Identifying Taylor rules in macro-finance models*

David Backus,† Mikhail Chernov,‡ and Stanley Zin§

August 15, 2013

Abstract

Identification problems arise naturally in forward-looking models when agents observe more than economists. We illustrate the problem in several macro-finance models with Taylor rules. When the shock to the rule is observed by agents but not economists, identification of the rule's parameters requires restrictions on the form of the shock. We show how such restrictions work when we observe the state directly, indirectly, or infer it from observables.

JEL Classification Codes: E43, E52, G12.

Keywords: forward-looking models; information sets; monetary policy; exponential-affine models.

*Preliminary. We welcome comments, including references to related papers we inadvertently overlooked. The project started with our reading of John Cochrane's paper on the same subject and subsequent emails and conversations with him and Mark Gertler. We thank them both. We also thank Patrick Feve and Gregor Smith for comments on an earlier draft, and participants in seminars at, and conference sponsored by, New York University, the Swedish House of Finance, and Toulouse. The latest version of the paper is available at: http://pages.stern.nyu.edu/~dbackus/Identification/ms/BCZ_trident_latest.pdf.

† Stern School of Business, New York University, and NBER; dbackus@stern.nyu.edu.

‡ Anderson School of Management, UCLA, and CEPR; mikhail.chernov@anderson.ucla.edu.

§ Stern School of Business, New York University, and NBER; stan.zin@nyu.edu.

1 Introduction

The field of macro-finance has the potential to give us deeper insights into macroeconomics and macroeconomic policy by using information from both aggregate quantities and asset prices. The link between bond-pricing and monetary policy seems particularly promising if central banks implement monetary policy through short-term interest rates, as they do in models with Taylor rules.

If the combination of macroeconomics and finance holds promise, it also raises challenges. We address one of them here: the challenge of identifying monetary policy parameters in a modern macroeconomic model. Identification problems arise in many economic models, but they play a particularly important role in assessments of monetary policy. If we see that the short-term interest rate rises with inflation, does that reflect the policy of the central bank or something else? How would we know? Since identification is a feature of models, the question is what we need in a model to be able to identify the parameters governing monetary policy.

We illustrate the problem and point to its possible resolution in a series of examples that combine elements of New Keynesian and macro-finance models. The source of the identification problem here is that we economists do not observe the shock to monetary policy, although the agents populating the model do. As a result, it's difficult, and perhaps impossible, to disentangle the systematic aspects of monetary policy from shocks to it. Without more information about the shock, we may not be able to identify the parameter tying, say, interest rate policy to the inflation rate. This issue about the difference between what agents and economists observe goes back (at least) to work by Hansen and Sargent (1980, 1991), who considered differences in the information sets of agents and economists in dynamic rational expectations models. In our examples, the issue is the Taylor rule shock. If agents observe it but economists do not, then we need restrictions on the shock to identify the Taylor rule's parameters. The identification problem in these models is different from the classical Cowles Commission work on simultaneous equation systems. The central issue is whether we observe the shock, not whether exogenous variables show up in the right configuration. If we observe the shock, identification follows immediately. If not, then we need restrictions on its coefficients.

We show how this works in a series of examples that illustrate how information about shocks interacts with dynamics in forward-looking rational expectations models. The first example is adapted from Cochrane (2011) and consists of the Fisher equation and a Taylor rule.

Later examples introduce exponential-affine pricing kernels and New Keynesian Phillips curves. We find it helpful to separate issues concerning the observation of shocks from those concerning the observation of the underlying state of the economy. If the state is observed, a clear conclusion emerges: we need restrictions on the shock to the Taylor rule to identify the rule's parameters. In our examples, we need one restriction for each parameter to be estimated. If the state is not observed, the same kinds of restrictions typically deliver identification.

If this seems clear to us now, it was not when we started. We thought, at first, that identification required shocks in other equations. We find, instead, that the conditions for identification don't change when we eliminate other shocks, or even when we eliminate the other equations. It depends entirely on what we know about the shock in the equation of interest, the Taylor rule. We also thought that knowledge of the term structure of interest rates might help with identification. We find, instead, that knowing the term structure can be helpful in observing the state, but knowing the state is not enough to identify the Taylor rule. What matters in all of these examples is whether we observe the shock to the Taylor rule, and if we do not, whether we have restrictions on its form.

Identification depends, then, on our ability to observe shocks or, failing that, to impose restrictions on their coefficients. The answer to the next question is less clear: What kinds of restrictions are plausible? Reasonable people can and will have different opinions about this. We leave it for another time.

2 The problem

Two examples illustrate the nature of identification problems in macro-finance models with Taylor rules. The first comes from Cochrane (2011). The second is an exponential-affine bond-pricing model. The critical ingredient in each is what we observe. We assume that economic agents observe everything, but we economists do not. In particular, we do not observe the shock to the Taylor rule. The question is how this affects our ability to infer the Taylor rule's parameters. We provide answers for these two examples and discuss some of the questions they raise about identification in similar settings.

2.1 Cochrane's example

Cochrane's example consists of two equations, an asset pricing relation (the Fisher equation) and a Taylor rule (which depends only on inflation):

$$i_t = E_t\pi_{t+1} \quad (1)$$

$$i_t = \tau\pi_t + s_t. \quad (2)$$

Here i_t is the (one-period) nominal interest rate, π_t is the inflation rate, and s_t is a monetary policy shock. The Taylor rule parameter $\tau > 1$ describes how aggressively the central bank responds to inflation. This model is extremely simple, but it's enough to illustrate the identification problem.

Let us say, to be specific, that the shock is a linear function of a state vector x_t, $s_t = d^\top x_t$, and that x_t is autoregressive,

$$x_{t+1} = Ax_t + Cw_{t+1}, \quad (3)$$

with A stable and disturbances $\{w_t\} \sim \text{NID}(0, I)$. Although simple, this structure is extremely helpful for clarifying the conditions that allow identification. It also allows easy comparison to models ranging from exponential-affine to vector autoregressions. For later use, we denote the covariance of innovations by $V_w = CE(ww^\top)C^\top = CC^\top$ and the covariance matrix of the state by $V_x = E(xx^\top)$, the solution to $V_x = AV_xA^\top + CC^\top$.

We solve the model by standard methods; see Appendix A. Here and elsewhere, we assume agents know the model and observe all of its variables. Equations (1) and (2) imply the forward-looking difference equation or rational expectations model

$$E_t\pi_{t+1} = \tau\pi_t + s_t.$$

The solution for inflation has the form $\pi_t = b^\top x_t$ for some coefficient vector b to be determined. Then $E_t\pi_{t+1} = b^\top E_t x_{t+1} = b^\top Ax_t$. Lining up terms, we see that b satisfies

$$b^\top A = \tau b^\top + d^\top \quad \Rightarrow \quad b^\top = -d^\top(\tau I - A)^{-1}. \quad (4)$$

This is the unique stationary solution if A is stable (eigenvalues less than one in absolute value) and $\tau > 1$ (the so-called Taylor principle). Equation (1) then gives us $i_t = a^\top x_t$ with $a^\top = b^\top A$.

