

Identifying monetary policy in macro-finance models*

David Backus,[†] Mikhail Chernov,[‡] Stanley Zin,[§] and Irina Zviadadze[¶]

July 17, 2018

Abstract

Identification problems arise in New Keynesian and macro-finance models when the Taylor rule includes both responses to observable variables like inflation and output, and a shock unseen by economists. Identification of the rule's parameters requires additional restrictions on this unobserved shock. We demonstrate how this can be accomplished in a macro term structure model using only long-run neutrality restrictions consistent with a wide variety of theories. The resulting Taylor rule is comparable to those commonly found in the literature. The unobserved shock is closely related to the slope factor of empirical term structure models.

JEL Classification Codes: E43, E52, G12.

Keywords: forward-looking models; identification; monetary policy; affine bond-pricing.

* 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. Earlier versions of this work by the first three authors circulated with the title "Identifying Taylor rules in macro-finance models". We are grateful to the two anonymous referees and the Editor, Andrew Karolyi, for their thoughtful comments, which have helped us improve the manuscript. We also thank Rich Clarida, Patrick Feve, Kei Kawai, Guillaume Roussellet, Ken Singleton, Gregor Smith, Roman Šustek, Tommy Sveen, and Tao Zha for comments on earlier drafts, and participants in seminars at ADEMU/CERGE-EI, BI Norwegian Business School, New York University, Stony Brook University, the Swedish House of Finance, Queen Mary University of London, and the Toulouse School of Economics. The latest version of the paper is available at https://sites.google.com/site/mbchernov/BCZ_trident_latest.pdf.

[†] Stern School of Business, New York University, and NBER; Deceased (June 12, 2016).

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

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

[¶] Stockholm School of Economics, and CEPR; irina.zviadadze@hhs.se.

1 Introduction

The field of macro-finance has the potential to give us deeper insights into macroeconomics and macroeconomic policy by combining information about aggregate quantities with 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 in the so-called “Taylor rule” of Taylor (1993). And when policy involves temporary departures from simple rules, such as recent *unconventional* monetary policy, it may even introduce an important source of financial-market risk.

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 and monetary policy shocks. If we see that the short-term interest rate rises with inflation, does that reflect the policy of the central bank, a temporary departure from that policy, the valuations of private agents, or something else? Can we tell the difference?

Some prominent scholars argue that the answer is no. Cochrane (2011, page 606) puts it this way: “The crucial Taylor rule parameter is not identified in the new-Keynesian model.” He devotes most of his paper to making the case. Joslin, Le, and Singleton (2013, page 597) make a related point about interpretations of estimated bond pricing models. In their words: “Several recent studies interpret the short-rate equation as a Taylor-style rule. ... However, without imposing additional economic structure, ... the parameters are not meaningfully interpretable as the reaction coefficients of a central bank.” Canova and Sala (2009), Carrillo, Feve, and Matheron (2007), and Iskrev (2010) also question aspects of the identification of New Keynesian models.

The contribution of this paper is twofold. First, we revisit this conclusion and propose an approach to monetary policy identification. Using a state-space model as a unifying framework, we find that what matters for identification of the parameters of the Taylor rule

is our knowledge of the structure of the shock to that one equation alone, specifically how it responds to changes in the current state. In equilibrium, all endogenous variables including the interest rate, the inflation rate, and real output growth, depend on the state. But the unobservable Taylor rule shock also depends on the state. Without more structure on that shock we have no way of decomposing an observed interest rate changes into the various components of the Taylor rule, such as changes in inflation, output growth, and the policy shock.

Second, we develop a specific set of identifying restrictions in the context of what Joslin, Priebsch, and Singleton (2014) call a “macro-finance affine term structure model” – an arbitrage-free term structure model that includes as factors, macroeconomic variables such as GDP growth and inflation. The restrictions are based on long-run neutrality of the monetary policy shock. Since short-run responses to policy are unconstrained, these identifying assumptions are consistent with a broad class of structural macroeconomic models. Estimation of the model results in estimates of the parameters of the systematic response of the Taylor rule to inflation and output growth, and estimates of the unobservable shock to the Taylor rule. We explore the dynamic correlations of both macro and financial variables to these policy shocks.

We find that the shock to the Taylor rule has relatively small and short-lived correlations with real GDP growth and inflation. On the other hand, it has a more substantial positive effect on the short rate and a negative effect on risk premiums embedded in long bond yields. In other words, it is closely related to the traditional “slope factor” of reduced-form term structure models, and the “conundrum” often expressed by monetary policy makers.

2 The identification problem

We use two examples to illustrate the nature of identification problems in macro-finance models with Taylor rules. The first comes from Cochrane (2011). The second is an affine bond-pricing model. The critical ingredient in each is what we observe. As the New Keynesian literature, 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.

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 here only on inflation):

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

$$i_t = \tau_0 + \tau \pi_t + s_{2t}. \tag{2}$$

Here i_t is the (one-period) nominal interest rate, r is the real interest rate (assumed constant in this example), π_t is the inflation rate, and s_{2t} is a monetary policy shock (the need for subscript 2 will be apparent shortly). The Taylor rule parameter $\tau > 1$ describes how aggressively the central bank responds to inflation.

The state of the economy is an n -dimensional vector x_t and the Taylor rule shock is a linear function of it: $s_{2t} = d_2^\top x_t$. The state follows the autoregressive process

$$x_{t+1} = Ax_t + Bw_{t+1}, \tag{3}$$

with A stable, B lower triangular with positive diagonal elements, and disturbances $\{w_t\} \sim$

NID(0, I). (All vectors and matrices conform in size to the dimension of x_t .) The unconditional covariance matrix of x_t , V_x , is the solution to $V_x = AV_xA^\top + BB^\top$. To maintain a clear distinction between theory and empirical applications, we assume that x_t is exogenous, and the parameters A and B are part of the structure of the economy. Since x_t and w_t have the same dimension, assuming that B is lower triangular is without loss of generality.

For simplicity we will assume that x_t is observable in our examples, hence, it can be used for estimation. Appendix A demonstrates that this has no bearing on the identification problem of interest. Replacing x_t with an estimate \hat{x}_t obtained through Kalman filtering using strictly observable variables, adds noise but does not change the fundamental structure of the identification problem.

We solve the model by standard methods; see Appendix B. Equations (1) and (2) imply the forward-looking difference equation or rational expectations model

$$E_t\pi_{t+1} = \tau_0 - r + \tau\pi_t + s_{2t}.$$

The solution for inflation has the form $\pi_t = b_0 + b^\top x_t$ for some coefficient vector b and intercept b_0 to be determined. Then $E_t\pi_{t+1} = b_0 + b^\top E_t x_{t+1} = b_0 + b^\top A x_t$. Lining up terms, we see that b satisfies $b^\top A = \tau b^\top + d_2^\top$ which we can solve for b ,

$$b^\top = -d_2^\top(\tau I - A)^{-1}. \tag{4}$$

Likewise, b_0 satisfies $b_0 = \tau_0 - r + \tau b_0$, which implies $b_0 = (\tau_0 - r)/(1 - \tau)$. 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 $i_t = a_0 + a^\top x_t$ where a satisfies $a^\top = b^\top A$, which implies

$$a^\top = -d_2^\top(\tau I - A)^{-1}A. \tag{5}$$

Likewise, we can solve for the intercept, $a_0 = (\tau_0 - \tau r)/(1 - \tau)$.

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_{2t} as the residual. The problem is evident in equation (4). There is no sense in which we can condition on π_t while allowing s_{2t} to vary, since they both depend on x_t . We need to distinguish the effect of the state on the interest rate (represented by τb^\top) from the effect of the shock (represented by d_2^\top). Least squares delivers a coefficient of $\text{Var}(\pi_t)^{-1} \text{Cov}(\pi_t, i_t) = (b^\top V_x b)^{-1} b^\top V_x a = \tau + (b^\top V_x b)^{-1} b^\top V_x d_2$, which is not in general equal to τ . Note that since the bias is not the result of the endogeneity of π_t , but rather the endogeneity of the omitted variable, s_{2t} , other regression methods like two-stage least squares will suffer the same fate. Instrumental variables estimation would require a valid instrument for s_{2t} , which we discuss below.

What about the intercept b_0 ? We can estimate b_0 with the mean of π_t , but since it depends on three unknown parameters, τ_0 , r , and τ , it neither separately identifies τ_0 and r , nor does it help in identifying τ . (Throughout the paper we use *identified* to mean that we can distinguish a unique value of a parameter from local alternatives, *i.e.*, the parameter is *locally point identified*.)

How then can we estimate τ ? The critical issue is whether we observe the shock s_{2t} . If 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_{2t} , then we can estimate the parameter vector d_2 . We now have all the components of b in (4) but τ , which we can infer. The Taylor rule parameter τ is not only identified, it is over-identified. If x_t has dimension n , we have n equations that each determine τ . In addition, we could identify r with the mean of $i_t - \pi_t$, then given τ and r , infer the value of τ_0 from either the mean of i_t or π_t . Note, however, that identifying τ_0 is always predicated on first identifying τ .

Suppose, however, that we do not observe s_{2t} . This is precisely the situation considered throughout the New Keynesian literature. If we do not observe s_{2t} , then we cannot estimate the coefficient vector d_2 . Equation (4) then has n equations in $n + 1$ unknowns, the n shock coefficients d_2 and the Taylor rule parameter τ , hence it cannot be solved for a unique values of these parameters. This is a concrete example of the identification issue faced by all empirical work based on forward-looking models with unobserved shocks.

