Clarida, Galí and Gertler (2000) and others find coefficients greater than one? In my simple example, in which $\hat{\phi}_0$ estimated the autocorrelation of the policy disturbance ρ_i, this was not possible. One might suspect that they had measured *something* interesting by finding a coefficient greater than one. Here, we see that estimated coefficients greater than one are perfectly possible, even in the context of the new-Keynesian model. And they are perfectly uninformative about the true ϕ.

2.4 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

$$\mathbf{y}_{t+1} = \mathbf{A}\mathbf{y}_t + \mathbf{C}\boldsymbol{\varepsilon}_{t+1} \tag{27}$$

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[2] 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,

$$\boldsymbol{\Lambda} = \left[\begin{array}{ccc} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{array}\right],$$

and $\mathbf{Q}$ is the corresponding matrix of eigenvectors.

Premultiplying (27) 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

$$\|\lambda_i\| > 1 : \; z_{it} = \sum_{j=1}^{\infty} \frac{1}{\lambda_i^j} E_t \xi_{t+j}^i = 0 \tag{28}$$

$$\|\lambda_i\| < 1 : \; z_{it} = \sum_{j=0}^{\infty} \lambda_i^j \xi_{t-j}^i \tag{29}$$

$$z_{it} = \lambda_i z_{it-1} + \xi_{it}. \tag{30}$$

Thus, we choose the unique locally-bounded equilibrium by setting the forward-looking 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 (29), denote by $\boldsymbol{\xi}_t^*$ the corresponding shocks, denote by $\boldsymbol{\Lambda}^*$ the

[2] 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.

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

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

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 (28) 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 $\mathbf{\Lambda}$ greater than one and no shocks by the new-Keynesian equilibrium selection criterion is observationally equivalent to the same no-shock equilibrium with $\mathbf{\Lambda}$ less than one.

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

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

which describes the 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. 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 central question for determinacy is whether eigenvalues of the *system* are greater than one. These are, in general, properties of the system as a whole, not just properties of a single parameter, set of parameters, or structural equation. For example, the lower boundary (24) in the simple model depends on structural parameters β, σ, γ as well as the policy rule parameter ϕ. Thus, even if it is possible to correctly identify the coefficients of the policy rule, we expect that we will not be able in general to identify other parameters (σ, γ), which control "determinacy" or "indeterminacy."

2.5 Hopes for identification

How can we get around these non-identification results? How *do* new-Keynesian modelers get around the results? How do their computer programs produce estimates, and presumably with non-singular Hessian matrices? Obviously, we must add some assumptions or restrictions on the models.

First, we can think about restricting the policy rule in order to remove the troublesome correlation of right hand variable with error terms, or at least to produce some valid instruments. More deeply, the central problem in (21) is that the parameters ϕ governing the response of interest rates to deviations from the equilibrium can't be seen since we never observe deviations from that equilibrium. Obviously, we can try restricting the policy rule so that this response to *deviations* from the * equilibrium are tied to some aspect of behavior *in* that equilibrium.

Second, I proved that the unstable *eigenvalues* in (31)-(31) cannot be identified from equilibrium time series. However, eigenvalues are not directly measured. Eigenvalues are functions of the underlying parameters, i.e. $\lambda(\beta, \gamma, \sigma, \phi_0, \phi_1)$. Perhaps the structure of the model links the stable eigenvalues to the unstable eigenvalues. Perhaps we we can identify the underlying structural parameters from the stable eigenvalues, and then *infer* the value of the unstable eigenvalues, or at least infer that they are greater than one in absolute value.

With these thoughts in mind, I both review the (often implicit) identification assumptions that new-Keynesian modelers have made, and I explore alternative assumptions one might make.

There are two basic approaches to identification and estimation. First, one can follow a single-equation approach, as in Clarida, Galí, and Gertler (2000): measure the Fed policy rule, without trying to measure the structural parameters (β, γ, σ) of the rest of the economy. Second, one can specify and estimate the whole model, and therefore exploit cross-equation restrictions that may give better identification. I consider each approach in turn.

3 Policy-rule identification

3.1 Policy rules in theory

Identification is a property of *models*, not a property of data. Therefore, before examining empirical specifications, we should examine what sort of policy rules are specified by new-Keynesian models. Only then can we understand the plausibility of restrictions that are imposed in actual empirical work. The crucial specifications I highlight are 1) the presence of a serially correlated stochastic intercept, which is a function of real disturbances to the economy 2) the presence of all endogenous variables on the right hand side and 3) the likely serial and cross-correlation of the monetary policy disturbance.

The stochastic intercept

Start with the simple Taylor rule (21),which I repeat here

$$i_t = i_t^* + \phi_0 \left(\pi_t - \pi_t^*\right) + \phi_1 \left(E_t \pi_{t+1} - \pi_{t+1}^*\right). \tag{31}$$

The first central ingredient of this rule is the time-varying intercept i_t^*. In this expression of the Taylor rule, i_t^* is the interest rate of the desired, $*$ equilibrium. It is a function of the underlying disturbances to the economy, as given in (17) and (16).

Many theoretical treatments and most empirical work are written with Taylor rule parameters that depend on actual variables, e.g. $\phi_0 \pi_t$, rather than deviations from equilibrium, e.g. $\phi_0 (\pi_t - \pi_t^*)$. This is not really a change in specification, it just means folding the π^* terms into a larger time-varying intercept, i.e.,

$$i_t = \left(i_t^* - \phi_0 \pi_t^* - \phi_1 E_t \pi_{t+1}^*\right) + \phi_0 \pi_t + \phi_1 E_t \pi_{t+1} \tag{32}$$

Since equilibrium i^*, π^* are functions of underlying disturbances, we can still understand the model Taylor rule as one with a time-varying intercept.

Woodford (2003) emphasizes the need for such a shifting intercept. The foundation of optimal "Wicksellian" policy is to allow the interest rate target to shift up and down following the "natural" rate of interest, determined by real disturbances to the economy. Equation (17) is a simple example of the time-varying intercept needed to achieve the no-gap equilibrium $y_t^* = \bar{y}_t$.

A stochastic intercept of the form (32) poses a serious challenge to empirical work. Very few estimates include the stochastic intercept, especially in single-equation approaches for which measuring disturbances to the other equations is impossible. The "stochastic intercept" is therefore a component of the error term. Since it is a function of the underlying economic disturbances, it is serially correlated, and it is correlated with the right hand variables (output, inflation, etc.), since equilibrium values of the latter are functions of the same disturbances.

The absence of a stochastic intercept is the same thing as the restriction that the central bank's response to inflation and output *in* equilibrium is the same as its response to inflation and output in alternative equilibria. This is easy to see just by writing

$$i_t = \bar{\imath} + \phi_0 \pi_t = \bar{\imath} + \phi_0 \pi_t^* + \phi_0 (\pi_t - \pi_t^*).$$

For this reason, it is a very appealing assumption. It gets right to the heart of the non-identification problem. It also is sensible that the Fed might want to enhance its credibility by tying its alternative-equilibrium threats to in-equilibrium behavior. With this kind of rule we can imagine agents learning about the Fed's responses by seeing data on an equilibrium. It's also clear how this assumption helps identification. If we solve models with this kind of policy rule, equilibrium output, inflation and interest rates vary as functions of current and expected future values of the disturbances $x_{dt}, x_{\pi t}, \bar{y}_t$, discounted by the forward-looking eigenvalues (e.g., Equation (23)) of the transition matrix. Thus, it's straightforward to see that equilibrium dynamics are affected by the choice of parameter ϕ_0. Finally, rules that must be written with shocks raise the question, how can the Fed measure and respond to shocks that we can't recover from endogenous variables?

Do we really need a stochastic intercept?

How harmful is it to assume away the stochastic intercept? The answer is that policy rules without a stochastic intercept can only generate a restricted set of equilibria, and policy rules that respond to shocks are a central feature of most interesting new-Keynesian equilibria. In some sense we knew this from reading: Clarida, Galí and Gertler (2000) and Woodford (2003, Ch4) calculate the variance of output and inflation using rules with no intercepts, and discuss the merits of larger ϕ for reducing such variance, while all the time equilibria with zero variance of output or inflation are sitting under our noses, as in Woodford's "Wicksellian" policies and the no-gap equilibrium $y_t - \bar{y}_t$ above, if only we will allow the policy rule to depend on disturbances directly.