Now consider estimation. Do we have enough information to estimate the Taylor rule parameter τ? We might try to estimate equation (2) by running a regression of i_t on π_t,

with the shock s_t as the residual. That won't work because s_t drives both i_t and π_t, and we need to distinguish its direct effect on i_t from its indirect effect through π_t. Least squares would deliver a coefficient of $\text{Var}(\pi)^{-1}\text{Cov}(\pi, i) = (b^\top V_x b)^{-1} b^\top V_x a$, which is not in general equal to τ.

How then can we estimate τ? The critical issue is whether we observe the shock s_t. Let us say that we — the economists — observe the state x_t, but may or may not observe the shock s_t or the coefficient vector d that connects it to the state. Because we observe x_t, we can estimate A and V_x. We can also estimate the parameter vectors a and b connecting the interest rate and inflation to the state. If we observe the shock s_t, then we can estimate the parameter vector d. We now have all the components of (4) but τ, which we can infer. Evidently the Taylor rule parameter is identified. In fact, it is over-identified. If x has dimension n, we have n equations that each determine τ.

However, if we don't observe s_t, and therefore do not know d, we're in trouble. In economic terms, we can't distinguish the effects on the interest rate of inflation (the parameter τ) and the shock (the coefficient vector d). In this model, there's not much we can do about that.

2.2 An exponential-affine example

Another perspective on the identification problem is that we can't distinguish the pricing relation (1) from the Taylor rule (2). Sims and Zha (2006, page 57) put it this way: "The ... problem ... is that the Fisher relation is always lurking in the background. The Fisher relation connects current nominal rates to expected future inflation rates and to real interest rates[.] ... So one might easily find an equation that had the form of the forward-looking Taylor rule, satisfied the identifying restrictions, but was something other than a policy reaction function." Cochrane (2011, page 598) echoes the point: "If we regress interest rates on output and inflation, how do we know that we are recovering the Fed's policy response, and not the parameters of the consumer's first-order condition?" We'll see exactly this issue in the next example, in which we introduce an exponential-affine model into the problem.

In finance, it's common to model interest rates with exponential-affine models, in which bond yields are linear functions of the state. In the macro-finance branch of this literature, the state includes macroeconomic variables like inflation and output growth. Examples include Ang and Piazzesi (2003), Chernov and Mueller (2012), Jardet, Monfort, and Pegoraro

(2012), Moench (2008), Rudebusch and Wu (2008), and Smith and Taylor (2009). In these models the short rate depends on, among other things, inflation.

An illustrative example follows from the log pricing kernel,

$$m^{\$}_{t+1} = -\lambda^\top \lambda/2 - \delta^\top x_t + \lambda^\top w_{t+1}, \tag{5}$$

and the linear law of motion (3). Here the nominal (log) pricing kernel $m^{\$}_t$ is connected to the real (log) pricing kernel m_t by $m^{\$}_t = m_t - \pi_t$. The one-period nominal interest rate is then

$$\begin{aligned} i_t &= -\log E_t \exp(m_{t+1} - \pi_{t+1}) && (6)\\ &= -\log E_t \exp(m^{\$}_{t+1}) = \delta^\top x_t. && (7) \end{aligned}$$

If we observe the state x_t, we can estimate δ by projecting the interest rate on it.

If the first element of x_t is the inflation rate, it's tempting to interpret equation (7) as a Taylor rule, with the first element of δ the inflation coefficient τ. But is it? The logic of equation (7) is closer to the asset-pricing relation, equation (1), than to the Taylor rule, equation (2). But without more structure, we can't say whether it's one, the other, or something else altogether. This is, of course, the point made by Sims and Zha and echoed by Cochrane. Joslin, Le, and Singleton (2013) make a similar point.

More formally, consider an interpretation of (7) as a Taylor rule (2). Since we observe inflation π_t and the state x_t, we can estimate the coefficient vector b connecting the two: $\pi_t = b^\top x_t$. Then the Taylor rule implies

$$i_t = \tau\pi_t + s_t = \tau b^\top x_t + d^\top x_t.$$

Equating our two interest rate relations gives us $\delta^\top = \tau b^\top + d^\top$. It's clear, now, that we have the same difficulty we had in the previous example: If we do not know the shock parameter d, we cannot infer τ from estimates of δ. If x_t has dimension n, we have n equations to solve for $n+1$ unknowns (d and τ).

If we interpret (7) as an asset pricing relation, then it's evident that we can't distinguish asset pricing (represented by δ) from monetary policy (represented by $\tau b + d$) without more information about the shock coefficients d. Generalizing the asset pricing relation from (2) to (7) has no effect on this conclusion.

2.3 Discussion

These examples illustrate the challenge we face in identifying the parameters of the Taylor rule, but they also suggest follow-up questions that might lead to a solution.

One such question is whether we can put shocks in other places and use them for identification. Gertler (private communication) suggests putting a shock in Cochrane's first equation, so that the example becomes

$$\begin{aligned} i_t &= E_t\pi_{t+1} + s_{1t} \\ i_t &= \tau\pi_t + s_{2t}. \end{aligned}$$

Can the additional shock identify the Taylor rule?

Suppose, as Gertler suggests, that s_{1t} and s_{2t} are independent. Then if s_{1t} is observed, we can use it as an instrument for π_t to estimate the Taylor rule equation, which gives us an estimate of τ. Given τ, we can then back out the shock s_{2t}. We'll see in the next section that this example is misleading in one respect — we do not need a shock in the other equation — but there are two conclusions here of more general interest. One is that identification requires a restriction on the Taylor rule shock. Here the restriction is independence, but in later examples other restrictions serve the same purpose. The other is that identifying τ and backing out the unobserved shock are complementary activities. Generally if we can do one, we can do the other.

A second question is whether we can use long-term interest rates to help with identification. The answer is no if the idea is to use long rates to observe the state. In exponential-affine models, the state spans bond yields of all maturities. In many cases of interest, we can invert the mapping and express the state as a linear function of a subset of yields. In this sense, we can imagine using a vector of bond yields to observe the state. We have seen, though, that observing the state is not enough. We observe the state in both examples, yet cannot identify the Taylor rule. We explore the issue of state observability further in Section 5.

3 Macro-finance models with Taylor rules

Macro-finance models, which combine elements of macroeconomic and asset-pricing models, bring evidence from both macroeconomic and financial variables to bear on our understand-

ing of monetary policy. It's not easy to reconcile the two, but if we do, we gain perspective that's missing from either approach on its own.

We show how Gertler's insight can be developed to identify the Taylor rule in such models. We use two examples, one based on a representative agent, the other on an exponential-affine model. We explore identification in these models when we observe the state, the short rate, and inflation, but not the shock to the Taylor rule. The identification issues are the same: we need one restriction on the shock to identify the (one) policy parameter.