Cochrane's example illustrates the challenges we might face in identifying the parameters of the Taylor rule, but given its stylized nature – most theoretical and empirical research uses more complicated models – one might question its generality. In particular, the shock to the Taylor rule in Cochrane's example is the only shock in the model, which is not a common assumption. We know that simple models in which both supply and demand are driven by a single shock suffer identification issues, so could the introduction of more shocks in other equations help with identification in this case too? For example, a model with a state-dependent real interest rate would introduce a shock in Cochrane's first equation, so that the example becomes

$$\begin{aligned} i_t &= r + E_t \pi_{t+1} + s_{1t} \\ i_t &= \tau_0 + \tau \pi_t + s_{2t}, \end{aligned}$$

where r is now the mean of the real interest rate. Can the presence of this additional shock identify the Taylor rule?

Suppose that s_{1t} and s_{2t} are independent – a common assumption in New Keynesian models. 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 estimate r and τ_0 as before, and back out the shock s_{2t} . It would seem, therefore, that the introduction of a shock to the first equation solved the identification problem.

This example, though perfectly logical, is misleading in one respect. Identification of τ is not a consequence of the presence of a shock in the first equation, but rather identification follows from the *restriction* we placed on the Taylor rule shock, s_{2t} , by assuming its independence with s_{1t} . Independence is just one example of a restriction on s_{2t} that can identify τ . Later examples will show how other restrictions can serve the same purpose.

Another general feature that this example demonstrates is that identifying τ and backing out the unobserved shock s_{2t} are complementary activities: if we can do one, we can do the other. We will return to this feature in the empirical application of Section 4.

2.2 An affine extension

The identification problem arises because we cannot distinguish the pricing relation for the nominal interest rate (1) from the Taylor rule (2). But what about long-term interest rates? Unlike the short rate, they are not the direct focus of monetary policy, yet they will respond to expected inflation. Can they help identify the Taylor rule?

We explore this possibility in an affine model, which has become the standard framework for empirical term-structure research. In the macro-finance branch of this literature, the state includes macroeconomic variables like inflation and output growth. Examples include Ang and Piazzesi (2003), Moench (2008), Rudebusch and Wu (2008), Smith and Taylor (2009), Chernov and Mueller (2012), Jardet, Monfort, and Pegoraro (2013), Hamilton and Wu (2012), Joslin, Le, and Singleton (2013), and Joslin, Priebsch, and Singleton (2014).

The model starts with the specification of the log pricing kernel,

$$-\log m_{t+1}^{\$} = a_0 + a^\top x_t + \lambda_t^\top \lambda_t / 2 + \lambda_t^\top w_{t+1}, \quad (6)$$

where $\lambda_t = \lambda_0 + \lambda x_t$. The nominal (log) pricing kernel $m_t^{\$}$ is connected to the real (log) pricing kernel m_t and inflation by $m_t^{\$} = m_t - \pi_t$. The one-period nominal interest rate is

then

$$i_t = -\log E_t \exp(m_{t+1} - \pi_{t+1}) \quad (7)$$

$$= -\log E_t \exp(m_{t+1}^{\$}) = a_0 + a^\top x_t. \quad (8)$$

Equations (7) and (8) are a more complex version of the Fisher equation – equation (1) – popular in the empirical finance literature because of its ability to capture dynamic risks. We can think of Cochrane’s example as a linear approximation of (7) with $E_t m_{t+1}$ equal to a constant r , which suppresses the impact of both real and nominal risk. Note that we can still estimate a_0 with the mean of i_t , and a by projecting i_t onto the state x_t .

Does the more general model in (8) help to identify parameters of a Taylor rule? Using observations on inflation π_t and the state x_t , we can estimate the intercept b_0 and coefficient vector b connecting the two: $\pi_t = b_0 + b^\top x_t$. Then the Taylor rule implies

$$i_t = \tau_0 + \tau\pi_t + s_{2t} = \tau_0 + \tau b_0 + \tau b^\top x_t + d_2^\top x_t.$$

Equating our two interest rate relations then gives us $a^\top = \tau b^\top + d_2^\top$ and $a_0 = \tau_0 + \tau b_0$. It is clear, now, that we have the same difficulty we had in the previous example: if we do not know the shock parameter d_2 , we cannot infer τ from estimates of a and b . If x_t has dimension n , we have n equations to solve for $n + 1$ unknowns (d_2 and τ).

The situation is no different if we rotate the state vector so that its first element is the inflation rate, $z_t = Tx_t = [\pi_t \ x_{2t} \ \dots \ x_{nt}]^\top$. In that case, it would be tempting to interpret an equation like (8) linking i_t to z_t , *i.e.* $i_t = \tilde{a}^\top z_t$ as a Taylor rule, with the first element of \tilde{a} equal to the inflation coefficient τ , and $[0 \ \tilde{a}_2 \ \dots \ \tilde{a}_n]^\top$ as the shock coefficients d_2 . However, the first element of \tilde{a} is actually $\tau + d_{21}$, and the vector we would like to interpret as d_2 is actually $[0 \ d_{22} \ d_{23} \ \dots \ d_{2n}]$. In other words, τ and d_2 are still not identified.

Sims and Zha (2006, page 57) anticipated this issue: “The Fisher relation is always lurking in the background ... one might easily find an equation that had the form of the ... Taylor rule ... 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?” And finally, Joslin, Le, and Singleton (2013, page 583) draw a very clear conclusion about affine macro-finance models: “the parameters of a Taylor rule are not econometrically identified.”

We find it more natural to interpret (8) as an asset pricing relation, analogous to (1), and complete the model by adding the Taylor rule (2). Here, too, it is evident that we cannot distinguish the systematic component of monetary policy (represented by τb) from shocks to policy (represented by d_2) without more information. Generalizing the asset pricing relation from (1) to (8) has no effect on this conclusion.

Now we add long-term interest rates to the model. Given the pricing kernel and the linear transition equation (3), the absence of arbitrage implies that the date- t price, $q_t^{(h)}$, of an h -period default-free pure-discount bond with a face value of 1, is log-linear in the state:

$$-\log q_t^{(h)} = \mathcal{B}_0^{(h)} + \mathcal{B}^{(h)} x_t, \quad (9)$$

where

$$\begin{aligned} \mathcal{B}^{(h)} &= a^\top (I - A^*)^{-1} (I - A^{*h}) \\ \mathcal{B}_0^{(h)} &= a_0 + \mathcal{B}_0^{(h-1)} + \mathcal{B}^{(h-1)} A_0^* - \mathcal{B}^{(h-1)} B B^\top \mathcal{B}^{(h-1)\top} / 2 \\ A^* &= A - B \lambda \\ A_0^* &= -B \lambda_0, \end{aligned}$$

where A_0^* and A^* are, respectively, the risk-neutral mean and persistence of x_t associated

with the pricing kernel $m_{t+1}^{\$}$.

Since $i_t = y_t^{(1)}$, we have initial conditions $\mathcal{B}_0^{(1)} = a_0$ and $\mathcal{B}^{(1)} = a^\top$. Continuously compounded yields for $h > 1$ are also linear in the state:

$$y_t^{(h)} = -\log q_t^{(h)}/h = \frac{1}{h} \left(\mathcal{B}_0^{(h)} + \mathcal{B}^{(h)} x_t \right). \quad (10)$$

We saw that even though we can estimate $a = \mathcal{B}^{(1)}$ by projecting i_t onto x_t , that is not sufficient to identify the Taylor rule parameters on which a is based. We can also estimate $\mathcal{B}^{(h)}$ by projecting $hy_t^{(h)}$ onto x_t . But since $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1} (I - A^{*h})$ this adds nothing to the identification problem: given an estimate for b , any configuration of the parameters τ and d_2 that leaves a unchanged, will also leave $\mathcal{B}^{(h)}$ unchanged. (Note that A^* can be identified without reference to the Taylor rule, since $A^* = A - B\lambda$ and values for A and B can be identified from the the dynamics of the state and λ from the cross-section of yields.) Information in longer-term yields may be important for efficient estimation of other parameters of the model or the state vector when it is unobserved, but the term structure of interest rates does not provide a solution to the problem of identifying the Taylor rule.

This is a more general property shared by the mathematical structure of forecasts. Denote a forecast of variable z_t at a horizon of h by $f_t^{(h)} = E_t z_{t+h}$, where $z_t = a_0 + a^\top x_t$, for some arbitrary coefficients (a_0, a) . The transition equation then implies $f_t^{(h)} = a_0 + a^\top A^h x_t$. (Recall that yields are just averages of forward rates, which are forecasts of future short-term interest rates using the risk-neutral dynamics of x_t .) A collection of forecasts can be used the same way we used yields. Or we could add forecasts to our collection of observables. Chernov and Mueller (2012), Chun (2011), and Kim and Orphanides (2012) are examples that use survey forecasts in state-space frameworks. The forecasts add useful information in all of these applications, but they do not resolve the identification problem.

3 Macro-finance models with Taylor rules

Both Cochrane’s original example and the affine term-structure extension of that model have a *reduced-form* quality. That is, given an abstract specification of the pricing kernel, equilibrium will impose restrictions on observables. But the dependence of the pricing kernel on deeper parameters of preferences and technologies is left unspecified. Macro-finance models often include this additional structure. Could that added structure provide the basis for identifying the Taylor rule?