Why can't the same equilibrium be achieved with a rule that responds only to variables and not to disturbances? After all, endogenous variables are functions of disturbances (that's the whole point of finding an equilibrium), so why can't we substitute out for disturbances in terms of endogenous variables and get rid of the stochastic intercept? The central problem in making this substitution is that the relationships between interest rates, output, and inflation *in* typical equilibria of the new-Keynesian model do not lie in the determinacy region. In order to generate these equilibria as unique locally-bounded equilibria, we are forced, as above, to imagine that the central bank responds more strongly to inflation or output that *deviates* from the equilibrium values than it responds to variation in equilibrium output or interest rates, which is equivalent to a stochastic intercept.

As the simplest example, suppose our equilibrium results in a constant nominal interest rate, $i_t^* = \bar{\imath}$, which means

$$i_t^* = \bar{\imath} + 0 \times \pi_t^*.$$

Obviously, the latter expression, $\phi_\pi = 0$, does not lie in the region of determinacy. To generate this equilibrium as a unique locally-determinate outcome, we have to write the policy rule with a stochastic intercept

$$i_t = (\bar{\imath} + \phi_\pi \pi_t^*) + \phi_\pi (\pi_t - \pi_t^*).$$

In the simple model of Section 2.1, as in any model with constant real interest rates, the Fisher relation

$$i_t^* = r + E_t \pi_{t+1}^* \tag{33}$$

means that we *always* see in equilibrium a coefficient of exactly one on expected future inflation. Again, to generate any equilibrium in such a model as the unique locally-bounded equilibrium, we will have to add a stochastic intercept,

$$i_t^* = r + E_t \pi_{t+1}^* + \phi_1 \left(E_t \pi_{t+1} - E_t \pi_{t+1}^* \right).$$

In the three-equation model, (14)-(15) give

$$i_t^* = r + \frac{1}{\sigma} \left(E_t y_{t+1}^* - y_t^* \right) + E_t \pi_{t+1}^* + \frac{1}{\sigma} x_{dt} \tag{34}$$

If there is no disturbance x_{dt}, we once again cannot write the equilibrium without a stochastic intercept. Coefficients $\phi_{\pi,1} = 1$, $\phi_{y,0} = -\phi_{y,1}$ also generate an eigenvalue of exactly one (see (47) below). Our only hope is to include x_{dt} and hope that when we substitute out for x_{dt} in terms of observables y and π, that the resulting relation lies in the determinacy region.

Now, equilibria y_t^* are in general driven by current and expected future values of the disturbances,

$$y_t^* = E_t \left[a(F)x_{dt} + b(F)x_{\pi t} + c(F)\bar{y}_t \right]; \quad Fx_t \equiv x_{t+1} \tag{35}$$

To keep the algebra manageable, I consider a simple example,

$$y_t^* = a\bar{y}_t + bx_{\pi t} + cx_{dt}. \tag{36}$$

Can we generate this equilibrium from an interest rate rule with a constant intercept and no loading on disturbances? Inflation is, substituting (36) into (12),

$$\pi_t^* - \beta E_t \pi_{t+1}^* = \gamma (a-1)\bar{y}_t + (1+\gamma b)\, x_{\pi t} + c\gamma x_{dt} \tag{37}$$

The real interest rate is not observable, and our objective is to substitute out the disturbances in the nominal interest rate equation. Thus, we have to find x_{dt} for equation (34) in terms of observables y_t and π_t from (36) and (37).

Alas, this cannot be done: we have two equations and three disturbances. Thus, even this simple equilibrium of the standard three-equation model can't be implemented without a stochastic intercept. More generally, the plan to recover disturbances from observable variables cannot always be accomplished; there may be too many disturbances or the functional forms may not be invertible.

To proceed, we will have to restrict the model some more. I assume $\bar{y}_t$ is constant through time and $a = 1$. (More generally, many authors use estimates of "potential output" to treat $\bar{y}_t$ as observable.) I also set $b = 0$, as adding $b \neq 0$ does not change the character of the example[3]. Now the example has specialized to

$$y_t^* - \bar{y} = cx_{dt} \tag{38}$$

$$\pi_t^* - \beta E_t \pi_{t+1}^* = x_{\pi t} + c\gamma x_{dt} \tag{39}$$

We *can* at last solve (38)-(39) for the disturbances $x_{\pi t}$ and x_{dt},

$$cx_{dt} = (y_t^* - \bar{y}).$$

Substituting into the interest rate equation (34),

$$i_t^* = r + \frac{1}{\sigma c}(y_t^* - \bar{y}) + \frac{1}{\sigma}\left(E_t y_{t+1}^* - y_t^*\right) + E_t \pi_{t+1}^*. \tag{40}$$

[3]With $b \neq 0$, we obtain

$$cx_{dt} = (1+\gamma b)(y_t^* - \bar{y}) - b\left(\pi_t^* - \beta E_t \pi_{t+1}^*\right).$$

$$i_t^* = r + \frac{1+\gamma b}{\sigma c}(y_t^* - \bar{y}) + \frac{1}{\sigma}\left(E_t y_{t+1}^* - y_t^*\right) + E_t \pi_{t+1}^* - \frac{b}{\sigma c}\left(\pi_t^* - \beta E_t \pi_{t+1}^*\right)$$

and the determinacy condition is

$$1 - \frac{(1-\beta)(1+2b\gamma)}{\sigma c \gamma} > 1,$$

which again can hold or not depending on b and c.

Now, we have it, a policy rule in standard form without a shifting intercept. (If we start with the more general (35), we end up with many leads and lags of output and inflation, not just the conventional t and $t+1$ terms.) The remaining question is, are these coefficients in the region of determinacy? Evaluating the basic determinacy condition, shown in (47) below, we need

$$1 - \frac{(1-\beta)}{\gamma}\frac{1}{\sigma c} > 1.$$

This condition can go either way. With the usual signs $\beta \in (0,1)$, $\gamma > 0$, $\sigma > 0$, if a value of $c > 0$ means the condition is violated, then $c < 0$ means that it holds. So, the point is made: some equilibria can be supported by a locally-deterministic Taylor rule with a nonstochastic intercept, i.e., with no direct loading on disturbances, and some cannot. The assumption of a fixed intercept, or that the interest rate rule away from equilibrium is the same as that in equilibrium, restricts the set of achievable equilibria.

Granted that we are ruling out equilibria, are we ruling out interesting equilibria? This is not a question I can answer in great generality, but many interesting equilibria of this model cannot be achieved as unique locally determinate outcomes without a stochastic intercept. The $y_t = \bar{y}_t$ "no-gap" equilibrium is one. If just before (38) we examine instead the special case

$$y_t^* = \bar{y}_t$$

then (37) becomes

$$\pi_t^* - \beta E_t \pi_{t+1}^* = b x_{\pi t}.$$

This system does not allow us to back out x_{dt} in (34). (We can also view $y_t = \bar{y}_t$ as a limiting case that requires infinite ϕ parameters.) The equilibrium with constant inflation is another example. From (12) such an equilibrium can be achieved with

$$y_t = \bar{y}_t - \frac{1}{\gamma} x_{\pi t}.$$

Equation (37) now just reads

$$\pi_t^* - \beta E_t \pi_{t+1}^* = 0$$

so once again we cannot find x_{dt} for in (34) from observables.

More generally, Woodford's (2003) discussion of optimal policies leads to the "neo-Wicksellian" conclusion that policy must respond directly to disturbances, i.e. with a stochastic intercept. When we rule out such policies, we are in effect assuming that the central bank is not following optimal policies of this sort.

To empirical policy rules

Models also allow the central bank to respond to output, and to every other endogenous variable. This helps to connect with the data – interest rate regressions seem to have output coefficients – and responses to additional variables can bring us closer to optimal policies that respond to structural disturbances, as above. Both theory and empirical work also favor interest rate smoothing, which is another way of saying that lags of output and inflation should be in the rule.