3.1 A representative-agent model

One line of macro-finance research combines representative-agent asset pricing with a rule governing monetary policy. Gallmeyer, Hollifield, and Zin (2005) is a prominent example. We simplify their model, using power utility instead of recursive preferences and a simpler law of motion for the state.

The model consists of equation (6) plus

$$\begin{aligned} m_t &= -\rho - \alpha g_t && (8)\\ g_t &= g + s_{1t} && (9)\\ i_t &= r + \tau\pi_t + s_{2t}. && (10) \end{aligned}$$

Equations (6) and (10) mirror the two equations of Cochrane's example. The former is a variant of equation (7), a more complex version of equation (1) that represents the finance component of the model. The latter is a Taylor rule, representing monetary policy. Equations (8) and (9) characterize the real pricing kernel. The first is the logarithm of the marginal rate of substitution of a power utility agent with discount rate ρ, curvature parameter α, and log consumption growth g_t. The second connects fluctuations in log consumption growth to a shock s_{1t}. As in Section 2, the state x_t follows the law of motion (3) and shocks are linear functions of it: $s_{it} = d_i^\top x_t$ for $i = 1, 2$. For simplicity, we choose r to reconcile the two interest rate equations, which makes mean inflation zero.

The solution now combines asset pricing with a forward-looking difference equation. We posit a solution of the form $\pi_t = b^\top x_t$. Solving (6) then gives us

$$i_t = \rho + \alpha g - V_m/2 + a^\top x_t, \qquad (11)$$

with

$$\begin{aligned} a^\top &= (\alpha d_1^\top + b^\top)A \\ V_m &= a^\top CC^\top a. \end{aligned}$$

Note that the short rate equation (11) now has a shock, as Gertler suggests. Equating (10) and (11) gives us

$$(\rho + \alpha g - V_m/2) + (\alpha d_1^\top + b^\top)Ax_t = r + (\tau b^\top + d_2^\top)x_t.$$

Lining up similar terms, we have $r = \rho + \alpha g - V_m/2$ and

$$(\alpha d_1^\top + b^\top)A = \tau b^\top + d_2^\top \quad \Rightarrow \quad b^\top = (\alpha d_1^\top A - d_2^\top)(\tau I - A)^{-1}.$$

As before, this gives us a unique stationary solution under the stated conditions: A stable and $\tau > 1$.

Now consider identification. Suppose we observe the state x_t, the interest rate i_t, the inflation rate π_t, and log consumption growth g_t, but not the shock s_{2t} to the Taylor rule. From observations of the state, we can estimate the autoregressive matrix A, and from observations of consumption growth we can estimate the shock coefficients d_1. We can also estimate a and b by projecting i_t and π_t on the state. With $a^\top = (\alpha d_1^\top + b^\top)A$ known, that leaves us to solve

$$a^\top = \tau b^\top + d_2^\top \tag{12}$$

for the Taylor rule's inflation parameter τ and shock coefficients d_2: n equations in the $n+1$ unknowns (τ, d_2). The identification problem is the same as Section 2: without further restrictions, the Taylor rule is not identified. This is Cochrane's conclusion in somewhat more general form.

We can, however, identify the monetary policy rule if we place one or more restrictions on the shock coefficients d_2. One example was mentioned earlier: choose d_1 and d_2 so that the shocks s_{1t} and s_{2t} are independent. We'll return to this shortly. Another example is a zero in the vector d_2 — what is traditionally termed an exclusion restriction. Suppose the ith element of d_2 is zero. Then the ith element of (12) is

$$a_i = \tau b_i.$$

If $b_i \neq 0$ (a regularity condition we'll come back to), this determines τ. Given τ, and our estimates of a and b, we can now solve (12) for the remaining components of d_2. We can do

the same thing with restrictions based on linear combinations. Suppose $d_2^\top e = 0$ for some known vector e. Then we find τ from $a^\top e = \tau b^\top e$. Any such linear restriction on the shock coefficient d_2 allows us to identify the Taylor rule.

Cochrane's example is a special case with shocks to consumption growth turned off: $d_1 = 0$. As a result, all the variation in inflation and the short rate comes from monetary policy shocks s_{2t}. Special case or not, the conclusion is the same: we need one restriction on d_2 to identify the (one) Taylor rule parameter τ. Note well: The restriction applies to the Taylor rule shock, and does not require a shock in the other equation.

3.2 An exponential-affine model

We take a similar approach to an exponential-affine model, adding a Taylor rule to an otherwise standard bond-pricing model. The model consists of a real pricing kernel, a Taylor rule, and the law of motion (3) for the state. The first two are

$$\begin{aligned} m_{t+1} &= -\rho - s_{1t} + \lambda^\top w_{t+1} \\ i_t &= r + \tau\pi_t + s_{2t}. \end{aligned}$$

As usual, the shocks are linear functions of the state: $s_{it} = d_i^\top x_t$. This model differs from the example in Section 2.2 in having a Taylor rule as well as a bond pricing relation.

We solve the model by the usual method. Given a guess $\pi_t = b^\top x_t$ for inflation, the nominal pricing kernel is

$$m_{t+1}^\$ = m_{t+1} - \pi_{t+1} = -\rho - (d_1^\top + b^\top A)x_t + (\lambda^\top - b^\top C)w_{t+1}.$$

The short rate follows from (7):

$$i_t = \rho - V_m/2 + (d_1^\top + b^\top A)^\top x_t,$$

where $V_m = (\lambda^\top - b^\top C)(\lambda - C^\top b)$. Equating this to the Taylor rule gives us $r = \rho - V_m/2$ and

$$d_1^\top + b^\top A = \tau b^\top + d_2^\top \quad \Rightarrow \quad b^\top = (d_1^\top - d_2^\top)(\tau I - A)^{-1}.$$

This is the unique stationary solution for b under the usual conditions.

Identification follows the same logic as the preceding example. Let us say, again, that we observe the state x_t, the short rate i_t, and inflation π_t. From the latter we estimate

$a = d_1^\top + b^\top A$ and b. That gives us n equations in the $n+1$ unknowns (τ, d_2). The model is identified only when we impose one or more restrictions on the parameter vector d_2 of the monetary policy shock. If, for example, the ith element of d_2 is zero, then τ follows from $a_i = \tau b_i$ as long as $b_i \neq 0$.

This model is a generalization of the previous one in which we've given the real pricing kernel a more flexible structure. Evidently the structure of the pricing kernel has little bearing on identification. We need instead more structure on the shock to the Taylor rule to compensate for not observing it.

3.3 Discussion

We have seen that one restriction suffices to identify the Taylor rule in these examples. The question is why this works. Is it similar to the use of exclusion restrictions in simultaneous equations models? Most econometrics textbooks illustrate exclusions of this kind with supply and demand. There we need a variable in one equation that's missing — excluded — from the other. To identify the demand equation, we need a variable in the supply equation that's excluded from demand. That's not the case here. We can identify the Taylor rule even when there are no shocks in the other equation if we have a restriction on the Taylor rule shock. The issue is not whether we have the right configuration of shocks across equations, but whether we observe them. When we don't observe the shock to the Taylor rule, we need additional structure in the same equation to deduce its parameters.