We use two representative-agent examples, one without any monetary non-neutralities and one with a Phillips curve, and a so-called structural VAR – a time-series model with parameters identified by imposing specific economic structure – to explore this question. We find that just like the more reduced-form examples above, deeper structural models alone will not restrict the Taylor rule shock, hence will not provide the identifying information we seek.

3.1 A representative-agent model without frictions

A growing body of macro-finance research combines representative-agent asset pricing with a rule governing monetary policy. Gallmeyer, Hollifield, Palomino, and Zin (2007) is a good example. We simplify their model, using power utility instead of Epstein-Zin recursive preferences, and a constant-volatility transition equation for the state.

The model consists of the previous bond-pricing relation, equation (7), plus

$$m_t = -\rho - \alpha g_t \tag{11}$$

$$g_t = g + s_{1t} \tag{12}$$

$$i_t = \tau_0 + \tau \pi_t + s_{2t}. \tag{13}$$

Again, equations (7) and (13) mirror the two equations of Cochrane's example. Equations (11) and (12) characterize the real pricing kernel based on a representative-agent's optimization problem and the equilibrium condition in a frictionless endowment economy. The first is the logarithm of the marginal rate of substitution of a power utility agent with time preference parameter ρ , relative risk aversion parameter α , and log consumption growth g_t . The second connects fluctuations in the agent's log consumption growth to the growth of a random endowment $g + s_{1t}$ with mean growth of g . As in Section 2, the state x obeys the transition equation (3) and shocks are linear functions of it: $s_j = d_j^\top x$ for $j = 1, 2$.

Once again, the solution combines equilibrium asset pricing with a forward-looking difference equation. We posit a solution of the form $\pi_t = b_0 + b^\top x_t$, then solve (7) to get

$$i_t = a_0 + a^\top x_t, \tag{14}$$

with

$$\begin{aligned} a_0 &= \rho + \alpha g + b_0 - V_m/2 \\ a^\top &= (\alpha d_1^\top + b^\top)A \\ V_m &= (\alpha d_1^\top + b^\top)BB^\top(\alpha d_1 + b). \end{aligned}$$

Note the obvious similarity of the short rate equation (14) and the example of Section 2.2.

Equating (13) and (14) gives us

$$(\rho + \alpha g + b_0 - V_m/2) + (\alpha d_1^\top + b^\top)Ax_t = (\tau_0 + \tau b_0) + (\tau b^\top + d_2^\top)x_t.$$

Lining up similar terms, we have $b_0 = (\tau_0 - \rho - \alpha g + V_m/2)/(1 - \tau)$ and

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

As before, this gives us a unique stationary solution when A is stable and $\tau > 1$.

Now consider identification. Suppose that along with the state x_t , we also observe 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 its mean g and the shock coefficients d_1 . We can also estimate a and b by projecting i_t and π_t onto the state. With a and b known, that leaves us to solve

$$a^\top = \tau b^\top + d_2^\top \tag{15}$$

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 in Cochrane's original example; without further restrictions, the Taylor rule coefficient τ is not identified. Similarly, the mean of $i_t - \pi_t$ will identify $\rho + \alpha g - V_m/2$, but the value of τ_0 cannot be identified from b_0 without first identifying τ .

We can, however, identify the monetary policy rule if we place one or more restrictions on its shock coefficients d_2 . One such case was mentioned earlier: choose d_1 and d_2 so that the shocks s_{1t} and s_{2t} are independent. We will return to this shortly. Another example is a zero in the vector d_2 – what is traditionally termed an exclusion restriction. Suppose the j th element of d_2 is zero. Then the j th element of (15) is

$$a_j = \tau b_j.$$

As long as $b_j \neq 0$, this determines τ . Given τ , and our estimates of a and b , we can now solve (15) for the remaining components of d_2 .

We can do the same thing with any restriction on d_2 . 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 restriction on the shock coefficient

d_2 allows us to identify the Taylor rule. An exclusion restriction sets the j th element of e equal to 1 and the others equal to 0. Independence of s_{1t} and s_{2t} is also a special case with $e = V_x d_1$. Likewise, independence of the innovations to these shocks is a linear restriction on d_2 with $e = BB^\top d_1$. These restrictions have no particular economic rationale at this point, but they illustrate how independence works as an identifying assumption. Similar “orthogonality conditions” for unobserved shocks appear throughout applied econometrics. In the New Keynesian literature (discussed below), the shocks are typically low-order ARMA models, assumed to be independent of the rest of the model. Independence serves as a set of restrictions on the shocks that identify the model parameters, including the parameters of the monetary policy rule.

Cochrane’s example has the shocks to consumption growth turned off: $d_1 = 0$. But as we now know, since this has nothing to do with s_{2t} , the identification conclusion is the same. We need one restriction on d_2 to identify the Taylor rule parameter τ . The general lesson is clear: Taylor rule identification requires a restriction that applies to the Taylor rule shock. Adding more structure to the other equations is irrelevant unless that structure implies such a restriction.

3.2 A model with a Phillips curve

Monetary policy could affect real variables such as consumption or output growth, as well as nominal variables such as the interest rate and inflation. To explore whether such non-neutralities impact the Taylor rule identification problem, we add a Phillips curve to the representative agent model of Section 3.1 and an output gap to the Taylor rule. As a result, output growth g_t becomes endogenous. New Keynesian 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), Iskrev (2010), King (2000), Shapiro (2008), Smets and Wouters (2007), Woodford (2003), and many others.

As in Gallmeyer, Hollifield, and Zin (2005), our model consists of the pricing relation in equation (7), the real pricing kernel in equation (11), and

$$\pi_t = \beta E_t \pi_{t+1} + \kappa g_t + s_{1t} \quad (16)$$

$$i_t = \tau_0 + \tau_1 \pi_t + \tau_2 g_t + s_{2t}. \quad (17)$$

The first equation is a Phillips curve. The second is a Taylor rule, which now includes an output growth term. In addition, we have the transition equation (3) for the state and the shocks $s_{jt} = d_j^\top x_t$ for $j = 1, 2$.

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 B. As others have noted, the conditions for a unique stationary solution are more stringent than before. We will assume that they are satisfied.

To solve the model we guess a solution for the two endogenous variables $\pi_t = b_0 + b^\top x_t$ and $g_t = c_0 + c^\top x_t$ (see Appendix B). Then the pricing relation gives us

$$i_t = \rho + \alpha c_0 + b_0 - V_m/2 + a^\top x_t$$

with $a^\top = (\alpha c^\top + b^\top)A$ and $V_m = (\alpha c^\top + b^\top)BB^\top(\alpha c + b)$. If we equate this to the Taylor rule and collect terms, we have $b_0 = (\tau_0 - \rho - \alpha c_0 + V_m/2 + \tau_2 c_0)/(1 - \tau_1)$ and

$$a^\top = \tau_1 b^\top + \tau_2 c^\top + d_2^\top. \quad (18)$$

Similarly, the Phillips curve implies

$$b^\top = \beta b^\top A + \kappa c^\top + d_1^\top. \quad (19)$$

The intercept terms are $b_0 = [\tau_0 - \rho + V_m/2]/[1 - \tau_1 + \alpha(1 - \beta)/\kappa - \tau_2(1 - \beta)/\kappa]$, and $c_0 = [\tau_0 - \rho + V_m/2]/[(1 - \tau_1)\kappa/(1 - \beta) + \alpha - \tau_2]$. Our goal is to solve these equations for the Taylor rule parameters (τ_0, τ_1, τ_2) .

Along with the state x_t , assume we observe the interest rate i_t , the inflation rate π_t , and log consumption growth g_t , but not the shocks (s_{1t}, s_{2t}) to the Phillips curve and Taylor rule, respectively. From the observables, we can estimate the autoregressive matrix A , the coefficient vectors (a, b, c) , and intercepts (a_0, b_0, c_0) . In equation (18), representing the Taylor rule, the unknowns are the policy parameters (τ_1, τ_2) and the coefficient vector d_2 for the shock. If we do not observe the shock, we need two restrictions on its coefficient vector d_2 to identify (τ_1, τ_2) .

Despite the additional economic structure, the logic for identification is the same: we need restrictions on the shock coefficients d_2 to identify the Taylor rule. All that changes is the *number* of restrictions we need. Since the Taylor rule has two parameters, we now need two restrictions.

The same logic applies to identifying the parameters of the Phillips curve. If we do not observe the shock s_{1t} , then two restrictions are needed to identify the parameters β and κ . The identification problem for the Phillips curve has the same structure as the Taylor rule, although in practice they have been treated separately. See the extensive discussions in Canova and Sala (2009), Gali and Gertler (1999), Iskrev (2010), Nason and Smith (2008), and Shapiro (2008).

Standard implementations of New Keynesian models typically use independent AR(1) or ARMA(1,1) shocks. See, for example, Gali (2008, ch 3) and Smets and Wouters (2007). In

our framework, an independent AR(1) amounts to $n - 1$ zero restrictions on the coefficient vectors d_j : none of the other state variables affect the shock. That is generally sufficient to identify the structural parameters of the model, including those of the Taylor rule. With respect to the Taylor rule, each element j for which $d_{2j} = 0$ leads, via equation (18), to an equation of the form $a_j = \tau_1 b_j + \tau_2 c_j$. As long as $(b_j, c_j) \neq (0, 0)$, any two such equations will identify the Taylor rule parameters (τ_1, τ_2) . Similar logic applies to the Phillips curve.