These thoughts leave us with specifications of the form

$$\begin{aligned} i_t &= \bar{\imath}_t + \rho_i i_{t-1} + ...\phi_{\pi,-1}\pi_{t-1} + \phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + ... \\ &\quad + \phi_{y,-1}y_{t-1} + \phi_{y,0}y_t + \phi_{y,1}E_t y_{t+1} + ... + \ x_{it} \end{aligned} \tag{41}$$

where I have used ... to indicate that more leads and lags may be added. Both theory and empirical work allow all endogenous variables to enter the policy rule. Optimal policy of course responds, in general, to everything available.

The error term includes genuine monetary policy disturbances as well as the left-out disturbances from the stochastic intercept. We certainly want to allow for such monetary policy disturbances, just as we allow for disturbances to preferences and technology of the other agents, if for no other reason than theories that predict 100% R^2 can quickly be rejected.

Monetary policy disturbances most likely represent left-out variables. At a fundamental level, the Federal reserve never announces "and then we flipped a coin and added 50 basis points for the fun of it." It always explains policy as a response to something – perhaps a momentary concern with exchange rates, or (as I write) with conditions in credit markets. This source of disturbances is at least serially correlated, and also likely to be correlated with the right hand variables such as inflation and output. Time-varying rules are often cited as a source of error, but they really are the same thing, since some variable caused the rule to shift. And even a genuine, exogenously-changing rule will lead to serially correlated error. Rudebush (2005) examines Fed Funds futures data to argue that persistence in Taylor rule shocks is due to persistent omitted variables, and not to partial adjustment to shocks.

Summary

In sum, new-Keynesian *models* specify policy rules that are a snake-pit for econometricians. The right hand variables are endogenous. As in the simple model above in which π_t jumps in response to the error x_t, many of the right hand variables are predicted to jump when there is an innovation to the error term, generating strong correlations. There is a time-varying intercept, which will be hard to distinguish from an error term, and which is also correlated with right hand variables since both are functions of underlying disturbances in equilibrium. The time-varying intercept and the true monetary policy disturbance x_{it} are serially correlated, making the use of lags as instruments invalid.

3.2 Clardia, Galí and Gertler

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

$$i_t = (1 - \rho_1 - \rho_2)\left\{r + (\phi_\pi - 1)\left[E_t(\pi_{t+1}) - \pi\right] + \phi_y E_t\left[\Delta y_{t+1}\right]\right\} + \rho_1 i_{t-1} + \rho_2 i_{t-2} \tag{42}$$

where

$$\begin{aligned} \pi &= \text{inflation target} \\ \Delta y_{t+1} &= \text{growth in output gap} \\ r &= \text{“long run equilibrium real rate”} \end{aligned}$$

(See (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.

An error term appears if we replace expected inflation and output with their ex-post realized values, writing

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

In this way, Clarida, Galí, and Gertler do not make the prediction of 100% R^2 that would normally come from assuming away the policy 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 estimate the model by instrumental variables observed at time t to remove correlation between ε_{t+1} and the ex-post values of the right hand variables π_{t+1} and Δy_{t+1}.

Second, the time-varying intercept is gone as well. Such an intercept would also look like a serially correlated error term, correlated with right hand variables in this single-equation estimation. As above, the absence of a time-varying intercept is a strong identifying restriction, with strong implications for model dynamics.

Clarida, Galí and Gertler (1998) consider a slightly more general specification that does include a monetary policy disturbance[4]. In this case, they specify (their equation 2.5, my notation)

$$i_t = (1 - \rho) \left[\alpha + \phi_\pi E (\pi_{t,t+n} | \Omega_t) + \phi_y E (y_t - \bar{y}_t | \Omega_t) \right] + \rho i_{t-1} + v_t \quad (44)$$

where $\bar{y}$ denotes potential output, separately measured, and Ω_t is the central bank's information set at time 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 (44) by instrumental variables, using lagged output, inflation, interest rates, and commodity prices as instruments.

The assumption that the disturbance v_t is unpredictable from any variable in the central bank's information set is the key to identification in this case. Clarida Galí and Gertler assume the central bank does not observe current output y_t. *Actual* output y_t jumps when v_t is revealed. In the new-Keynesian model, the error term represents the time-varying intercept and the other variables that the Fed may look at, so of course assuming an i.i.d. error is a very strong restriction.

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

3.3 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 regressions. Giannoni and Woodford (2005 p. 36-37) nicely lay out 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
>
> $$i_t = \bar{\imath} + \sum_{k=1}^{n} \phi_{ik}(i_{t-k} - \bar{\imath}) + \sum_{k=0}^{n_w} \phi_{wk}\hat{w}_{t-k} + \sum_{k=0}^{n_w} \phi_{\pi k}(\pi_{t-k} - \bar{\pi}) + \sum_{k=0}^{n_y} \phi_{yk}\hat{Y}_{t-k} + \varepsilon_t \quad (45)$$
>
> where i_t is the Federal funds rate in period t; ϕ_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 [(45)], 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 [(45)] 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 a stochastic intercept, responding to structural disturbances in the economy, and other endogenous variables that the Fed may respond to.

Second, Giannoni, Rotemberg, and Woodford 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 a common *empirical* assumption, but it is an especially surprising result of a new-Keynesian *model*, because in $\hat{w}_t, \pi_t$ $\hat{Y}_t$ are endogenous variables. From the very simplest model in this paper, endogenous variables have jumped when there is a monetary policy (or any other) disturbance. The whole *point* of new-Keynesian, forward-looking solutions is that endogenous variables jump, so that disturbances do not lead to hyperinflations. 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 the 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. These additional lags help to produce realistic dynamics in the predicted impulse-response functions.) Needless to say, while logically consistent, this is an even stronger assumption. Are wages, prices, and output really fixed one to two quarters in advance in our economy, and therefore unable to react within the quarter to monetary policy disturbances?

Finally, 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 does jump in this model are expectations of future values of these variables, among others $\hat{w}_{t+1} = E_t\hat{w}_{t+1}, \pi_{t+1} = E_t\pi_{t+1}$, and $\hat{Y}_{t+2} = E_t\hat{Y}_{t+2}$ as well as the state variable $E_t\mu_{t+1}$ (marginal utility of consumption). All of these variables are determined at date t. Now, we see another assumption in the policy function (45) – none of these future variables are present in the policy rule. In contrast to the vast 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, that reaction is absent here.

In sum, Giannoni and Woodford identify the Taylor rule in their model, by virtue of classic a-priori identifying assumptions. There are two assumptions about Fed behavior: 1) The disturbance and time-varying component of the intercept are not serially correlated, or predictable by any variables at time $t-1$, and 2) The Fed does not react to expected future output, or wage, price inflation, or other state variables. There is an assumption about the economy: 3) Wages, prices, and output are fixed a period in advance.

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 the above 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. Exploiting the full predictions of the model offers greater hope for identification. It also offers the only way to really check determinacy. The eigenvalues of new-Keynesian models are almost always complex functions of many structural parameters, not just those of the policy rule. There is now an exploding literature on estimating fully-specified new-Keynesian models. I review the major contributions, and then I study the identification possibilities in a three-equation model.

4.1 Flavors of identification

Identification comes in several flavors, and it's worth being explicit about them so we know what we're looking for. Denote the set of parameters by θ. This includes structural parameters (β, γ, σ), policy-rule parameters (ϕ) and parameters describing the evolution of disturbances (ρ). There is also a set of observables, which I denote B. In the model I will study below, the observable implications of the model are completely summarized by the VAR transition matrix for output, inflation, and interest rates, so B represents this matrix.

1. *General identification.* The strongest possible form of identification occurs if there is a one-to-one mapping between θ and B. Then you know that no matter how the empirical exercise turns out (estimation of B), you are guaranteed to know the structural

parameters uniquely. Often, identification occurs not completely in general but except for measure zero sets of parameters. For example, most of the identification I present below relies on two disturbances having different AR(1) coefficients.

2. *Identification given a particular estimate.* Weaker forms of identification suffice for many purposes. We are often not interested in *arbitrary* observables B. Often, we are only interested in one particular estimate, for example the transition matrix B^* that comes from a particular data set. It is enough to say that we we have uniquely measured the parameters if there is only one θ^* corresponding to *this* B^*. A unique θ for each B *near* B^* is also interesting for distributional questions. Rejection regions can expand if there are many θ corresponding to some B that are statistically near B^*, even if there is only one θ^* corresponding to an estimate B^* itself.