The same logic applies to Gertler's example in Section 2.3, where we used independence of the two shocks to identify the Taylor rule. Doesn't that involve shocks in the second equation? Well, yes, but the critical feature of independence here is the restriction it places on the Taylor rule shock. The shocks are uncorrelated, hence independent, if $d_2^\top V_x d_1 = 0$. But that's a linear restriction $d_2^\top e = 0$ on the coefficient vector d_2 of the Taylor rule shock. In this case, $e = V_x d_1$. The same holds for restrictions on innovations to the shocks. They're independent and uncorrelated if $d_2^\top V_w d_1 = 0$. In the representative agent model of Section 3.1, such a restriction is easily implemented. If we observe consumption growth (9), then we also observe s_{1t} and can use it to estimate d_1 and compute the restriction on d_2, the coefficient vector of the Taylor rule shock. We give numerical examples in Appendix B. These restrictions have no particular economic rationale in this case, but they illustrate how independence works. Similar "orthogonality conditions" appear throughout the New Keynesian literature.

A similar question arises with restrictions on interest rate coefficients. Suppose we know that a linear combination of interest rate coefficients is zero: $a^\top e = 0$ for some known e. Then (12) gives us a restriction connecting the Taylor rule shock and its coefficient vector: $\tau b^\top e + d^\top e = 0$. One interpretation is that we've used a restriction from another part of the model to identify the model. We would say instead that any such restriction on interest rate behavior implies a restriction on the Taylor rule, which identifies the policy rule for the usual reasons.

Another difference from traditional simultaneous equation methods is that single-equation estimation methods generally won't work. We need information about the whole model to deduce the Taylor rule. In the model of Section 3.1, for example, we need to estimate an interest rate equation to find a and an inflation equation to find b, before applying (12) to find τ. This reflects what Hansen and Sargent (1980, page 37) call the "hallmark" of rational expectations models: cross-equation restrictions connect the parameters in one equation to the parameters in the others.

4 A model with a Phillips curve

We apply the same logic to an example with a stronger New Keynesian flavor. We add a Phillips curve to the model and an output gap to the Taylor rule. Models with similar features are described by Carrillo, Feve, and Matheron (2007), Canova and Sala (2009), Christiano, Eichenbaum, and Evans (2005), Clarida, Gali, and Gertler (1999), Cochrane (2011), Gali (2008), King (2000), Shapiro (2008), Woodford (2003), and many others. Our contribution is a modest one: to add an asset pricing relation, which we think of as an illustration of how asset prices might be introduced into the analysis.

Despite the additional economic structure, the logic for identification is the same: we need restrictions on the shock coefficients to identify the Taylor rule. What changes is that we need two restrictions, one for each of the two parameters of the rule. We face similar issues in identifying the Phillips curve. If its shock isn't observed, we need restrictions to identify its parameters.

Our model consists of equations (6) and (8) plus

$$\begin{aligned} \pi_t &= \beta E_t \pi_{t+1} + \kappa g_t + s_{1t} \\ i_t &= r + \tau_1 \pi_t + \tau_2 g_t + s_{2t}. \end{aligned}$$

The first equation is a New Keynesian Phillips curve and the second is a Taylor rule, which now includes an output term. In addition, we have the law of motion (3) for the state, and the shocks $s_{it} = d_i^\top x_t$ for $i = 1, 2$.

We now have a two-dimensional rational expectations model in the forward-looking variables π_t and g_t. The solution of such models is described in Appendix A, but we can illustrate the challenges of identification with our usual guess-and-verify procedure. We guess a solution that includes $\pi_t = b^\top x_t$ and $g_t = c^\top x_t$. Then the pricing relation gives us

$$i_t = \rho - V_m/2 + a^\top x_t$$

with $a^\top = (\alpha c^\top + b^\top)A$ and $V_m = a^\top CC^\top a$. If we equate this to the Taylor rule and collect terms, we have $r = \rho - V_m/2$ and

$$a^\top = \tau_1 b^\top + \tau_2 c^\top + d_2^\top. \tag{13}$$

The Phillips curve implies

$$b^\top = \beta A b^\top + \kappa c^\top + d_1^\top. \tag{14}$$

As others have noted, the conditions for this to have a unique stationary solution are more stringent than before. We'll assume that they're satisfied.

Let's turn to identification. Suppose we, the economists, observe the state x_t, the interest rate i_t, the inflation rate π_t, and log consumption growth g_t. From them, we can estimate the autoregressive matrix A and the coefficient vectors (a, b, c). In the Taylor rule, the unknowns are the policy parameters (τ_1, τ_2) and the coefficient vector d_2 for the shock. If we observe the shock s_{2t}, equation (13) gives us n equations to solve for τ_1 and τ_2. As long as the dimension n of the state is at least two, the Taylor rule parameters are identified. If we do not observe the shock, then we need two restrictions on its coefficient vector d_2. The conclusion is the same as before, but with two parameters to identify we need two restrictions on the vector of shock coefficients.

The same logic applies to identifying the parameters of the Phillips curve. If we observe the shock s_{1t}, equation (14) gives us n equations to solve for the parameters β and κ. If we do not observe its shock s_{1t}, then two restrictions are needed to identify its two parameters. The identification problem for the Phillips curve has the same structure as the Taylor rule, although in practice they've been treated separately. For more on the subject, see the extensive discussions in Canova and Sala (2009), Gali and Gertler (1999), Nason and Smith (2008), and Shapiro (2008).

Standard implementations of New Keynesian models typically assume independent AR(1) shocks. See, for example, Gali (2008, ch 3). In our framework, this amounts to $n-1$ zero restrictions on the coefficient vectors d_i. That's generally sufficient to identify the structural parameters of the model, including those of the Taylor rule, but more restrictive than necessary. With respect to the Taylor rule parameter, each element i for which $d_{2i}=0$ leads, via equation (13), to an equation of the form $a_i = \tau_1 b_i + \tau_2 c_i$. As long as $(a_i, b_i, c_i) \neq (0,0,0)$, any two such equations will identify the Taylor rule parameters (τ_1, τ_2). Similar logic applies to the Phillips curve.

5 Observing the state

We have seen, in a number of examples, that to identify the parameter(s) of the Taylor rule, we need information about its shock. If we do not observe the shock, then we must impose restrictions on its form.

Our approach so far is predicated on observing the state. What happens if we observe the state indirectly? Or a noisy signal of the state? We show that neither changes the nature of the identification problem.

5.1 Indirect observation of the state

In many applications, the state variable is latent: we don't observe it directly, but we may be able to infer something about it. Examples include dynamic factor models, exponential-affine bond-pricing models, and structural vector autogressions.