3.3 Vector autoregressions

There is an influential body of research that uses vector autoregressions (VARs) to characterize the dynamic effects of exogenous shocks to the economy, often interpreted as shocks to monetary policy. See, for example, the many studies cited by Christiano, Eichenbaum, and Evans (1999, Sections 3 and 4) and Watson (1994, Section 4). As Watson (1994, page 2898) puts it: “[VARs] provide answers to the ‘impulse’ and ‘propagation’ questions often asked by macroeconomists.”

This approach to identification of shocks is different than what we have discussed so far. But since VAR models fit nicely into a state-space framework, we can compare the two approaches directly, and ask whether VAR identification can also be interpreted as Taylor rule identification. That is, will the identification of a so-called policy shock in the context of a VAR correspond to the identification of the Taylor rule shock s_{2t} , and hence solve the problem posed by Cochrane?

A dynamic model in applied monetary economics might be expressed as a system of equations of the form

$$Dy_t = C_1 y_{t-1} + w_t, \tag{20}$$

where y_t is a vector of observable macroeconomic variables that includes the object of

policy (such as the short-rate in the Taylor rule example), w_t is a vector of fundamental economic shocks of the same dimension, $\{w_t\} \sim \text{NID}(0, I)$, and D is nonsingular. (Note that for simplicity, we arbitrarily truncate the number of the lags in (20) after the first and we suppress intercept terms, but the extension to any finite-order of lags with intercepts is straightforward.) By multiplying (20) by D^{-1} , this model can be written as a vector autoregression (VAR),

$$y_t = A_1 y_{t-1} + u_t, \tag{21}$$

where $A_1 = D^{-1}C_1$, $\{u_t\} \sim \text{NID}(0, \Sigma)$, and $\Sigma = D^{-1}(D^{-1})^\top$. Standard regression methods provide estimates of A_1 and Σ . Note, however, that each element of u_t is now a linear combination of *all* the fundamental shocks, w_t . The dynamic response of y_{t+j} , $j \geq 0$, to a unit change in the exogenous shock w_t , other things equal, is given by $A_1^j D^{-1}$. But since D is not separately identified from A_1 and Σ , these fundamental *impulse responses* also remain unidentified unless we impose more structure on D .

Identifying restrictions imposed on D can take a variety of forms. A popular choice is a recursive identification scheme that assumes D is lower triangular. But one might also consider other restrictions on D such as the long-run restrictions of Blanchard and Quah (1989), or the sign restrictions of Uhlig (2005). And for some questions it is sufficient to restrict just certain blocks of D , rather than the entire matrix. For our purposes, these will all work in much the same way: they result in a VAR that identifies the endogenous impulse response to a fundamental exogenous shock.

To see the relationship to Taylor rule identification, consider Cochrane's example of Section 2.1. Define the vector of endogenous variables $y_t = [\pi_t \ i_t]^\top$. The equilibrium restrictions in

(4) and (5) imply that y_t is linear in x_t ,

$$y_t = Rx_t, \quad R = \begin{bmatrix} b^\top \\ a^\top \end{bmatrix} = \begin{bmatrix} -d_2^\top (\tau I - A)^{-1} \\ -d_2^\top (\tau I - A)^{-1} A \end{bmatrix}. \quad (22)$$

For simplicity, assume that the dimension of x_t , is the same as y_t . The dynamics of x_t in equation (3) imply an equation for y_t of same form as (21),

$$y_t = (RAR^{-1})y_{t-1} + RBw_t.$$

Note that without further restrictions on the model, changes to either shock, w_{1t} or w_{2t} , affect both π_t and i_t directly, and the impulse responses of π_{t+j} and i_{t+j} , $j \geq 0$, through $(RAR^{-1})^j RB$. In other words, there is nothing in the model that implies that RB has the lower-triangular structure commonly used in VAR shock identification.

Christiano, Eichenbaum, and Evans (1999, Section 6) make a similar point: “Why did we not display or interpret the [relevant equation of a VAR as a monetary policy rule]? The answer is that these parameters are not easily interpretable.” Why, you might ask? They continue: “In [two of our] examples the decision maker reacts to a variable that is not in the econometrician’s data set. The policy parameters are a convolution of the parameters of the rule ... and the projection of the missing data onto the econometrician’s data set.” That’s the essence of our problem: “the missing data” (the shock s_{2t}) is not in “the econometrician’s data set.” As a consequence, the “convolutions” we see in a VAR will be the regression matrix $A_1 = RAR^{-1}$ and the covariance matrix $\Sigma = RB B^\top R^\top$.

The structure of the Taylor rule in (2), leads Christiano, Eichenbaum, and Evans (1999, Section 4) to consider what seems like a natural definition of a “policy shock” in a VAR model: the change in the policy variable that is unrelated to changes in other endogenous variables. In the context of Cochrane’s example, the policy variable is i_t and the other

endogenous variable is π_t , so that the policy shock is defined such that the upper-right element of RB is zero. This associates the policy shock with the structural shock w_{2t} , which will now have a direct effect only on the policy variable i_t , and *not* the other endogenous variable π_t . The other structural shock, w_{1t} , can have a direct effect on both variables. Note that the policy shock will still have an effect on the dynamics of π_{t+j} , $j > 0$, through the upper-right element of the matrix $(RAR^{-1})^j$, since assuming that RB is lower triangular does not imply that RAR^{-1} is also lower triangular.

It is important to note that this definition of a policy shock is *not* an implication of the Taylor rule in (2), which places no restrictions on the correlation structure between the Taylor rule shock, s_{2t} , and inflation, π_t . The VAR definition of a policy shock, therefore, is an *additional* assumption.

Does this additional assumption identify τ and d_2 ? The answer is no, at least not without more information. None of the VAR parameters correspond directly to the Taylor rule parameters. The VAR will result in estimates of the matrices A_1 and Σ , which in theory are both functions of the deeper structural parameters, τ , d_2 , A , and B : A_1 is equal to RAR^{-1} , the lower-triangular Choleski decomposition of Σ is now equal RB , and R depends on τ , d_2 , and A through (22). But we cannot recover these structural parameters from VAR estimates alone, hence we cannot identify τ and d_2 without more information.

4 Empirical application

The analysis thus far gives us a deeper understanding of how models can fail to meet the requirements for Taylor rule identification. Although it does not lead directly to some optimal solution to this fundamental problem, it does give us a framework for understanding and evaluating the content of any particular identification scheme. We need to restrict the factor loadings on the Taylor rule shock, d_2 , without much theoretical guidance. It would

seem sensible, therefore, to find restrictions that are consistent with as broad a class of theoretical models as possible. In this section we detail how this can be done in the context of a macro term structure model and demonstrate how long-run neutrality of the policy shock for both real quantities and real asset prices provide exactly the type of robust identification we seek. We begin by estimating an affine macro term structure model, then we show how long-run neutrality restrictions lead to identification. Finally, we explore the empirical properties of our identified Taylor rule and the unobservable policy shock.

4.1 Affine term structure estimation

We estimate an affine term structure model as outlined in Section 2.2 using quarterly US data from 1982Q3 to 2017Q2. Along with discount bond yields of various maturities, the system of equations is expanded to include two key macroeconomic variables, real GDP growth and inflation, as in a wide variety of macro term-structure models. The sample period is chosen to be sufficiently far from Volcker’s regime shift to allow for short-term adjustments and to establish credibility for Taylor-rule-based monetary policy. We use continuously compounded default-free pure-discount bond yields as measured by Gürkaynak, Sack, and Wright (2007), along with real GDP growth rates from the National Income and Product Accounts, and core CPI inflation from the Bureau of Labor Statistics. Figure 1 displays these standard data for our sample period.

GMM estimation of the parameters of this model is detailed in Appendix C. The results are summarized in Table 1. Our system of equations has seven variables: the nominal interest rate, 4 longer nominal yields corresponding to maturities of 8, 20, 28, and 40 quarters, real GDP growth, and inflation. We assume that these seven variables are driven by four common factors, *i.e.*, $n = 4$. The linear relationships between yields and the state variable is given in Section 2.2. In addition, we assume that the macro factors are also linear functions of the state variable: inflation is given by $\pi_t = b_0 + b^\top x_t$ and GDP growth is

given by $g_t = c_0 + c^\top x_t$. Since these are unrestricted relationships, our reduced-form model is consistent with a wide variety of both New Keynesian and neoclassical structural macro models.

The affine term structure model has 45 parameters: 26 parameters governing the dynamics of the state space (A and B), 9 more parameters governing the nominal pricing kernel (A^* , a_0 and λ_0), and 10 more parameters governing the macro variables (b_0 , c_0 , b , and c). These parameters are just-identified using 45 first- and second-moment restrictions implied by the theory. Since the goal in this empirical exercise is Taylor rule identification rather than an exhaustive analysis of the term structure, we do not impose additional over-identifying restrictions on our estimation.

The estimated parameter values have a number of noteworthy features. The risk-neutral dynamics encoded in the non-zero elements of A^* , are more persistent than the actual dynamics of the statespace in A . The absolute values of the 4 eigenvalues of A , which govern the persistence in the process for x_t , are 0.9761, 0.8648, 0.6329, and 0.6329 (the last pair corresponds to complex conjugates). The diagonals of A^* are 0.9862, 0.9250, 0.8609, and 0.6775, and they are very precisely estimated – a consequence of the additional information in the cross-equation restrictions implied by the absence of arbitrage – and all significantly different than zero.

Many of the off-diagonal elements of the matrix B are significantly different from zero, suggesting that our vector of state variables does not have orthogonal innovations. This will play an important role when we explore dynamic correlations below.