2a. *Local identification at a particular estimate.* Even restricting attention to one particular B^*, there is a weak and a strong form of identification. A set of parameters θ^* is *locally identified* if there is no other set of parameters θ within a small neighborhood of θ^* that produce the same B^*. This case is important, because it's easy to check for in estimation, even when models must be solved numerically. We need simply to check that the matrix $dvec\left[(B(\theta))\right]/d\theta'$ has the same rank as the number of parameters in θ. Equivalently, when estimation results from minimizing an objective, we can check that the matrix of second derivatives of the objective function with respect to parameters is full rank.

2b. *Global identification at a particular estimate.* We are particularly concerned with a somewhat stronger kind of identification. Often, a set of parameters is locally identified (at least we know the author's computer programs converged and did not report a singular Hessian), and the parameters are well inside the region of determinacy. We want to know, though, is there a set of parameters *far away* from these parameters, in the region of indeterminacy, that produces the same B^*? A derivative test will not tell us the answer to this question. For example, suppose the observable implications of a model with parameter θ could be wrapped up in the autoregression of a single observable y_t, of the form

$$y_t = \theta^2 y_{t-1} + \varepsilon_t$$

Given a measurement of the autoregression coefficient, the parameter θ is *locally* identified. However, $\theta = -\theta^*$ would work just as well.

5. *Identification of regions.* For determinacy questions, less information may suffice. A parameter may not be either globally or locally identified; there may be a large set of θ^* that have the same implications for B^*. To test for determinacy, it is *sufficient* to uniquely identify all the structural parameters, but it is not *necessary.* If we can show that the entire set of θ^* consistent with the given B^* lies in the region of determinacy, then we can identify determinacy without identifying all the parameters. Lubik and Schorfheide (2004), reviewed below, investigate this kind of identification.

6. *Reasonable identification.* Finally, we may have many parameters that produce the same dynamics B^*, and some of these may in the region of indeterminacy. However, we may feel that the parameters that violate indeterminacy are unreasonable, either from a-priori economic reasoning or other considerations.

7. *Weak identification.* We often say parameters are "weakly identified" when, though there is a one-to-one correspondence between θ and B at or near B^*, very small changes in B imply very large changes in θ.

4.2 Identification in current estimates

4.2.1 Rotemberg, Giannoni and Woodford

Rotemberg and Woodford (1997, 1998, 1999) and Giannoni and Woodford (2005) go on from OLS estimation of the policy rule (45) to estimate structural parameters. They solve the full model, and search for parameters that come closest to matching the response functions of the endogenous variables to the monetary policy shock.

In the most extensive identification discussion, Rotemberg and Woodford (1998) report (p. 21) that in fact the remaining structural parameters are not separately identified

> In fact, one can show that only four combination of the structural parameters can be identified from the impulse responses to a monetary policy shock, given a particular feedback rule for the monetary policy. The parameters β, κ and σ are each identified, but only a single function of ω and ψ is, rather than either of these being identified independently.

Their response is to fix a-priori the remaining parameters (κ, σ and ψ , p. 24) a-priori and only estimate three parameters. (They use the word "calibrate" and refer to other studies, but if this is the model of the world, then the other studies can't identify them either.)

4.2.2 Ireland

Ireland (2007) is an excellent recent example of a full-system approach. It is estimated by maximum likelihood, so in this sense we know that Ireland exploits all the information the data has about his model.

Ireland specifies a policy rule (his Equation (7), p.9, in my notation)

$$\begin{aligned} i_t &= i_{t-1} + \phi_\pi \left(\pi_t - \pi_t^*\right) + \phi_y \left(\Delta y_t - \Delta y\right) + v_t \\ v_t &= \rho_v v_{t-1} + \sigma_v \varepsilon_{vt} \\ \pi_t^* &= \pi_{t-1}^* - \delta_\theta \varepsilon_{\theta t} - \delta_z \varepsilon_{zt} + \sigma_\pi \varepsilon_{\pi t} \end{aligned}$$

$\varepsilon_{\theta t}$ is the "cost push shock" and ε_{zt} is the "technology shock" that appear in the supply or Phillips-curve equations.

This is a much more general policy rule than studied so far: it includes a serially correlated disturbance v_t, and through π_t^* it contains a "stochastic intercept," linked to the disturbances to other equations of the model. It is clearly hopeless to estimate such a rule by single-equation methods. However, by estimating an entire system we can hope to measure the shocks such as ε_θ and ε_z that enter the stochastic intercept of the policy rule, and to distinguish them from the monetary policy disturbance $\varepsilon_{\pi t}$. By estimating the dynamics of the entire model, we can hope to identify ϕ_π and ϕ_y despite the evident correlation between $\pi_t, \Delta y_t$ and v_t.

For given parameter values, Ireland solves the complete model as I have outlined in section 2.4 above. With a full law of motion for observables, he can construct the likelihood function and maximize it. Ireland's approach is implemented in the DYNARE[5] package of computer programs.

There is a global and a local identification issue 1: What does the model solution algorithm do when we want to evaluate parameters that generate insufficient (or overabundant) eigenvalues greater than one, i.e. a region of indeterminacy? 2: Are the structural parameters locally identified, i.e. given that we are in the interior of a region in which we have just enough eigenvalues greater than one, can one identify all the parameters or just a subset of them?

In Ireland's case, the first answer is that the estimate is *constrained* to lie in the region of local determinacy. Ireland subtracts a very large number from the log-likelihood for any parameter set that leads to local indeterminacy, and does not attempt to solve the model for such parameters. Thus, we don't know if other parameters, in the indeterminacy region, can give the same dynamics. Ireland's estimate also displays symptoms of local lack of identification, that several parameters must be imposed a-priori. On p. 13, Ireland finds analytically that θ (Dixit-Stiglitz elasticity of substitution) and ϕ (price adjustment cost) are not separately identified, so only the composite parameter $\psi = (\theta - 1)/\phi$. One of ϕ or θ must be calibrated. On p. 15, Ireland finds that the estimates of ψ are unreasonable, so he calibrates that as well[6]. 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.

4.2.3 Smets and Wouters

Smets and Wouters (2003) is an early and influential full-model estimates. Computer programs implementing their approach are also available on the DYNARE website, so we are likely to seem many more estimates following their approach. They solve an extended new-Keynesian model, fit to European data.

As a common sign of local identification problems, they fix several parameters a-priori, including the discount factor β, the depreciation rate, the power on capital in the Cobb-Douglas production function, and the "parameter capturing the markup in wage setting as this parameter is not identified" (p., 1141)

Smets and Wouters take a Bayesian approach, specifying priors for the other parameters. Most interesting, from our perspective, are the parameters of the policy rule.

[5] http://www.cepremap.cnrs.fr/dynare/

[6] On both issues, I thank Peter Irealnd for showing me how the estimates work.

The prior for the inflation coefficient is normal with a mean of 1.70 and a standard error of 0.10 (their Table 1). If we sample from a priors seven standard deviations from the boundary for indeterminacy, the chance of drawing an indeterminate parameter ($\phi < 1$) is $\Phi(-7) = 1.3 \times 10^{-12}$. This calculation is just suggestive, of course. Smets and Wouters' policy function responds to the output gap, and to changes in inflation and output gap, and determinacy is a property of the whole model. Still, it is quite likely that Smets and Wouters' simulation never even asked about parameters in the indeterminacy region. Even if it did, multiplying by priors on the order of 1.3×10^{-12}, it's clear that they do not address the question "could the same observational implications be generated by a different set of parameters in the indeterminacy region?"

The difference between prior and posterior is a measure of how much the data have to say about a parameter. Tellingly, 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) examine the Smets-Wouters model in detail. They find that changing priors affects the 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." They call this "over-parameterized," which is the same as "underidentified." Most interestingly for my quest, when they substitute a uniform prior between 1 and 4 on the inflation coefficient in monetary policy ϕ_π (Table 1) for Smets and Wouters' tight normal, their estimate is the upper bound 4.00 (Table 3).

4.2.4 Summary

From reading, then, we don't know much about whether the current generation of new-Keynesian model estimates is identified. We have some indications of weak local identification. But we don't know much at all about my central question, whether there are parameters from the indeterminacy region that give the same predictions for observables. Most computer programs for solving new-Keynesian models will not produce solutions if there are insufficient or excessive eigenvalues greater than one, so this question is typically not even asked.