In some of these models, the state is not completely determined: linear transformations of the state are observationally equivalent. Consider a model based on the linear law of motion (3). If the state x_t is not observed directly, then it is indistinguishable from a model with state $\hat{x}_t = Tx_t$, where T is an arbitrary square matrix of full rank. The law of motion is

$$\hat{x}_{t+1} \;=\; TAT^{-1}\hat{x}_t + TCw_{t+1} \;=\; \widehat{A}\hat{x}_t + \widehat{C}w_{t+1}. \tag{15}$$

where $\widehat{A} = TAT^{-1}$ and $\widehat{C} = TC$. By the same logic, the shocks are $s_{it} = \hat{d}_i^\top \hat{x}_t$ with $\hat{d}_i^\top = d_i^\top T^{-1}$.

We can transform the rest of the model the same way. Consider the representative agent model described in Section 3.1. For any choice of state $\hat{x}_t$, we can estimate the associated

parameter vectors for the observables. The short rate is then related to the state by $i_t = r + \hat{a}^\top \hat{x}_t$ with $\hat{a}^\top = a^\top T^{-1}$. Similarly, inflation is related by $\pi_t = \hat{b}^\top \hat{x}_t$ with $\hat{b}^\top = b^\top T^{-1}$. The Taylor rule then implies

$$\hat{a}^\top \quad = \quad \tau \hat{b}^\top + \hat{d}_2^\top,$$

the analog of equation (12) for the transformed state.

Consider the problem of identifying τ when we observe the transformed state $\hat{x}_t$ but not the original state x_t or the transforming matrix T. The identification problem is the same as before: we need one restriction on $\hat{d}_2$ to identify the single Taylor rule parameter τ. The only question is whether the restrictions we place on d_2 are intelligible when we translate them to $\hat{d}_2$. Consider a general linear restriction $d_2^\top e = 0$, where at least one element of e is non-zero. This restriction can be expressed $\hat{d}_2^\top T e = 0$, so the restricting vector is $\hat{e} = Te$. But if we don't know T, can we deduce $\hat{e}$?

There are at least two cases where the restrictions translate naturally to the transformed state. These cases have nearly opposite economic interpretations, so they offer a range of choices that can lead to identification.

In the first case, suppose the Taylor rule shock is uncorrelated with the other shock. Such "orthogonality conditions" are standard in the New Keynesian literature, where most shocks are independent of the others. We saw an example in Section 3.3. In terms of the original state x_t, the restriction takes the form $d_2^\top V_x d_1 = 0$. In terms of the transformed state $\hat{x}_t$, we have

$$\hat{d}_2^\top E(\hat{x}\hat{x}^\top)\hat{d}_1 \quad = \quad (d_2^\top T^{-1})(T V_x T^\top)(d_1^\top T^{-1})^\top \quad = \quad d_2^\top V_x d_1.$$

It's clear that this restriction is invariant to linear transformations of the state. We show how this might work in practice in Appendix B.

In the second case, optimal monetary policy dictates a connection between shocks s_{1t} and s_{2t}. See, for example, Gali (2008) and Woodford (2003). When a monetary authority minimizes an objective function, all variables of interest are affected by by s_{1t}. As a result, an optimal policy rule will make s_{2t} proportional to s_{1t}. If we express this by $s_{2t} = k s_{1t}$ for some constant k, it implies the restriction $d_2^\top - k d_1^\top = 0$ in terms of the original state variable x_t. In terms of the transformed state $\hat{x}_t$, the restriction is

$$\hat{d}_2^\top - k\hat{d}_1^\top \quad = \quad d_2^\top T^{-1} - k d_1^\top T^{-1} \quad = \quad (d_2^\top - k d_1^\top)T^{-1} \quad = \quad 0,$$

which is independent of the transformation T.

These restrictions apply to a given transformed state $\hat{x}_t$. One way to determine $\hat{x}_t$ is to choose a set of observable variables of the same dimension. We give some examples in Appendix B. Another is to select a specific transformation. This typically involves restrictions on both C and A. For the former, a common example is $T = C^{-1}$, which gives each state equation an independent innovation with unit variance. For the latter, the typical approach is to consider rotations: transformation by orthogonal matrices. Rotations that deliver lower-triangular $\widehat{A}$ are a common choice, but not a necessary one. Variants of this approach are used in dynamic factor models (Bai and Wang, 2012, and Bernanke, Boivin, and Eliasz, 2005), exponential-affine term structure models (Joslin, Singleton, and Zhu, 2011), and structural vector autoregressions (Leeper, Sims, and Zha, 1996, and Watson, 1994).

5.2 Noisy observation of the state

In other cases we may observe the state with noise, even after appropriate transformation. One example is a state with dimension $n > 2$ with observations of only the two variables i_t and π_t. Another is a larger collection of observables in which each variable is measured with error.

The Kalman filter deals with precisely this problem. Here we have three sets of variables. The state x_t follows the usual linear law of motion (3). Endogenous variables z_t are related to the state by $z_t = Bx_t$, where each row of B represents one of the model's endogenous variables. In our examples, z_t includes the one-period interest rate i_t, the inflation rate π_t, and so on. Finally, we have a set of variables y_t that we observe, which can include some or all of z_t.

A state-space model adds a measurement equation for the observables to the law of motion for the state:

$$y_t \quad = \quad Gx_t + Hv_t. \tag{16}$$

We posit measurement errors $v_t \sim \text{NID}(0, I)$ that are independent of w_t, but could easily incorporate arbitrary correlation between them. The observables are likely to include at least some of the endogenous variables, including those mentioned above, and possibly some of the state variables, too. In modern "data-rich" applications, the dimension of y_t is large relative to that of x_t, but the filter works even in the opposite situation.

Our question becomes how well observations of y_t substitute for observations of the state x_t. The question is a classical one and has a standard answer based on the Kalman filter: use y_t to estimate the state x_t and proceed as before using the estimate in place of the state. The Kalman filter is widely used in macroeconomics, so our treatment will be brief. Some of the essentials are given in Appendix C.

Two conditions guarantee that the state is, for our purposes, observed. First, we need the law of motion (3) to generate a state of dimension n, not a lower dimensional subspace. It's sufficient to assume that C has rank n, but when (3) is in companion form that's not realistic. In formal terms, we need (A, C) to be controllable, so that we can attain any x_t in $\mathbb{R}^n$ with some realization of disturbances w_t. If that's not the case, we should reduce the dimension of our problem. Second, and more critically, we need (A, G) to be observable; for the matrix

$$\mathcal{O} \quad = \quad \begin{bmatrix} G \\ GA \\ \vdots \\ GA^{n-1} \end{bmatrix}$$

to have rank n. Roughly speaking: the state x_t must generate enough variation in the observed variables y_t that we can (approximately) reverse engineer the state. With these two conditions, we can estimate the state from observations of the infinite history $y^t = (y_t, y_{t-1}, \ldots)$.

The state typically isn't recovered exactly, but the Kalman filter generates recursive estimates from the history of observables, which we might represent by

$$x_{t|t} \quad = \quad E\left(x_t | y^t\right).$$

The state itself therefore includes an error: $x_t = x_{t|t} + \varepsilon_t$ with the error ε_t orthogonal to the history y^t.