The inflation rate has significant and positive loadings (the values of b) for all 4 factors. On the other hand, real GDP growth has significant and positive loadings (the values of c) for only the first three factors. This suggests that the model is capturing a purely nominal feature in the data.

The average price of risk, λ_0 , appears to be different from zero for three of the four factors: the second factor (the second most persistent under the risk-neutral distribution), is both small and statistically insignificant. The average price of risk for the first factor (the most persistent factor) is negative and close to zero, but it is statistically significant, whereas two less persistent factors have average prices of risk that are larger and with opposite signs.

4.2 Identification through long-run neutrality restrictions

As we have seen, when we add a Taylor rule to the model, our just-identified system becomes under-identified, and we need to add identifying restrictions. We use the version of the rule that depends on both inflation and real GDP growth given in equation (17). We need to impose at least two more restrictions on the parameters of our empirical model to identify the two key parameters, τ_1 and τ_2 .

Macro models that include a shock to the Taylor rule do not attach any deep structural interpretation to that shock. Conceptually, there are many possibilities that seem reasonable. For example, real-time measurement error in inflation and output, a high-frequency risk premium (the policy rate is typically an overnight federal funds rate whereas macro models work at a quarterly frequency), changes in the preferences of the median voter on policy-setting committees like the FOMC, or the market microstructure of interactions between the central bank and its network of private brokers, all sound like plausible models of a policy shock. However, each of these structural models is likely to result in a shock that has a different relationship with the state variable, and hence, a different impact on both macro variable and asset prices. Since we are not imposing much economic structure beyond the absence of arbitrage on our model, we would like the identification restrictions we introduce to be equally robust. We would ideally like our reduced-form model to be consistent with as large a set of structural models as possible.

The first restriction we impose is a standard assumption in both New Keynesian and neo-classical models: shocks to monetary policy may have short-run real consequences but they do not have a permanent impact on the level of real output. This is analogous to the long-run restriction used in structural VAR models popularized by Blanchard and Quah (1989). Denote the log of the level of real GDP as y_t and note that in our specification, y_t is a unit-root process, $y_t = y_{t-1} + g_t$, which implies

$$\begin{aligned} y_{t+n} &= y_{t-1} + g_t + g_{t+1} + \cdots + g_{t+n} \\ &= y_{t-1} + nc_0 + c^\top [x_t + x_{t+1} + \cdots + x_{t+n}]. \end{aligned}$$

The conditional covariance of g_{t+j} and s_{2t} for $j \geq 0$ is

$$\begin{aligned} \text{Cov}_{t-1}(g_{t+j}, s_{2t}) &= E_{t-1}[c^\top x_{t+j} x_t^\top d_2] \\ &= c^\top A^j B B^\top d_2, \end{aligned} \tag{23}$$

which implies

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}_{t-1}(y_{t+n}, s_{2t}) &= \sum_{j=0}^{\infty} c^\top A^j B B^\top d_2 \\ &= c^\top (I - A)^{-1} B B^\top d_2. \end{aligned}$$

The assumption that this long-run covariance is zero implies a linear restriction on d_2 ,

$$c^\top (I - A)^{-1} B B^\top d_2 = 0, \tag{24}$$

that we can use to help identify the parameters of the Taylor rule. In principal, it would seem reasonable to add other real quantities to the model, *e.g.*, consumption and investment, and impose similar restrictions on their long-run response to s_{2t} . But given the high correlation of other real variables with real output, such restrictions would not add much

new information beyond (24). Instead, we make use of our pricing-kernel model and explore asset prices as the source for additional identifying restrictions.

The second restriction we impose is the natural analog to (24) applied to real asset prices: shocks to monetary policy may have short-run real consequences for asset markets, but they do not have a permanent impact on the level of real asset prices. Since we have already assumed that long-run real quantities are unaffected by the policy shock s_{2t} , what this assumption adds is a restriction on the real marginal utility of wealth. Following Alvarez and Jermann (2005), Hansen and Scheinkman (2009), and Hansen (2012), we decompose shocks to the real marginal utility of wealth, *i.e.*, the real pricing kernel, into permanent and transitory components. In Appendix D we detail this decomposition for our affine model. We show that the permanent component of the real pricing kernel is also a linear function of the state, x_t . The assumption that the current policy shock is uncorrelated with the permanent component of the real marginal utility of wealth implies another linear restriction d_2 that takes the form

$$[(b^\top - a^\top)(I - A^*)^{-1} - \lambda_0^\top B^{-1}]BB^\top d_2 = 0. \quad (25)$$

Note that given the dependence on A^* and λ_0 , a well-specified pricing-kernel model is essential to the construction of this second restriction.

4.3 Empirical properties of the Taylor rule

Given these two additional restrictions, we can identify the Taylor rule parameters, τ_0 , τ_1 , and τ_2 , as well as the factor loadings for the shock to the Taylor rule, d_2 . Note that once we have identified τ_1 and τ_2 , identification of τ_0 and d_2 follows from equation (17): $\tau_0 = a_0 - \tau_1 b_0 - \tau_2 c_0$ and $d_2 = a - \tau_1 b - \tau_2 c$. Estimates of these parameters are presented in Table 1.

It is both reassuring and surprising that although we have adopted a novel approach for identification, the estimates for the Taylor rule parameters are quite conventional. The coefficient on π_t is 2.5771, which safely satisfies the Taylor-principal stability condition, $\tau_1 > 1$. It is larger than Taylor’s original specification of $\tau_1 = 1.5$, suggesting a somewhat more aggressive policy. Our estimate of τ_2 is 0.3972, which is quite close to Taylor’s value of $\tau_2 = 0.5$, however, the units are difficult to compare directly as we use output growth rather than deviations from a potential-output trend.

The estimates for the factor loadings for the policy shock, d_2 , are significant for only three of the four factors. The loading on the first (the most persistent) factor is small and insignificant, which stands in contrast to real GDP growth which had a significant loading on that factor, but not the fourth. Again, this reinforces the interpretation that one of the dimensions of the model is purely nominal.

To get a sense of the sign and the magnitude of these policy shocks, the upper panel of Figure 2 plots the Taylor rule with and without the shock s_{2t} . We see large and persistent negative shocks in the early 1990s followed by large and persistent positive shocks in the late 1990s. The zero lower bound shows up as large positive shocks in 2009-2010, followed by persistent negative values for the shock over the last five years of our sample.

Since we have identified the Taylor rule shock as part of our dynamic macro-term structure model, we can use that model to get better understanding of how policy shocks affect the economy. The Taylor rule shock is affected by the entire vector of innovations, w_t . To measure the average dynamic response of an endogenous variable, say $\pi_t = b_0 + b^\top x_t$, to a Taylor rule shock, $s_{2t} = d_2^\top x_t$, we calculate a dynamic covariance

$$\begin{aligned} \text{Cov}_{t-1}(\pi_{t+j}, s_{2t}) &= \text{Cov}_{t-1}(b^\top x_{t+j}, d_2^\top x_t) \\ &= b^\top A^j B B' d_2, \end{aligned}$$

for $j \geq 0$. In Table 3 we graph this covariance – in the form of a correlation coefficient – for GDP growth, inflation, the short rate, the two-year and ten-year discount bond yields, and the ten-year forward premium, *i.e.*, the difference between the ten-year forward rate and the short rate. (The factor loadings for the ten-year forward rate are $\mathcal{B}^{(40)} - \mathcal{B}^{(39)}$.)

GDP growth has a small positive correlation with the policy shock, but it is neither significant nor persistent. Similarly, inflation has a small negative correlation that is marginally significant at short horizons, but which also dies out quickly. Medium-term bond yields – in this case the two-year yield – has virtually no correlation with the policy shock. However, at both the short and the long ends of the maturity structure, the response is quite different. The short rate has a significant positive correlation and the long yield – in this case the 10-year yield has a significant negative correlation with the shock to the Taylor rule, that is still significant at a one-year horizon. The combined effect is a near perfect negative correlation between the slope of the yield curve – measure here by the 10-year forward premium – and the monetary policy shock, and this correlation dies out even more slowly, as depicted in the lower right panel of Figure 3. In other words, the policy shock exhibits a strong negative correlation with the risk premium on long bonds. A positive shock to the Taylor rule results in a higher short-rate and a lower forward rate on long bonds.

Without a structural model linking risk premiums to Taylor rule shocks we can only speculate on the source of these strong correlations. However, these results do cast some light on the so-called conundrum that has puzzled policy makers in the past: long rates often move in ways that appear disconnected from the policy rate. Alan Greenspan (1994) attributed the increase in long yields in early 1994 to expectations of increases in future values of g_t and π_t : “In early February, we thought long-term rates would move a little higher as we tightened. The sharp jump in [long] rates that occurred appeared to reflect the dramatic rise in market expectations of economic growth and associated concerns about possible inflation pressures.” What Figure 2 suggests is that since the tightening in late 1993 and

early 1994 was warranted by values of GDP growth and inflation alone, *i.e.*, small values of s_{2t} , monetary policy was not a big contributor to the risk premiums that were driving the long-end of the yield curve.

A decade later Greenspan (2005) once again voiced puzzlement regarding the behavior of long yields: “Long-term interest rates have trended lower in recent months even as the Federal Reserve has raised the level of the target federal funds rate by 150 basis points. Historically, even distant forward rates have tended to rise in association with monetary policy tightening... For the moment, the broadly unanticipated behavior of world bond markets remains a conundrum.” What Figure 2 suggests once again is that the monetary policy tightening of 2004-05 was warranted by values of GDP growth and inflation alone. The implied small values of s_{2t} suggests that whatever was driving risk premiums at the long-end of the yield curve, it had little to do with policy shocks.