This is not a criticism of the authors. They are not interested in testing for determinacy, or 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. 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. In the effort to match dynamics, these are natural and harmless simplifications. Lack of identification is, as expressed by Ontaski and Williams (2004), almost a feature not a bug, as it means the model's ability to match dynamics is "robust." The models are and meticulously specified and long and hard, so I do not mean to criticize them by saying they should have asked additional questions.

The question remains, though: *Are* parameters identified in the full-system approach? This certainly becomes an important question when we move past matching dynamics to welfare calculations and optimal-policy calculations. Parameters that are not identified or weakly identified can have vastly different welfare consequences. What kinds of assumptions must we make to identify parameters, and those assumptions reasonable? Can the full-system approaches escape the non-identification theorems I outlined above? In particular, can we identify enough parameters to infer that the "unstable" eigenvalues are in fact large enough, so a full-system estimate might be able to document a shift from an indeterminate to a determinate regime more convincingly than a single-equation estimate?

4.3 Identification in three-equation models

The Rotemberg-Woodford, Giannoni-Woodford, Smets-Wouters and Ireland models are large, complex, and must be solved numerically. Analytically characterizing identification and determinacy in these models is an algebraic nightmare, and it is doubtful results of any generality would emerge. Therefore, I turn to an expanded three-equation model that is just simple enough to be analytically tractable, yet complex enough to illustrate interesting and nontrivial possibilities.

I consider a fairly general case of the standard three-equation model:

$$
\begin{aligned}
i_t &= \phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + \phi_{y,0}y_t + \phi_{y,1}E_t y_{t+1} + x_{it} + \theta_y x_{yt} + \theta_\pi x_{\pi t} \\
y_t &= E_t y_{t+1} - \sigma\left(i_t - E_t\pi_{t+1}\right) + x_{yt} \\
\pi_t &= \beta E_t\pi_{t+1} + \gamma y_t + x_{\pi t} \\
x_{it} &= \rho_i x_{it-1} + \varepsilon_{it} \\
x_{yt} &= \rho_y x_{t-1} + \varepsilon_{yt} \\
x_{\pi t} &= \rho_\pi x_{\pi t-1} + \varepsilon_{\pi t}
\end{aligned}
$$

Comments on specification

To specify this model, I have introduced three disturbances. We need at least three disturbances to break stochastic singularities between the three observable variables y_t, π_t, i_t. Identification can be purchased by such singularities, but not very convincingly. For example, if we write a policy rule $i_t = \phi_\pi \pi_t$, then we will certainly be able to identify ϕ_π, but we will also predict an easily-rejectable 100% R^2 in the policy-rule regression. On the other hand, three disturbances can be a restrictive assumption. King's (2000) model (10)-(12) has two disturbances in the Phillips curve, a separate potential output $\bar{y}_t$ as well as a disturbance $x_{\pi t}$, which I have merged into one. Limiting the model to three disturbances allows us to recover disturbances from the three observable variables, which is convenient but not necessarily realistic. Thus, to the extent that we obtain identification, sensitivity to additional disturbances is obviously a concern.

I specify AR(1) processes for the disturbances. This too is a restrictive assumption. It means that current values of the disturbances x_t will be state variables. Additional dynamics in the disturbances act exactly like additional disturbances, giving a system with more state variables. My review of Lubik and Schorfheide (2004) below finds that their identification comes only from restricting the disturbance process, so any identification found here really needs to be checked in the same direction. Most of my point, however, is a cautionary tale about how identification can be lost, and these are simply more possibilities for that point which I do not explore.

I allow a serially correlated disturbance x_{it} in the policy rule (46). I also allow the disturbances to enter the policy rule, as my above discussion suggests they should, and as Ireland (2007) does. I allow the standard responses to current and expected future inflation and to current and expected future output.

The system (46)-(46) results in an AR(1) representation for observable variables

$$\begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ i_{t+1} \end{bmatrix} = \mathbf{B} \begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} + \mathbf{C} \begin{bmatrix} \varepsilon_{yt+1} \\ \varepsilon_{\pi t+1} \\ \varepsilon_{\imath t+1} \end{bmatrix} \tag{46}$$

I allow arbitrary cross-correlation among the shocks ε_t, so no implications flow from the **C** matrix of observables. Everything the model predicts about data is captured in the transition dynamics of the **B** matrix.

The **B** matrix has 9 elements, so at best we can hope to identify 9 parameters. We have 6 structural parameters, β, γ, σ, ρ_i, ρ_y, ρ_π, and 6 policy rule parameters $\phi_{\pi,0}$, $\phi_{\pi,1}$, $\phi_{y,0}$, $\phi_{y,1}$, θ_y, θ_π so we know ahead of time that they can't all be identified. However, the exercise is still not trivial: We will be able to see what is and what is not identified; how some special cases achieve identification, and how generalizing assumptions removes some of those identification possibilities. For example, it's possible that the policy rule parameters are all identified, and some structural parameters are not identified, or vice versa. It's possible that the specific combinations of policy rule parameters required to establish local determinacy are identified even if the individual parameters are not identified. It's possible that by "calibrating" the structural parameters, we can identify all the policy parameters and establish determinacy. We can only answer these sorts of questions by solving the model.

Solving the model

I proceed exactly as in the simple example of the last section. I relegate the tedious algebra to the appendix. I write (46)-(46) in the standard form, (27). The eigenvalues of the transition matrix are ρ_i, ρ_y, ρ_π and

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

where

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

The condition $\|\lambda\| > 1$ gives rise to a complicated region of determinacy, characterized in the appendix. (Cochrane 2007 also plots some subsets of the determinacy region.) One simple and commonly-studied boundary occurs with $\sigma\phi_{y,1} \neq 1$, real roots $b^2 - 4ac > 0$, and the case $\lambda = 1$ (rather than $\lambda = -1$). In this case, the determinacy condition is

$$\left(\phi_{\pi,0} + \phi_{\pi,1}\right) - \frac{(1-\beta)}{\gamma}\left(\phi_{y,1} + \phi_{y,0}\right) > 1. \tag{47}$$

In the first term, we see the "Taylor principle" that inflation coefficients sum to more than one. However, the second term shows that output responses can substitute for inflation responses. In fact, determinacy can be achieved with no inflation response at all, if the output response is strong enough. Thus, to see if the system lies in this well-studied boundary, we will need to consider all of the ϕ coefficients as well as the structural parameters $(\beta - 1)/\gamma$.

The equilibrium dynamics of the observables can be described by

$$\mathbf{z}_t = \boldsymbol{\varrho}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_{zt}$$

where

$$\boldsymbol{\varrho} \equiv \begin{bmatrix} \rho_i & 0 & 0 \\ 0 & \rho_y & 0 \\ 0 & 0 & \rho_\pi \end{bmatrix}$$

and then observables follow

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = \mathbf{Q}\mathbf{z}_t.$$

With three equations and three AR(1) disturbances, we can recover the disturbances from the observables. The $\mathbf{z}$ variables and their shocks are just scaled versions of the x disturbances and their shocks,

$$\begin{aligned} \alpha_i z_{it} &= -\sigma x_{it} \\ \alpha_y z_{yt} &= (1 - \sigma\theta_y)\, x_{yt} \\ \alpha_\pi z_{yt} &= -x_{\pi t} \end{aligned}$$

where

$$\begin{aligned} \alpha_i &\equiv (1-\rho_i)(1-\rho_i\beta) + \sigma\gamma\left(\phi_{\pi,0} + \rho_i\left(\phi_{\pi,1} - 1\right)\right) + \sigma(1-\beta\rho_i)\left(\phi_{y,0} + \rho_i\phi_{y,1}\right) \\ \alpha_y &\equiv \left(1-\rho_y\right)\left(1-\rho_y\beta\right) + \sigma\gamma\left(\phi_{\pi,0} + \rho_y\left(\phi_{\pi,1} - 1\right)\right) + \sigma\left(1-\beta\rho_y\right)\left(\phi_{y,0} + \rho_y\phi_{y,1}\right) \\ \alpha_\pi &\equiv (1-\rho_\pi)(1-\rho_\pi\beta) + \sigma\gamma\left(\phi_{\pi,0} + \rho_\pi\left(\phi_{\pi,1} - 1\right)\right) + \sigma(1-\beta\rho_\pi)\left(\phi_{y,0} + \rho_\pi\phi_{y,1}\right) \end{aligned}$$