Now think about how we use the state: we need it to estimate the coefficient vectors tying endogenous variables to the state. Consider a specific endogenous variable z_{it} connected to the state by $z_{it} = b^\top x_t$, where $b^\top$ is the ith row of B. The inflation rate is a good example. Then

$$z_{it} \quad = \quad b^\top x_t \quad = \quad b^\top x_{t|t} + b^\top \varepsilon_t.$$

If z_{it} is one of the observables, then the last term is noise, orthogonal to the history y^t and therefore to $x_{t|t}$. We can estimate b by projecting z_{it} on $x_{t|t}$. By doing this for each observable variable, we can estimate coefficient vectors as before using $x_{t|t}$ in place of x_t.

5.3 Discussion

The challenges of observing the state raise a number of questions. One is how to observe the state in practice. In some cases, we may find that the state is spanned by some collection of observables. We think financial variables and forecasts are likely to be helpful here.

Models that have an exponential-affine structure, including all of the models in this paper, have a structure that makes forward rates natural state variables. If q_t^h is the price at date t if a claim to one dollar at $t+h$, then continuously-compounded forward rates are defined by $f_t^h = \log(q_t^h/q_t^{h+1})$. The short rate is $i_t = f_t^0$. In all of our models, the short rate takes the form (ignoring the intercept) $i_t = a^\top x_t$ and forward rates are $f_t^h = a^\top A^h x_t$. The vector f_t of the first h forward rates has the form

$$f_t \;=\; \left[\begin{array}{c} f_t^0 \\ f_t^1 \\ \vdots \\ f_t^{h-1} \end{array}\right] \;=\; \left[\begin{array}{c} a^\top \\ a^\top A \\ \vdots \\ a^\top A^{h-1} \end{array}\right] x_t \;=\; U x_t.$$

You may notice a resemblance to the observability condition here, in this case the observability of $(A, a^\top)$. If the dimension of the state and the number of forward rates are the same, and U is invertible, we find x_t by inverting the mapping from states to forward rates: $x_t = U^{-1} f_t$. If not, we can use them as inputs to the Kalman filter. Yields, defined by $i_t^h = h^{-1} \sum_{j=1}^h f_t^{j-1}$, are linear functions of forward rates, so they can be used the same way.

Forecasts have a similar mathematical structure. Even noisy measurements of conditional expectations can be helpful. Consider the forecast of variable z_{it} h periods ahead. If $z_{it} = b^\top x_t$ for some coefficient vector b, then a forecast might be expressed

$$F_t z_{i,t+h} \;=\; E_t z_{i,t+h} + u_{t,t+h} \;=\; b^\top A^h x_{t+h} + u_{t,t+h},$$

with some distribution of noise $u_{t,t+h}$. Surveys provide forecasts for a range of h's, each of which can be a useful input into a Kalman filter estimate of the state. Chernov and Mueller (2012), Chun (2011), and Kim and Orphanides (2012) are examples that use survey forecasts in state-space frameworks.

Another question is whether we can proceed if the observability condition fails, leaving us with hidden states. In some cases, it may be possible to proceed with the observable subspace of the state. Suppose, for example, that $x_{t|t}$ has dimension $k < n$. We can still

project endogenous variables on this estimate of the state and use these projections to estimate the Taylor rule. We simply need restrictions on the projection of the Taylor rule shock onto this subspace. The number of restrictions needed for identification does not change: one in the examples of Section 3, two in the example of Section 4.

6 Conclusion

Identification is always an issue in applied economic work, perhaps nowhere more so than in the study of monetary policy. That's still true. We have shown, however, that (i) the problem of identifying the systematic component of monetary policy (the Taylor rule parameters) in macro-finance models stems from our inability to observe the nonsystematic component (the shock to the rule) and (ii) the solution is to impose restrictions on the shock. We are left where we often are in matters of identification: trying to decide which restrictions are plausible, and which are not.

A Solution of rational expectations models

Consider the class of forward-looking linear rational expectations models,

$$\begin{aligned} z_t &= \Lambda E_t z_{t+1} + D x_t \\ x_{t+1} &= A x_t + C w_{t+1}. \end{aligned}$$

Here x_t is the state, Λ is stable (eigenvalues less than one in absolute value), A is also stable, and $w_t \sim \text{NID}(0, I)$. The goal is to solve the model and link z_t to the state x_t.

One-dimensional case. If z_t is a scalar and the shock is $s_t = d^\top x_t$, we have

$$z_t = \lambda E_t z_{t+1} + d^\top x_t. \tag{17}$$

Substituting repeatedly gives us

$$z_t = \sum_{j=0}^{\infty} \lambda^j d^\top E_t x_{t+j} = d^\top \sum_{j=0}^{\infty} \lambda^j A^j x_t = d^\top (I - \lambda A)^{-1} x_t.$$

The last step follows from the matrix geometric series if A is stable and $|\lambda| < 1$. Under these conditions, this is the unique stationary solution.

The same solution follows from the method of undetermined coefficients, but the rationale for stability is less clear. We guess $z_t = b^\top x_t$ for some vector b. Then the difference equation tells us

$$b^\top x_t = b^\top \lambda A x_t + d^\top x_t.$$

Collecting terms in x_t gives us $b^\top = d^\top (I - \lambda A)^{-1}$.

This model is close enough to the examples of Sections 2 and 3 that we can illustrate their identification issues in a more abstract setting. Suppose we observe the state x_t and the endogenous variable z_t, but not the shock s_t. Then we can estimate A and b. Equation (17) then gives us

$$b^\top = \lambda b^\top A + d^\top.$$

If x has dimension n, we have n equations in the $n + 1$ unknowns (λ, d); we need one restriction on d to identify the parameter λ.

Multi-dimensional case. If z_t is a vector, as in Section 4, we have

$$z_t = \Lambda E_t z_{t+1} + D^\top x_t.$$

Repeated substitution gives us

$$z_t = \sum_{j=0}^{\infty} \Lambda^j D A^j x_t.$$

That gives us the solution $z_t = Bx_t$ where

$$B \;=\; D + \Lambda BA \;=\; \sum_{j=0}^{\infty} \Lambda^j D A^j.$$

The solution is

$$\mathrm{vec}(B) \;=\; (I - A^\top \otimes \Lambda)^{-1} \mathrm{vec}(D).$$

See, for example, Anderson, Hansen, McGrattan, and Sargent (1996, Section 6) or Klein (2000, Appendix B). The same sources also explain how to solve rational expectations models with endogenous state variables.

B Numerical examples

We illustrate some of the issues raised in the paper with numerical examples of the model in Section 3.1: a representative agent with power utility and given consumption growth. We show how identification works when we observe the state and when we observe only a linear transformation of the state. In each case we use an orthogonality restriction on the shock to the Taylor rule.