In contrast, Bernanke (2013) expresses a greater appreciation for the relationship between risk premiums and the large positive Taylor rule shocks implied by the zero lower bound in 2009-10: “Two changes in the nature of this interest rate risk have probably contributed to a general downward movement of the term premium in recent years. First, the volatility of Treasury yields has declined, in part because short-term rates are pressed up against the zero lower bound and are expected to remain there for some time to come. Second, the correlation of bond prices and stock prices has become increasingly negative over time, implying that bonds have become more valuable as a hedge against risks from holding other assets.” Although it is not obvious from just looking at the lower panel of Figure 1 that risk premiums on long bonds were small around this period, it is true that on average, positive values of s_{2t} are coincident with small values of risk premiums.

Backus and Wright (2007) summarize the situation: “We think the evidence points to a declining term premium as the primary source of the recent fall in long forward rates.... In this sense, we follow a long line of work in suggesting that expectations-hypothesis intuition,

based on constant term premiums, is likely to be misleading not only in this case but more generally.” And they call for a research program to fill this gap in our knowledge: “The next step, in our view, should be to develop models in which macroeconomic policy and behavior can be tied more directly to the properties of interest rates. We might want to know, for example, whether changes in the volatility of output or in the nature or communication of monetary policy had an impact on the behavior of long rates in the recent past. Neither of this is possible with existing models.” The work summarized in this section is a small step in this direction.

5 Conclusion

Identification is always an issue in applied economic work, perhaps nowhere more so than in the study of monetary policy. That is still true. We have shown, however, that (i) the problem of identifying the systematic component of monetary policy (the Taylor rule parameters) in New Keynesian and 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 estimate a macro-finance term structure model and use it, along with long-run restrictions that are consistent with a wide-variety of both New Keynesian and Classical monetary models, to identify the parameters of the Taylor rule and the process for the shocks to this policy rule. We find that the shock has little or no effect on real GDP growth, and only a small effect on inflation. However, it has a more substantial effect on the short interest rate and a persistent negative correlation with risk premiums at the long end of the maturity structure. To arrive at a deeper understanding of the causes and consequences of these empirical facts, the challenge we now face is to develop plausible structural models of the sources of both policy shocks and risk premiums that are capable of accounting for the

strong connection we see in the data between the Taylor rule shock and the slope of the yield curve.

A State-space models and Kalman filtering

Our examples assume that the state is observable and can form the basis for econometric estimation. But what happens if we observe the state indirectly? Or observe only a noisy signal of the state? These questions lead us to state-space models, in which we add to the transition equation for the state a so-called measurement or observation equation connecting an unseen state to a collection of observable variables.

State-space models provide a method for estimating the unobserved state through Kalman filtering. Does this affect any of our conclusions about identification? The answer is no. The Kalman filter is a recursive algorithm for computing the distribution of x from observations of y , and through x the distribution of y ; see, among many others, Anderson and Moore (1979, Chapters 3-4), Boyd (2009, Lecture 8), and Hansen and Sargent (2013, Chapter 8).

The classic state-space framework consists of the transition equation (3) and a related measurement equation for observables,

$$v_t = C_0 + Cx_t + Du_t. \tag{26}$$

The measurement errors $u_t \sim \text{NID}(0, I)$ are independent of the w 's.

A state-space model is a description of the distribution of observables v , but this distribution is invariant to linear transformations of the state x . Consider a model with state $\tilde{x} = Tx$, where T is an arbitrary square matrix of full rank. The transformed model is

$$\begin{aligned} \tilde{x}_{t+1} &= TAT^{-1}\tilde{x}_t + TBw_{t+1} = \tilde{A}\tilde{x}_t + \tilde{B}w_{t+1} \\ v_t &= C_0 + CT^{-1}\tilde{x}_t + Du_t = C_0 + \tilde{C}\tilde{x}_t + Du_t, \end{aligned}$$

where $\tilde{A} = TAT^{-1}$, $\tilde{B} = TB$, and $\tilde{C} = CT^{-1}$. The observational equivalence of models based on x and \tilde{x} raises new identification issues that are not related to those we discussed

earlier. These issues are generally managed by choosing a canonical form. See, for example, the extensive discussions in De Schutter (2000), Gevers and Wertz (1984), and Hinrichsen and Pratzel-Wolters (1989). Variants of this approach are used in dynamic factor models (Bai and Wang, 2015; Bernanke, Boivin, and Eliasch, 2005; Boivin and Giannoni, 2006; Stock and Watson, 2012) and affine term structure models (Joslin, Singleton, and Zhu, 2011). Given a canonical form we can generally estimate the matrices (A, B, C_0, C, D) .

One of the intermediate outputs of the estimation process described below is a series of estimates (conditional means) of the state:

$$\hat{x}_{t|s} = E(x_t | v^s),$$

where $v^s = (v_s, v_{s-1}, \dots)$ is a history of measurements. The Kalman filter produces, among other things, $\hat{x}_{t|t}$ and $\hat{x}_{t|t-1}$. Given such estimates of the state, we can identify the parameters of a forward-looking model just as in the case of observable x_t . The errors in these estimates are orthogonal to observed variables by construction, so projections of observables on estimates of the state produce the same parameter values in population.

Note that our theoretical models will restrict C_0 and C to be functions of deeper parameters. Standard Kalman filtering arguments lead to a multivariate normal one-step-ahead conditional distribution, $p(v_{t+1} | v^t; A, B, C_0, C, D)$, for the observable vector

$$v_{t+1} | v^t \sim \mathcal{N}(C_0 + C\hat{x}_{t+1|t}, C\Sigma_t C^\top + D),$$

where

$$\begin{aligned}
\hat{x}_{t|s} &= E(x_t|v^s) \\
&= (A - K_t C)\hat{x}_{t|t-1} + K_t v_t \\
K_t &= A\Sigma_t C^\top (C\Sigma_t C^\top + D)^{-1} \\
\Sigma_{t+1} &= A\Sigma_t A^\top + BB^\top - A\Sigma_t C^\top (C\Sigma_t C^\top + D)^{-1} C\Sigma_t A,
\end{aligned}$$

and $\hat{x}_{1|0}$ is set equal to 0, the mean of the ergodic distribution of x_t , and $\Sigma_0 = \text{Var}(x)$ its variance, $\text{Var}(x) = A\text{Var}(x)A^\top + BB^\top$. From these conditional distributions, we can construct the likelihood function for a sample of size T :

$$p(v^T|A, B, C_0, C, D) = \prod_{t=1}^T p(v_{t+1}|v^t; A, B, C_0, C, D).$$

In term-structure applications, v_t is partitioned into a vector of macroeconomic variables, z_t , whose relationship to x_t is unconstrained, and a vector of multi-period bond yields, y_t , whose relationship to x_t is constrained by the absence of arbitrage across bonds of different maturity. We partition the parameters of the observation equation, C_0 , C , and D , into C_{0z} , C_z , D_z , C_{0y} , C_y , and D_y , corresponding to the partition of v_t . Since the parameters in C_{0z} , C_z , and D_z are unconstrained by the term-structure model, they could be estimated by MLE without further complication. When x_t is unobservable, the information provided by bond yields, y_t , will be helpful nonetheless. However, the elements of C_{0y} and C_y , are nonlinear functions of A , B , and the parameters of the pricing kernel in equation (6). Estimation of the full system, therefore, requires imposing no-arbitrage restrictions on C_{0y} and C_y .

The free parameters of the model are then a_0 , C_{0z} , C_z , A , and the non-zero elements of B , A^* and D , which implies a likelihood function

$$p(v^T|a_0, C_{0z}, C_z, A, B, A^*, D) = \prod_{t=1}^T p(v_{t+1}|v^t; a_0, C_{0z}, C_z, A, B, A^*, D),$$

that can form the basis for method of moments and maximum likelihood estimation or a Bayesian posterior distribution.

B Solutions of forward-looking models

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

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

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 . (For simplicity, we have suppressed intercept terms so z_t and x_t should be interpreted as deviations from long-run means.)

One-dimensional case. If z_t is a scalar we have

$$z_t = \lambda E_t z_{t+1} + d^\top x_t, \tag{27}$$

for a some vector d . Repeated substitution 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 obvious. We guess $z_t = h^\top x_t$ for some vector h . The difference equation

tells us

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

Collecting coefficients of x_t gives us $h^\top = d^\top (I - \lambda A)^{-1}$.

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 $d^\top x_t$. Then we can estimate A and h . Equation (27) then gives us

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

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 3.2, repeated substitution gives us

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

That gives us the solution $z_t = H x_t$ where

$$H = \sum_{j=0}^{\infty} \Lambda^j D A^j = D + \Lambda H A$$

or

$$\text{vec}(H) = (I - A^\top \otimes \Lambda)^{-1} \text{vec}(D).$$

See, for example, Anderson, McGrattan, Hansen, 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.

C Macro Term Structure Model Estimation

We adopt a canonical form of a state space model in which an unobserved state, x_t , follows the process

$$x_t = Ax_{t-1} + Bw_t,$$

where $w_t \sim iid \mathcal{N}(0, I)$, and B is a lower triangular matrix with positive elements on its diagonal. Calendar time from t to $t + 1$ is a quarter of a year (3 months). The observed macro variables are inflation, π_t , and GDP growth, g_t , which are linear functions of the unobserved state:

$$\begin{aligned}\pi_t &= b_0 + b^\top x_t \\ g_t &= c_0 + c^\top x_t.\end{aligned}$$

Bond yields also depend linearly on x_t , and their loadings satisfy the arbitrage-free restrictions given in equation (10).