We can equivalently write the VAR representation (46) of the observables as

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = \mathbf{Q}\boldsymbol{\varrho}\mathbf{Q}^{-1}\begin{bmatrix} y_{t-1} \\ \pi_{t-1} \\ i_{t-1} \end{bmatrix} + \mathbf{Q}\boldsymbol{\varepsilon}_{zt}. \tag{48}$$

The columns of the $\mathbf{Q}$ matrix are

$$\mathbf{Q}_{:,1} = \begin{bmatrix} 1-\beta\rho_i \\ \gamma \\ \gamma\rho_i - \frac{1}{\sigma}(1-\rho_i)(1-\rho_i\beta) \end{bmatrix} \tag{49}$$

$$\mathbf{Q}_{:,2} = \begin{bmatrix} 1-\beta\rho_y \\ \gamma \\ \frac{1}{1-\sigma\theta_y}\left\{ \begin{array}{c} \left[\left(1-\rho_y\right)\left(1-\beta\rho_y\right)-\sigma\gamma\rho_y\right]\theta_y \\ +\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)+\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right) \end{array}\right\} \end{bmatrix} \tag{50}$$

$$\mathbf{Q}_{:,3} = \begin{bmatrix} \sigma\rho_\pi+\sigma(1-\beta\rho_\pi)\theta_\pi-\sigma\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right) \\ (1-\rho_\pi)+\sigma\gamma\theta_\pi+\sigma\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right) \\ \left\{ \begin{array}{c} \left[\sigma\gamma\rho_\pi-(1-\rho_\pi)(1-\rho_\pi\beta)\right]\theta_\pi \\ +(1-\rho_\pi)\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)+\rho_\pi\sigma\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right) \end{array}\right\} \end{bmatrix} \tag{51}$$

Now, we can summarize everything the data has to say about parameters of this model by the 9 elements of the VAR transition matrix $\mathbf{B}=\mathbf{Q}\boldsymbol{\varrho}\mathbf{Q}^{-1}$. (See (46) and (48)). We can also first eigenvalue-decompose $\mathbf{B}$, and summarize everything the data has to say about parameters by the 3 eigenvectors and 6 independent elements of the eigenvector matrix. The latter approach is clearly easier algebraically. If these expressions for the eigenvector matrix $\mathbf{Q}$ seem complex, try inverting $\mathbf{Q}$! In this VAR system, $\mathbf{Q}_{ij}\rho_j^k$ gives the response of variable i to a shock k at the k horizon. Thus, by studying the elements of $\mathbf{Q}$ we are also studying what the impulse-response function can tell us about system parameters.

In sum, then, our identification question is this: Given $\boldsymbol{\rho}$ and $\mathbf{Q}$ matrices, corresponding to a set of parameters $\rho^*,\beta^*,\gamma^*,\sigma^*,\phi^*,\theta^*$, are there other parameters $\rho,\beta,\gamma,\sigma,\phi,\theta$ that generate the same $\mathbf{Q}$?

What's identified?

1. The persistence parameters ρ_i,ρ_y,ρ_π and all the structural parameters of the economy β,γ,σ are identified. The three persistence parameters are easily identified from the eigenvalues of the VAR transition matrix $\mathbf{B}$. Taking ratios of the first and second rows, the top left block $\mathbf{Q}_{1:2,1:2}$ lets us identify $(1-\beta\rho_i)/\gamma$ and $\left(1-\beta\rho_y\right)/\gamma$. Since we know ρ_i and ρ_y, and presuming they are different, we can separately identify β and γ. $\mathbf{Q}_{3,1}$ now contains only the unknown σ, so σ is now identified. (Global identification in this system requires ruling out special cases such as $\rho_i=\rho_y$, and I will not take up space listing the special cases.)

If we only study responses to the monetary policy disturbance x_{it}, we only see the first column of $\mathbf{Q}$, i.e., (49). That column allows us to identify $(1-\beta\rho_i)/\gamma$, but not β and γ separately. Giannoni and Woodford (2005) and Rotemberg and Woodford (1999) were similarly only able to identify certain combinations of structural parameters, but not the individual parameters. Here, we could fix β a-priori as they do, and then estimate γ and σ. However, by examining the response to both monetary policy shocks and demand (y) shocks, we can in this system estimate β, γ and σ separately. This result suggests that the Rotemberg-Woodford model's limited identification of structural parameters did not result from the determinacy/indeterminacy issues I study, and could be easily fixed by examining more responses. This sort of result is a nice argument in favor of a full-system approach to identification and estimation.

2. The response to the monetary policy shock $\mathbf{Q}_{:,1}$ *contains no information about the parameters of monetary policy* ϕ *and* θ. Sensibly, perhaps, to have any hope of identifying monetary policy, we must examine how the economy, including the central bank, responds to *other* shocks.

3. Knowledge of the structural parameters $\beta, \sigma, \gamma, \rho_i, \rho_\pi, \rho_y$ *would not help to identify the policy rule parameters* ϕ, θ. One might think, based on counting parameters and the 9 degrees of freedom, that "calibrating" some of the structural parameters to sensible values would leave more degrees of freedom for identifying policy rule parameters. However, the ρ are identified from the eigenvalues of the transition matrix, and β, σ, γ are identified from the first column and first two rows of the second column of $\mathbf{Q}$, and these do not contain policy parameters anywhere. Hence, the identification of policy parameters cannot be helped by better knowledge of structural parameters.

Policy parameters and determinacy

Now, we can move on to the policy parameters ϕ and θ. $\mathbf{Q}_{3,2}$ (50) and the third column $\mathbf{Q}_{:,3}$ (51) contain all of our information about policy parameters. However, these elements of $\mathbf{Q}$ are all linear functions of the parameters ϕ and θ, and $\mathbf{Q}_{3,3}$ is a linear combination of $\mathbf{Q}_{1,3}$ and $\mathbf{Q}_{2,3}$: $-\frac{1}{\sigma}(1-\rho_\pi)\mathbf{Q}_{1,31} + \rho_\pi \mathbf{Q}_{2,3} = \mathbf{Q}_{3,3}$. Thus, all our information about the ϕ and θ parameters comes down to two restrictions,

$$\frac{\left[\left(1-\rho_y\right)\left(1-\beta\rho_y\right)-\sigma\gamma\rho_y\right]\theta_y+\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)+\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right)}{\gamma\left(1-\sigma\theta_y\right)}=\frac{\mathbf{Q}_{3,2}}{\mathbf{Q}_{2,2}} \tag{52}$$

$$\frac{\left(1-\rho_\pi\right)+\sigma\gamma\theta_\pi+\sigma\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right)}{\sigma\rho_\pi+\sigma\left(1-\beta\rho_\pi\right)\theta_\pi-\sigma\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)}=\frac{\mathbf{Q}_{1,3}}{\mathbf{Q}_{2,3}}. \tag{53}$$

4. If we know ϕ*, we can measure* θ. Note that only θ_y enters (52), derived from the responses to the x_y disturbance, and only θ_π enters (53), derived from the responses to the x_π disturbance. Thus, the most natural interpretation of these equations is this: If we somehow know the ϕ parameters, then (52)-(53) allow us to measure the θ parameters. If we *know* how the central bank would respond to the endogenous variables – precisely the determinacy question (the θ do not enter (47) or any other determinacy conditions) – then we we could *measure* how the central bank responds to shocks.

However, what we're after is to see how the ϕ parameters might be identified, or at least seem to be identified. With 2 linear equations and 6 unknowns, we're obviously not going to identify everything. But, if we fix some parameters a-priori we can identify others.