We give the model a two-dimensional state and and choose parameter values $\tau = 1.5$, $\alpha = 5$, and

$$A \;=\; \begin{bmatrix} 0 & 1 \\ -0.05 & 0.9 \end{bmatrix}, \qquad C \;=\; \begin{bmatrix} 0.0078 & 0 \\ -0.0004 & 0.0003 \end{bmatrix}.$$

The consumption growth shock is governed by $d_1^\top = (1, 0)$. The monetary policy shock is $d_2^\top = (\delta, 1)$, with δ chosen to make s_2 uncorrelated with s_1. These inputs imply

$$V_x \;=\; \begin{bmatrix} 0.6432 & -0.0069 \\ -0.0069 & 0.0258 \end{bmatrix} \cdot 10^{-4},$$

so $d_2^\top V_x d_1 = 0$ implies $\delta = 0.0108$.

These are the inputs, the source of data that we can use to estimate the Taylor rule. The question is whether we can do that under different assumptions about observability of the state.

State observed. Suppose, first, that we observe the state x_t. Then we can use observations of the interest rate i_t and inflation rate π_t to recover the coefficient vectors

$$a \;=\; \begin{bmatrix} -0.3152 \\ 10.4566 \end{bmatrix}, \qquad b \;=\; \begin{bmatrix} -0.2174 \\ 6.3044 \end{bmatrix}.$$

Similarly, observations of consumption growth allow us to recover d_1. We do not observe the Taylor rule shock s_{2t}, so its coefficient vector d_2 remains unknown. A least squares estimate of the Taylor rule here gives us $\tau = 1.6510$ which, of course, isn't the value that generates the data.

Identification requires a restriction on d_2. We know d_2 satisfies the orthogonality condition $d_2^\top e = 0$ with $e = V_x d_1$. With our numbers, $e^\top = (0.6364, -0.0069) \cdot 10^{-4}$. We post-multiply (12) by e to get $a^\top e = \tau b^\top e$. This implies $\tau = 1.5$, the value we started with. We can now recover d_2 from the same equation: $d_2^\top = a^\top - \tau b^\top = (0.0108, 1.0000)$.

This is simply the procedure we outlined in Section 3.1, but it gives us a concrete basis of comparison for situations in which we don't directly observe the state.

State observed indirectly. Now suppose we don't observe the state, but we observe enough variables to deduce a linear transformation of the state. We consider two examples.

In our first example, we observe the interest rate and inflation rate and use them as our transformed state: $\hat{x}_t = (i_t, \pi_t)^\top$. Then $\hat{x}_t = T x_t$ with

$$T = \begin{bmatrix} a^\top \\ b^\top \end{bmatrix}.$$

This has something of the flavor of a structural vector autoregression, albeit a simple one. The law of motion for the transformed state is equation (15) with

$$\widehat{A} = \begin{bmatrix} -4.6170 & 9.1006 \\ -2.8044 & 5.5170 \end{bmatrix},$$

which is easily estimated.

Now consider identification. If we regress i_t, π_t, and log consumption growth on $\hat{x}_t$, we get the coefficient vectors

$$\hat{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \hat{d}_1 = \begin{bmatrix} 22.07 \\ -36.60 \end{bmatrix}.$$

From observations of $\hat{x}_t$, we can estimate its covariance matrix

$$V_{\hat{x}} = \begin{bmatrix} 0.2932 & 0.1775 \\ 0.1775 & 0.1075 \end{bmatrix} \cdot 10^{-3}.$$

Finally, the orthogonality condition in these coordinates is $\hat{d}_2^\top \hat{e} = 0$ with $\hat{e} = V_{\hat{x}} \hat{d}_1$. With our numbers, we have $\hat{e}^\top = (-0.2718, -0.1812) \cdot 10^{-4}$. As before, we apply the orthogonality condition to equation (12), which gives us $\tau = \hat{a}^\top \hat{e} / \hat{b}^\top \hat{e} = 1.5$, the number we started with. It is clear from this that we are still able to recover the Taylor rule from this linear transformation of the state.

In our second example, we use the first two forward rates as the state: $\hat{x}_t = (f_t^0 = i_t, f_t^1)^\top$. As we saw in Section 5.3, forward rates are connected to the original state x_t by $\hat{x}_t = Tx_t$, where

$$T = \begin{bmatrix} a^\top \\ a^\top A \end{bmatrix}.$$

The same series of calculations gives us τ.

C Elements of Kalman filtering

We outline some of the essential elements of Kalman filtering. Hansen and Sargent (2013) is a standard reference for economists. Anderson and Moore (1979) and Boyd (2009) are good technical references.

The starting point is the state-space system

$$\begin{aligned} x_{t+1} &= Ax_t + Cw_{t+1} \\ y_t &= Gx_t + Hv_t. \end{aligned}$$

Here $\{w_t\}$ and $\{v_t\}$ are vectors of independent standard normals — independent element by element, with each other, and across time. We refer to x_t as the state and y_t as the measurement. The state has dimension n, the measurement dimension p. In the technical literature, it's common to use w_t in place of w_{t+1} in the first equation, but since neither x_{t+1} nor w_{t+1} is ever observed, it's a convention without content.

Controllability and observability. We say (A, C) is *controllable* if

$$\mathcal{C} = \begin{bmatrix} C & AC & \cdots & A^{n-1}C \end{bmatrix}$$

has rank n. The word controllable is misleading in this context; the idea is simply that w_t generates variation across all n dimensions of x_t. We say (A, G) is *observable* if

$$\mathcal{O} = \begin{bmatrix} G \\ GA \\ \vdots \\ GA^{n-1} \end{bmatrix}$$

has rank n. The issue is similar to the previous one. The idea is that the condition guarantees us that observing the history of y_t is enough to come up with a full-rank estimate of x_t.

Controllability example. Here's one with x_t of dimension two and w_t of dimension one that fails:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}, \; C = \begin{bmatrix} c_1 \\ 0 \end{bmatrix} \;\Rightarrow\; \mathcal{C} = \begin{bmatrix} c_1 & a_{11}c_1 \\ 0 & 0 \end{bmatrix},$$

which has rank $1 < n = 2$. Here the innovation w_t never generates variation in x_{2t}, so we don't span the whole two-dimensional state. However, if a_{21} is nonzero we get controllability, because w_t affects x_{2t} with a one-period lag through its impact on x_{1t}. A similar example is an AR(2) in companion form.

Observability example. The logic is similar. Suppose x_t is n-dimensional and the nth column of G consists of zeros. There's no direct impact of the nth state variable on the observations y_t. In the bond-pricing literature, this might be a case in which one of the state variables doesn't appear in bond yields of any maturity. Nevertheless, the nth state variable might be (indirectly) observable if it feeds into other state variables: if a_{jn} is nonzero for some $j \neq n$. Here's an example similar to our previous one:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_1 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathcal{O} = \begin{bmatrix} g_1 & 0 \\ a_{11}g_1 & a_{12}g_1 \end{bmatrix},$$

which has rank two. Since $a_{12} = 0$, the condition fails.