Following Hamilton and Wu (2012), we assume that bond yields for $n = 1$, *i.e.*, the short rate, and $n = 8, 20, 28$, where maturity is also measured in quarters, can be used to identify the state. Stack these 4 variables into the vector $z_t = [i_t \ y_t^{(8)} \ y_t^{(20)} \ y_t^{(28)}]^\top$. Rewrite the dynamics for the state as

$$\begin{aligned}z_t &= R_0 + Rx_t = R_0 + R(Ax_{t-1} + Bw_t) \\ &= \tilde{R}_0 + \tilde{A}z_{t-1} + \tilde{B}w_t,\end{aligned}\tag{28}$$

where $R_0 = [a_0 \ \mathcal{B}_0^{(8)}/8 \ \mathcal{B}_0^{(20)}/20 \ \mathcal{B}_0^{(28)}/28]$, $R = [a \ \mathcal{B}^{(8)}/8 \ \mathcal{B}^{(20)}/20 \ \mathcal{B}^{(28)}/28]^\top$,

$\tilde{A} = RAR^{-1}$, $\tilde{R}_0 = (I - \tilde{A})R_0$, $\tilde{B} = RB$, and we adopt the normalization $a^\top = [1 \ 1 \ 1 \ 1]$. This rotation of the state space eliminates the need for Kalman filtering the unobserved state.

A 3×1 vector of the two macro variables along with the long bond yield, $y_t = [\pi_t \ g_t \ y_t^{(40)}]^\top$, is also linear in the state variable:

$$\begin{aligned} y_t &= G_0 + Gx_t + u_t \\ &= \tilde{G}_0 + \tilde{G}z_t + u_t, \end{aligned} \tag{29}$$

where $u_t \sim iid\mathcal{N}(0, \sigma_u^2 I)$. The parameters before the rotation of the state space are $G_0 = [b_0 \ c_0 \ \mathcal{B}_0^{(40)}/40]^\top$, $G = [b^\top \ c^\top \ \mathcal{B}^{(40)\top}/40]^\top$. The parameter σ_u^2 is the measurement error variance, and the parameters $\mathcal{B}_0^{(n)}$ and $\mathcal{B}^{(n)}$ are functions of n , A^* , A_0^* and B . The parameters after the rotation are $\tilde{G} = GR^{-1}$, and $\tilde{G}_0 = G_0 - \tilde{G}R_0$.

We can now write the one-step-ahead conditional distribution for the full system:

$$v_{t+1}|v^t \sim \mathcal{N}(C_0 + Cv_t, \Omega\Omega^\top),$$

where $v_t = [z_t \ y_t]^\top$, $C_0 = [\tilde{R}_0^\top \ \tilde{G}_0 + \tilde{G}\tilde{R}_0]^\top$, and

$$C = \begin{bmatrix} \tilde{A} & 0 \\ \tilde{G}\tilde{A} & 0 \end{bmatrix}, \quad \Omega = \begin{bmatrix} \tilde{B} & 0 \\ \tilde{G}\tilde{B} & \sigma_u I \end{bmatrix},$$

which we use to calculate the likelihood function for our sample.

The structural parameters of our model are just-identified from the reduced-form parameters C_0 , C , and Ω . Specifically, this 7-variable system has 46 parameters that are functions of the 46 parameters of our structural model. We apply natural conjugate priors to this Gaussian likelihood function to get a joint posterior for C_0 , C , and $\Omega\Omega^\top$. We repeatedly sample parameter values from this posterior distribution using MCMC methods. For draw, we solve

the 46-equations-in-46-unknowns nonlinear system to obtain the structural parameters. We use the median values of the empirical distributions from this sampling exercise as starting values for just-identified GMM estimation.

We estimate 45 parameters – all but σ_u^2 – to exactly match 45 moment restrictions implied by equations (28) and (29):

$$\begin{aligned}
E(z_t - \tilde{R}_0 - \tilde{A}z_{t-1}) &= 0 \\
E(z_t - \tilde{R}_0 - \tilde{A}z_{t-1})z_{t-1}^\top &= 0 \\
E(y_t - \tilde{G}_0 - \tilde{G}z_t) &= 0 \\
E(y_t - \tilde{G}_0 - \tilde{G}z_t)z_t^\top &= 0 \\
E(z_t - \tilde{R}_0 - \tilde{A}z_{t-1})(z_t - \tilde{R}_0 - \tilde{A}z_{t-1})^\top &= \tilde{B}\tilde{B}^\top
\end{aligned}$$

Since the variance of the sample analogues of these moment restrictions depends only on parameters of the model that we have already estimated, asymptotic standard errors require no further estimation. Derivatives of the moment restrictions are also evaluated at our estimated parameter values. The measurement error variance, σ_u^2 , is estimated using the estimated residuals from the u_t equations.

Average dynamic response. The Taylor rule shock is affected by the entire vector of innovations, w_t . To measure the average dynamic response of observable variables to a Taylor rule shock that is the result of simultaneous innovations in all values of w_t , we calculate a dynamic covariance

$$\text{Cov}_{t-1}(y_{t+j}, s_{2t}) = R_l A^j B B' d_2, \quad j \geq 0,$$

where R_l is the l -th row of R . To calculate a dynamic correlation we scale this conditional covariance with the product of the conditional standard deviations of each variable.

D Long-run real asset prices

Following Alvarez and Jermann (2005), Hansen and Scheinkman (2009), and Hansen (2012), we decompose shocks to the real marginal utility of wealth, *i.e.* the real pricing kernel, into permanent and transitory components. Denote as M_{t+1} the real marginal utility of wealth, and a multiplicative decomposition into permanent and transitory components as $M_{t+1} = M_{t+1}^P M_{t+1}^T$. In growth rates, this implies a comparable decomposition for the real pricing kernel

$$m_{t+1} = \frac{M_{t+1}}{M_t} = \frac{M_{t+1}^P}{M_t^P} \frac{M_{t+1}^T}{M_t^T} = m_{t+1}^P m_{t+1}^T.$$

The positive eigenfunction, v_{t+1} , of the real pricing kernel satisfies the equation

$$E_t m_{t+1} v_{t+1} = e^\rho v_t, \tag{30}$$

where ρ is a positive eigenvalue. M_{t+1}^P will be a martingale when $m_{t+1}^P = e^{-\rho} m_{t+1} v_{t+1} / v_t$, since by construction, $E_t m_{t+1}^P = 1$. In other words, $\log m_{t+1}^P$ is the additive shock to the unit root process for $\log M_{t+1}^P$. Any change in $\log m_{t+1}^P$ creates a permanent change in the real marginal utility of wealth.

To derive $\log m_{t+1}^P$, we start with the log of the real pricing kernel, $\log m_{t+1}$, given by the log of the nominal kernel, $\log m_{t+1}^\$,$ plus the rate of inflation, π_{t+1} :

$$\begin{aligned} \log m_{t+1} &= \log m_{t+1}^\$ + \pi_{t+1} \\ &= -a_0 - a^\top x_t - \lambda_t^\top \lambda_t / 2 - \lambda_t^\top w_{t+1} + b_0 + b^\top x_{t+1} \\ &= (b_0 - a_0) - \lambda_t^\top \lambda_t / 2 + (b^\top A - a^\top) x_t + (b^\top B - \lambda_t^\top) w_{t+1}. \end{aligned} \tag{31}$$

Given the affine structure of the pricing kernel, we guess a log-linear solution for the eigen-

function, $\log v_t = k^\top x_t$, and solve for k using (30) and (31):

$$\rho + k^\top x_t = \log E_t \exp[\log m_{t+1} + k^\top x_{t+1}],$$

which implies

$$k^\top x_t = [(b^\top + k^\top)A - a^\top - (b^\top + k^\top)B\lambda]x_t.$$

Using risk-neutral dynamics, $B\lambda = A - A^*$, the solution for k^\top is

$$k^\top = (b^\top A^* - a^\top)(I - A^*)^{-1}.$$

The innovation to $\log m_{t+1}^P$, therefore, is given by

$$\begin{aligned} \eta_{t+1} &= [b^\top - \lambda_t^\top B^{-1} + (b^\top A^* - a^\top)(I - A^*)^{-1}]Bw_{t+1} \\ &= [(b^\top - a^\top)(I - A^*)^{-1} - \lambda_t^\top B^{-1}]Bw_{t+1}. \end{aligned}$$

The conditional covariance of the innovation to $\log m_{t+1}^P$ with the innovation to the Taylor rule shock, $d_2^\top Bw_{t+1}$ is

$$E_t(\eta_{t+1}w_{t+1}^\top B^\top d_2) = [(b^\top - a^\top)(I - A^*)^{-1} - \lambda_t^\top B^{-1}]BB^\top d_2.$$

When the shock to the Taylor rule has no permanent impact on the real marginal utility of wealth, this correlation must be equal to 0, which implies a set of linear restrictions on d_2 – one for each value of λ_t . Since we need only one of these for our Taylor rule identification exercise, we use the restriction for the average price of risk, $\lambda_t = E\lambda_t = \lambda_0$.