5. If we set $\theta_y = 0$ *and* $\theta_\pi = 0$*, then we can identify any two of* $\phi_{\pi,0}, \phi_{\pi,1}, \phi_{y,0}, \phi_{y,1}$. Most empirical specifications assume away policy responses to shocks, so this is an interesting case to examine. Equations (52)-(53) become

$$
\begin{aligned}
\gamma\left(\phi_{\pi,0}+\rho_y\phi_{\pi,1}\right)+\left(1-\beta\rho_y\right)\left(\phi_{y,0}+\rho_y\phi_{y,1}\right) &= \gamma\frac{\mathbf{Q}_{3,2}}{\mathbf{Q}_{2,2}} \\
\frac{\mathbf{Q}_{1,3}}{\mathbf{Q}_{2,3}}\sigma\left(\phi_{\pi,0}+\rho_\pi\phi_{\pi,1}\right)+\sigma\left(\phi_{y,0}+\rho_\pi\phi_{y,1}\right) &= \sigma\rho_\pi\frac{\mathbf{Q}_{1,3}}{\mathbf{Q}_{2,3}}-(1-\rho_\pi)
\end{aligned}
$$

Setting any two ϕ to zero, we retain a linearly independent system (except, as usual for peculiar parameter sets such as $\rho_\pi = \rho_y$.) For example, we could identify responses to current output and inflation, $\phi_{\pi,0}$ and $\phi_{y,0}$; or we could identify inflation responses $\phi_{\pi,0}$ and $\phi_{\pi,1}$.

This is a remarkable result, worth highlighting. The policy rule *does* have a disturbance x_{it}, and that disturbance is serially correlated, and correlated with the other shocks in the economy. The right hand variables of the Taylor rule can move contemporaneously and so are correlated with its disturbance. The policy rule cannot be estimated by single-equation methods. Yet if we restrict the policy rule to a two-parameter form, for example,

$$i_t = \phi_{\pi,0}\pi_t + \phi_{\pi,1}E_t\pi_{t+1} + x_{i,t+1} \tag{54}$$

then we *can* identify both $\phi_{\pi,0}$ and $\phi_{\pi,1}$ by examining the dynamics of the whole system. The *presence* of the all three shocks is key to this identification.

6. We cannot identify ϕ, nor can we identify whether the system is determinate, if we allow the standard four responses $\phi_{\pi,0}, \phi_{\pi,1}, \phi_{y,0}, \phi_{y,1}$. The two linear equations (54) and (54) obviously can't separately identify all four of $\phi_{\pi,0}, \phi_{\pi,1}, \phi_{y,0}, \phi_{y,1}$. In particular, the apparently attractive identification of (54) falls apart if we allow any output responses $\phi_{y,0}$ or $\phi_{y,1}$.

Perhaps, however, the set of observationally-equivalent ϕ parameter do not range to far, so if we start with a parameter set ϕ^* well inside the region of determinacy, all the alternative possibilities also lie in that region? Alas, this hope is dashed as well. Equations (54)-(54) are linear, so unidentified parameters can range over the whole real line.

The only hope is that the two linear combinations of ϕ that (54)-(54) allow us to identify are also the linear combinations we have to test, as in (47). Alas, they are different, in a revealing way. We can see from (54)-(54) that the linear combinations we can identify are weighted by persistence parameters such as ρ_y, ρ_π. These parameters do not appear in determinacy conditions. (The explicit expression for the identified linear combinations of ϕ is not pretty or revealing, so I relegate it to the appendix.)

In sum, even without bringing up responses to disturbances θ, i.e. "stochastic intercepts," if we allow the possibility that the central bank responds to the standard current and expected future output and inflation, we cannot identify that the system is determinate.

7. We cannot identify any of the ϕ responses if we allow policy to respond to shocks, $\theta_\pi \neq 0$ and $\theta_y \neq 0$. Any identification of policy-rule parameters necessary for determinacy must come by assuming away policy responses to disturbances. This result is clear by simple inspection of (54)-(54). Even if we restrict the policy rule as much as possible, say by only allowing $\phi_{\pi,0} \neq 0$, we still cannot identify three parameters with these two linear equations. In particular, different possibilities for the θ imply different, observationally-equivalent values of ϕ.

Perhaps we can identify whether the system lies in the region of determinacy, even if we cannot identify the parameters? Again, that hope fails: (54)-(54) are linear restrictions on ϕ, so as we vary θ_y and θ_π, ϕ varies over the whole real line.

Perhaps escaping the region of determinacy requires economically implausible parameters? To answer this question, I express $\mathbf{Q}$ as a function of an initial set of parameters ϕ^*, θ^*, and I express (54)-(54) in terms of deviations from those parameters. To keep the algebra relatively simple, I present the case in which we only consider $\phi_{\pi,0} \neq 0$, i.e., I start at a given $\phi^*_{\pi,0}$, I set $\phi^*_{\pi,1} = 0, \phi^*_{y,0} = 0, \phi^*_{y,1} = 0, \theta^*_\pi = 0, \theta^*_y = 0$, and I consider variation only in $\phi_{\pi,0}, \theta_\pi, \theta_y$. (The monstrous formula for the general case is in the Appendix)

In this example, the observationally-equivalent changes $\tilde{\phi}_{\pi,0} = \phi_{\pi,0} - \phi^*_{\pi,0}, \tilde{\theta}_\pi = \theta_\pi - \theta^*_\theta, \tilde{\theta}_y = \theta_y - \theta^*_y$ obey

$$
\begin{aligned}
0 &= \left[\sigma\gamma\phi^*_{\pi,0} + \left(1-\rho_y\right)\left(1-\beta\rho_y\right) - \sigma\gamma\rho_y\right]\tilde{\theta}_y + \gamma\tilde{\phi}_{\pi,0} \\
0 &= \left[\sigma\gamma\phi^*_{\pi,0} + \left(1-\rho_\pi\right)\left(1-\beta\rho_\pi\right) - \sigma\gamma\rho_\pi\right]\tilde{\theta}_\pi - \left(1-\rho_\pi\right)\tilde{\phi}_{\pi,0}.
\end{aligned}
$$

So, suppose we start with $\phi^*_{\pi,0} = 2$, well inside the region of determinacy $\phi_{\pi,0} > 1$. How much θ_y and θ_π do we need to assume to find an indeterminate $\phi_{\pi,0} < 1$ that is observationally equivalent? Using $\sigma = 1, \gamma = 1, \beta = 0.95, \rho_\pi = \rho_y = 0.7, \phi^*_{\pi,0} = 2$, we have

$$
\begin{aligned}
0 &= 1.40 \times \tilde{\theta}_y + \tilde{\phi}_{\pi,0} \\
0 &= 1.40 \times \tilde{\theta}_\pi - 0.3 \times \tilde{\phi}_{\pi,0}.
\end{aligned}
$$

Thus, to reduce $\phi_{\pi,0}$ by one, we need to decrease θ_y by $1/1.4 = 0.71$ and we need to increase θ_π by $0.3/1.4 = 0.21$. θ_y and θ_π capture how much the interest rate responds to shocks to output y and inflation π, so they have the same units as ϕ_y and ϕ_π. These are therefore quite sensible values. And, of course, smaller values are needed if one can also look over the range of $\phi_{\pi,1}, \phi_{y,0}, \phi_{y,1}$ as well.

Summary

The full-system approach starts with great promise: by fully exploiting all the predictions of the model, we may be able to identify parameters or combinations that single-equation methods cannot identify. And, some of that promise is fulfilled: in this example, the full-system approach can identify all the structural parameters $\beta, \gamma, \sigma, \rho$. If we assume away policy response to shocks $\theta_y = \theta_\pi = 0$, and only allow two responses to inflation and output, then we can also identify the policy rule parameters, even though there is a serially-correlated policy disturbance, correlated with the right hand variables of the policy equation.

Alas, the policy-rule identification, and identification of determinacy, falls apart if the central bank follows even slightly generalized rules. If we allow the standard four response coefficients $\phi_{\pi,0}, \phi_{\pi,1}, \phi_{y,0}, \phi_{y,1}$, then we cannot identify the parameters individually, nor if the system is in the determinate region. For example, estimates of $\phi_{\pi,0}$ and $\phi_{\pi,1}$ in the determinacy region are observationally equivalent to estimates that include $\phi_{y,0} \neq 0, \phi_{y,1} \neq 0$ for which the system is not determinate. Allowing policy to respond directly to the shocks in "Wicksellian" fashion removes all hope of identifying the policy parameters ϕ on which determinacy rests. Allowing more disturbances or more general stochastic processes for those disturbances can only make matters worse.