D Stochastic volatility

We consider a representative agent model with stochastic volatility: the conditional variance is itself a stochastic process.

The model is similar to the one in Section 3.1, with these changes. The law of motion for x_t is now

$$x_{t+1} = Ax_t + C_t w_{t+1},$$

where

$$\begin{aligned} C_t &= C\sqrt{S_t}, \quad S_{iit} = \alpha_i + \beta_i v_t, \quad S_{ijt} = 0, \\ v_t &= (1-\nu)v + \nu v_{t-1} + \sigma_v \sqrt{v_{t-1}}\varepsilon_t^v, \quad \sigma_v^2/2 < (1-\nu)v \end{aligned}$$

The state now includes x_t and v_t. Shocks are $s_{it} = d_i^\top x_t + d_{iv} v_t$.

We solve by guess and verify. If we guess $\pi_t = b^\top x_t + b_v v_t$, the pricing relation gives us

$$i_t = \rho + \alpha g - a^\top C_t C_t^\top a/2 + a^\top x_t + \nu(b_v + \alpha d_{1v})v_t + (b_v + \alpha d_{1v})^2 v_t/2,$$

where a is the same as before. The interest rate is therefore $i_t = a^\top x_t + a_v v_t$, with $a_v = (b_v + \alpha d_{1v})^2/2$. Equating it to the Taylor rule implies

$$\begin{aligned} a^\top &= \tau b^\top + d_2^\top \\ a_v = -a^\top CBC^\top a/2 + \nu(b_v + \alpha d_{1v}) + (b_v + \alpha d_{1v})^2/2 &= \tau b_v + d_{2v}, \end{aligned}$$

where B is a diagonal matrix with elements β_i.

With respect to identification this has the same structure as the constant volatility version. What's new is the presence of volatility as a state variable. If we think monetary policy does not respond to volatility, then we can use the restriction $d_{2v} = 0$ to achieve identification. Normandin and Phaneuf (2004) do something similar.

References

Anderson, Brian D.O. and John B. Moore, 1979, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice Hall.

Anderson, Evan W., Lars Peter Hansen, Ellen R. McGrattan, and Thomas J. Sargent, 1996, "Mechanics of forming and estimating dynamic linear economies," in H.M. Amman, D.A. Kendrick, and J. Rust, *Handbook of Computational Economics, Volume I*, 171-252.

Ang, Andrew, and Monika Piazzesi, 2003, "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics* 50, 745-787.

Bai, Jushan, and Peng Wang, 2012, "Identification and estimation of dynamic factor models," manuscript, April.

Bernanke, Ben, Jean Boivin, and Piotr S. Eliasz, 2005, "Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach," *Quarterly Journal of Economics* 120, 387-422.

Boyd, Stephen P., 2009, "Lecture slides for EE363: Linear Dynamical Systems," posted online at http://www.stanford.edu/class/ee363/.

Canova, Fabio, and Luca Sala, 2009, "Back to square one: Identification issues in DSGE models," *Journal of Monetary Economics* 56, 431-449.

Carrillo, Julio A., Patrick Feve, and Julien Matheron, 2007, "Monetary policy or persistent shocks: a DGSE analysis," *International Journal of Central Banking* 3, 1-38.

Chernov, Mikhail, and Philippe Mueller, 2012, "The term structure of inflation expectations," *Journal of Financial Economics*, 106, 367-394.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans, 2005, "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy* 113, 1-45.

Chun, Albert, 2011, "Expectations, bond yields, and monetary policy," *Review of FInancial Studies*, 24, 208-247.

Clarida, Richard, Jordi Gali, and Mark Gertler, 1999, "The science of monetary policy," *Journal of Economic Literature* 37, 1661-1707.

Cochrane, John H., 2011, "Determinacy and identification with Taylor rules," *Journal of Political Economy* 119, 565-615.

Gali, Jordi, 2008, *Monetary Policy, Inflation, and the Business Cycle*, Princeton, NJ: Princeton University Press.

Gali, Jordi, and Mark Gertler, 1999, "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics* 44, 195-222.

Gallmeyer, Michael, Burton Hollifield, and Stanley Zin, 2005, "Taylor rules, McCallum rules and the term structure of interest rates," *Journal of Monetary Economics* 52, 921-950.

Hansen, Lars, and Thomas Sargent, 1980, "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and Control* 2, 7-46.

Hansen, Lars, and Thomas Sargent, 1991, "Two difficulties in interpreting vector autoregressions," *Rational Expectations Econometrics*, ed. L.P. Hansen and T.J. Sargent, San Francisco: Westview Press.

Hansen, Lars, and Thomas Sargent, 2013, *Recursive Models of Dynamic Linear Economies*, Princeton: Princeton University Press.

Joslin, Scott, Anh Le, and Kenneth Singleton, 2013, "Gaussian macro-finance term structure models with lags," *Journal of Financial Econometrics*, forthcoming.

Joslin, Scott, Kenneth Singleton, and Haoxiang Zhu, 2011, "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 24, 926-970

Kim, Don, and Athanasios Orphanides, 2012, "Term structure estimation with survey data on interest rate forecasts," *Journal of Financial and Quantitative Analysis*, 47, 241-271.

King, Robert, 2000, "The new IS-LM model: Language, logic, and limits," *Federal Reserve Bank of Richmond Economic Quarterly*, 45-103.

Klein, Paul, 2000, "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics & Control* 24, 1405-1423.

Leeper, Eric, Christopher Sims, and Tao Zha, 1996, "What does monetary policy do?" *Brookings Papers on Economic Activity*, 1-63.

Moench, Emanuel, 2008, "Forecasting the yield curve in a data-rich environment: A no-arbitrage factor augmented VAR approach," *Journal of Econometrics* 146, 26-43.

Nason, James M., and Gregor W. Smith, 2008, "The New Keynesian Phillips curve: Lessons from single-equation econometric estimation," *Federal Reserve Bank of Richmond Economic Quarterly*, 361-395.

Normandin, Michel, and Louis Phaneuf, 2004, "Monetary policy shocks: Testing identification conditions under time-varying volatility," *Journal of Monetary Economics* 51, 1217-1243.

Rudebusch, Glenn, and Tao Wu, 2008, "A macro-finance model of the term structure, monetary policy, and the economy," *Economic Journal* 118, 906-926.

Shapiro, Adam Hale, 2008, "Estimating the New Keynesian Phillips curve: A vertical production approach," *Journal of Money, Credit and Banking* 40, 627-666.

Sims, Christopher, and Tao Zha, 2006, "Were there regime switches in US monetary policy?" *American Economic Review* 96, 54-81.

Smith, Josephine, and John Taylor, 2009, "The term structure of policy rules," *Journal of Monetary Economics* 56, 907-917.

Watson, Mark, 1994, "Vector autoregressions and cointegration," in *Handbook of Econometrics, Volume IV*, ed. R.F. Engel and D.L. McFadden, Elsevier.

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