5 Lubik and Schorfheide; Testing regions

Lubik and Schorfheide (2004) test for determinacy vs. indeterminacy in the simple model I presented above. They 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 (2006), reviewed below, make this point with a series of examples.)

Lubik and Schorfheide explain their ideas in the same single-equation setup as I use above, simplifying even further by assuming a white noise monetary policy disturbance

$$x_t = \varepsilon_t$$

(i.e., $\rho = 0$). The equilibrium is characterized again by (4) which becomes

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

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

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

Thus, 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, than 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* $\{i_t, \pi_t\}$, *and for any* ϕ, *one can construct an* x_t *process that generates the given process for the observables* $\{i_t, \pi_t\}$. *If* $\phi > 1$, *the observables are generated as the unique bounded forward-looking solution. In either case, given the process* $\pi_t = a(L)\varepsilon_t$ *we construct* $x_t = b(L)\varepsilon_t$ *with*

$$b_j = a_{j+1} - \phi a_j, \tag{55}$$

or, in lag operator notation,

$$b(L) = \left(L^{-1} - \phi\right) a(L) - a(0)L^{-1}. \tag{56}$$

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. Thus, 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., so that Equation (5) holds,

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

It's easy enough to check that (55) is correct:

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

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

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

The derivation of (55)-(56) takes a bit more algebra. I use the Hansen and Sargent (1980) formulas for the moving average representation of an expected discounted value. I present the derivation in the Appendix.

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 (56) follows.

i_t is just given by $i_t = r + E_t(\pi_{t+1})$, 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 (55),

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

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 (56), 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 $(1-\phi L^{-1})$ 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).

6 Taylor on Taylor rules

A natural question is, "how does Taylor think Taylor rules work?" The best instance I can find to answer this question is Taylor (1999). Taylor adopts a "simple model" (p. 662, in my notation)

$$\begin{aligned} y_t &= -\sigma(i_t - \pi_t - r) + u_t \\ \pi_t &= \pi_{t-1} + \gamma y_{t-1} + e_t \\ i_t &= \bar{\imath} + \phi_\pi \pi_t + \phi_y y_t \end{aligned}$$

This is a classic "old-Keynesian model," in that all the forward-looking terms are absent. (The companion paper, Cochrane 2007 discusses it a bit more fully.) As a result, $\phi_\pi > 1$ is the condition for *stable* dynamics, in which we solve for endogenous variables as a function of *past* shocks. Formally, with $\bar{\imath} = r$, the standard form of the model is

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \sigma\gamma\frac{1-\phi_\pi}{1+\sigma\phi_y} & \sigma\frac{1-\phi_\pi}{1+\sigma\phi_y} \\ \gamma & 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} \frac{1}{1+\sigma\phi_y} & \sigma\frac{1-\phi_\pi}{1+\sigma\phi_y} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ e_t \end{bmatrix} \qquad (57)$$

The eigenvalues of the transition matrix are

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

$\phi_\pi > 1$ and $\phi_y > 0$ generate λ_1 *less* than one, so we solve for y_t, π_t as a function of past shocks.

Identification in Taylor's model is similar to identification in familiar "old-Keynesian" models, as reviewed by Sims (1980). It's easy if we restrict the stochastic process of the disturbances. For instance, if we assume u_t and e_t are unpredictable from time $t-1$ information, then we can simply run a first-order VAR and recover $\sigma, \gamma, \phi_\pi, \phi_y$ from the VAR regression coefficients in (57). If we do not make any assumptions about error terms, identification is more difficult, as we cannot easily separate disturbance dynamics from model dynamics. However, this is an old and familiar issue (Sims (1980)), having nothing to do with multiple equilibria.

Unsurprisingly, identification of parameters depends on the model that those parameters describe. The theorem is not "the policy rule parameters are unidentified," but "Parameters crucial to determinacy are unidentified in new-Keynesian models."

7 Related literature

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

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

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

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

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

$$\begin{bmatrix} u_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 0.05 \\ 0.02 \\ 0.05 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0.05 \\ -0.5 & 1 & -0.25 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}. \tag{58}$$

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 (58) 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 at some date in the future it will lead to an explosion, even if it does not lead to an explosion under the current regime. 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) also 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 after 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 there will be variation of nominal interest rates, output and inflation in equilibrium, 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,

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

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.

8 Conclusions and implications

My main motivation for looking at identification in new-Keynesian models is to assess a central empirical success: 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. My main point, in this context, is that this important historical episode is misinterpreted. We cannot, through new-Keynesian Taylor-rule glasses, read either regressions or full-model estimates as evidence that the US moved from an "indeterminate" regime in the 1970s to a "determinate" regime thereafter. And even if we did, a theory that has absolutely nothing to say about inflation in the 1970s, other than "inflation is indeterminate, so any value can happen" is surely a bit lacking. Beyond this motivation, however, the analysis reveals a more generally unsettling lack of identification in the empirical implementation of new-Keynesian models.

Clarida, Galí, and Gertler not only find coefficients greater than one, they find coefficients a *lot* greater than one, as do most authors who run such regressions. Their coefficients range from a baseline 2.15 (Table IV) to as much as 3.13 (Table V). These coefficients imply that if the US returned to 12% inflation (a 10 percentage point rise), the Federal reserve would raise the Federal Funds rate to a value between $5.25+21.5 = 26.75\%$ and $5.25 + 31.3 = 36.55\%$, implying astronomical real rates. If these predictions seem implausibly large, even for the current inflation-oriented Federal Reserve, then digesting them as something less than structural helps a great deal.

The identification issue stems from the heart of all new-Keynesian models, which is the need to specify explosive dynamics in order to rule out multiple equilibria, the latter generated by the fact that the model only determines expectations of future variables. Yet the dynamics of equilibrium variables are by nature stationary, i.e. "locally-bounded," so cannot possibly reveal such explosions.

This fact means that endogenous variables must jump in response to disturbances. Alas such jumps mean that right hand variables of a policy rule must jump when there is a monetary policy disturbance, so that equation will be exquisitely hard to estimate. One can only get around this central prediction by strong and arbitrary assumptions, in particular that the central bank does not respond to many variables.

In this paper, I take the new-Keynesian modeling rules as given to ask about identification. In particular, I do not question the restriction to the unique locally-bounded equilibrium. A companion paper, Cochrane (2007), argues that as a matter of theory, there is no economic reason to ignore nominally-explosive equilibria. Thus, the equilibrium-selection device that causes all the trouble is invalid anyway, and we might as well stop dreaming up unobservable behavior to exploit it.

9 References

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

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

Beyer, Andreas, and Roger Farmer 2006, "On the Indeterminacy of Determinacy and Indeterminacy: Comments on the paper 'Testing for Indeterminacy' by Thomas Lubik and Frank Schorfheide" Forthcoming, *American Economic Review.*

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

Clarida, Richard, Jordi Galí, and Mark Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115 (1), 147-180

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, (6) ,1033-1067.

Clarida, Richard, Jordi Galí, and Mark Gertler, 1999, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* 37, 1661-1707.

Cochrane, John H., 2007, "Inflation Determination with Taylor Rules: a Critical Review," Manuscript, University of Chicago.

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.

Davig, Troy and Eric Leeper, 2005, "Generalizing the Taylor Principle," Manuscript, Indiana University at Bloomington.

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.)

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," Manuscript, Boston College, forthcoming *Journal of Money, Credit and Banking.*

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

King, Robert G., 2000 "The New IS-LM Model: Language, Logic, and Limits," Federal Reserve Bank of Richmond *Economic Quarterly* 86 (3), 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.

Loisel, Olivier, 2007, "Bubble-Free Interest-Rate Rules," Manuscript, Banque de France and CEPREMAP.

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

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.

Ontaski, Alexei, and Noah Williams, 2004, "Empirical and Policy Performance of a Forward-Looking Monetary Model," Manuscript, Columbia University and Princeton University.

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.

Rudebusch, Glenn D., 2005, "Monetary Policy Inertia: Fact or Fiction?," Federal Reserve Bank of San Francisco Working Paper 2005-19.
http://www.frbsf.org/publications/economics/papers/2005/wp05-19bk.pdf

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.

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.

Taylor, John B. 1999 "A Historical Analysis of Monetary Policy Rules" In Taylor, John B., Ed. *Monetary Policy Rules,* Chicago: University of Chicago Press.

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