References

- Alvarez, Fernando, and Urban J. Jermann, 2005, "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica*, 73, 1977-2016.
- Anderson, Brian D.O. and John B. Moore, 1979, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice Hall.
- Anderson, Evan W., Ellen R. McGrattan, Lars Peter Hansen, and Thomas J. Sargent, 1996, "Mechanics of forming and estimating dynamic linear economies," in H.M. Amman, D.A. Kendrick, and J. Rust, eds., *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.
- Backus, David K., and Jonathan H. Wright, 2007, "Cracking the conundrum," Finance and Economics Discussion Series 2007-46, Federal Reserve Board.
- Bai, Jushan, and Peng Wang, 2015, "Identification and Bayesian estimation of dynamic factor models," *Journal of Business and Economic Statistics* 33, 221-240.
- Bernanke, Ben S., 2013, "Long-Term Interest Rates," Remarks at the Federal Reserve Bank of San Francisco, March 1.
- Bernanke, Ben, Jean Boivin, and Piotr S. Elias, 2005, "Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach," *Quarterly Journal of Economics* 120, 387-422.
- Blanchard, Olivier Jean, and Danny Quah, 1989, "The dynamic effects of aggregate demand and supply disturbances," *American Economic Review* 79, 655-673.
- Boivin, Jean, and Marc Giannoni, 2006, "DSGE models in a data-rich environment," NBER

Working Paper 12772.

- Boyd, Stephen P., 2009, "Lecture slides for EE363: Linear Dynamical Systems," posted online at <http://www.stanford.edu/class/ee363/>, accessed March 17, 2015.
- 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 inertia 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 J., Martin Eichenbaum, and Charles L. Evans, 1999, "Monetary policy shocks: What have we learned and to what end?" in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics, Volume 1*, 65-148.
- Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, 2005, "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy* 113, 1-45.
- Chun, Albert Lee, 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.
- De Schutter, Bart, 2000, "Minimal state-space realization in linear system theory: an overview," *Journal of Computational and Applied Mathematics* 121, 331-354.

- Gali, Jordi, 2008, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, 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 F., Burton Hollifield, and Stanley E. Zin, 2005, "Taylor rules, McCallum rules and the term structure of interest rates," *Journal of Monetary Economics* 52, 921-950.
- Gallmeyer, Michael F., Burton Hollifield, Francisco J. Palomino, and Stanley E. Zin, 2007, "Arbitrage-free bond pricing with dynamic macroeconomic models," *Federal Reserve Bank of St Louis Review* 89, 305-326.
- Gevers, Michel, and Vincent Wertz, 1984, "Uniquely identifiable state-space and ARMA parametrizations for multivariable linear systems," *Automatica* 20, 333-347.
- Greenspan, Alan, 1994, "Statement by Alan Greenspan, Chairman, Board of Governors of the Federal Reserve System, before the Committee on Banking, Housing, and Urban Affairs, May 27, 1994." *Federal Reserve Bulletin*, July.
- Greenspan, Alan, 2005, "Semiannual Monetary Policy Report to the Congress," February 16.
- Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, "The US Treasury yield curve: 1961 to the present," *Journal of monetary economics* 54, 2291-2304.
- Hamilton, James D., and Jing Cynthia Wu, 2012, "Identification and estimation of Gaussian affine term structure models," *Journal of Econometrics* 168, 315-331.
- Hansen, Lars Peter, 2012, "Dynamic value decomposition within stochastic economies," *Econometrica* 80, 911-967.
- Hansen, Lars Peter, and Thomas J. Sargent, 2013, *Recursive Models of Dynamic Linear*

Economies, Princeton: Princeton University Press.

Hansen, Lars Peter, and Jose Scheinkman, 2009, "Long term risk: an operator approach," *Econometrica* 77, 177-234.

Hinrichsen, D., and D. Pratzel-Wolters, 1989, "A Jordan canonical form for reachable linear systems," *Linear Algebra and Its Applications* 122/123/124, 489-524.

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

Jardet, C., Monfort, A., and Pegoraro, F., 2013, "No-arbitrage near-cointegrated VAR(p) term structure models, term premia and GDP growth," *Journal of Banking & Finance* 37, 389-402.

Joslin, Scott, Anh Le, and Kenneth J. Singleton, 2013, "Gaussian macro-finance term structure models with lags," *Journal of Financial Econometrics* 11, 581-609.

Joslin, Scott, Marcel Pribsch, and Kenneth J. Singleton, 2014, "Risk premiums in dynamic term structure models with unspanned macro risks," *Journal of Finance* 69, 1197-1233.

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

Kim, Don H., and Athanasios Orphanides, 2012, "Term structure estimation with survey data on interest rate forecasts," *Journal of Financial and Quantitative Analysis* 47, 241-272.

King, Robert G., 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.

- 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.
- Rudebusch, Glenn D., 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 chain approach," *Journal of Money, Credit and Banking* 40, 627-666.
- Sims, Christopher A., and Tao Zha, 2006, "Were there regime switches in US monetary policy?" *American Economic Review* 96, 54-81.
- Smets, Frank, and Rafael Wouters, 2007, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American Economic Review* 97, 587-606.
- Smith, Josephine M., and John B. Taylor, 2009, "The term structure of policy rules," *Journal of Monetary Economics* 56, 907-917.
- Stock, James H., and Mark W. Watson, 2012, "Disentangling the channels of the 2007-09 recession," *Brookings Papers on Economic Activity*, Spring, 81-135.
- Taylor, John B., 1993, "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- Uhlig, H., 2005, "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics* 52, 381-419.
- Watson, Mark W., 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: Foundations of a theory of monetary policy*,

Princeton, NJ: Princeton University Press.

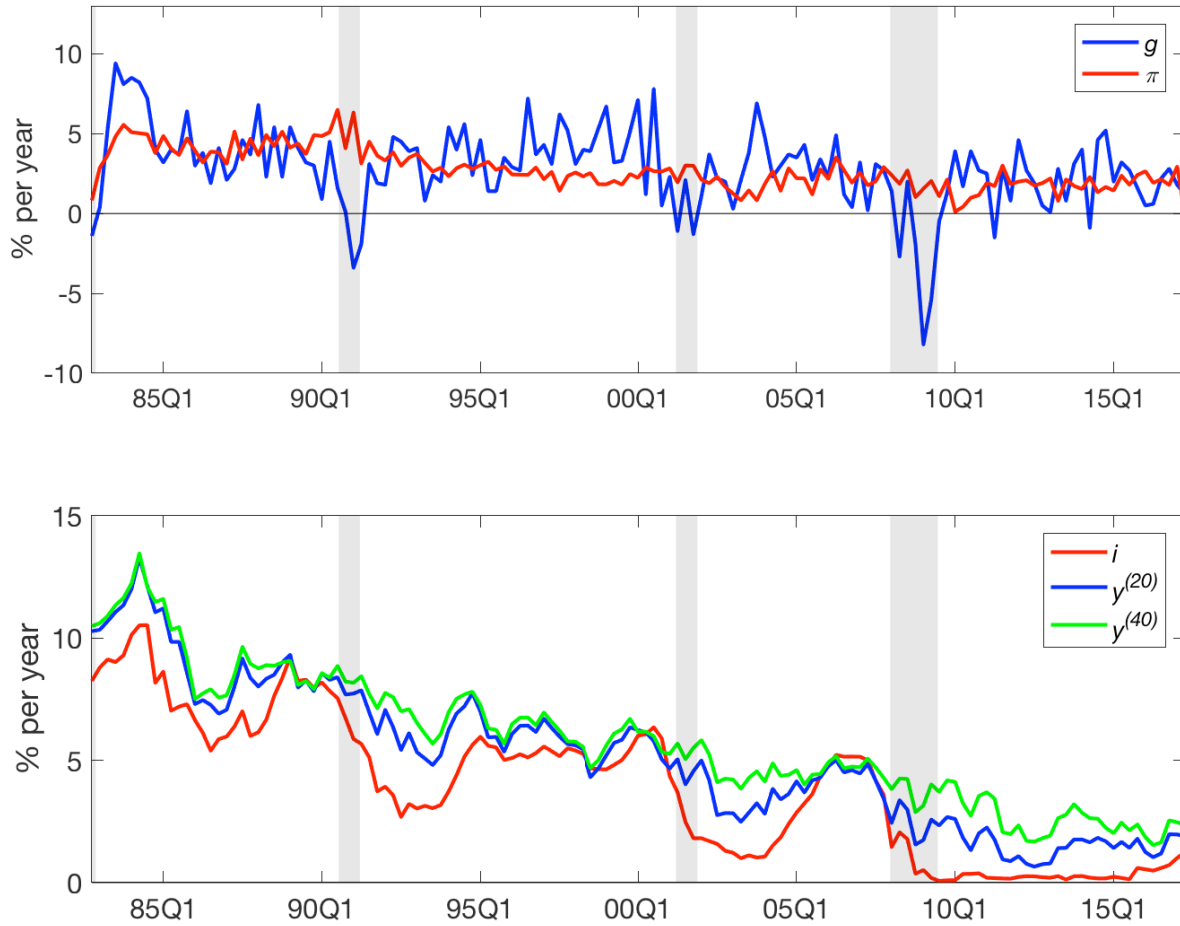
Table 1. GMM Estimation

<i>A</i>				<i>B</i>			
1.0354 (0.0256)	0.1261 (0.0428)	0.0896 (0.0480)	0.1897 (0.0397)	0.0026 (0.0003)	0	0	0
-0.0701 (0.0035)	0.8205 (0.0613)	0.1993 (0.0229)	0.0974 (0.0234)	-0.0050 (0.0026)	0.0057 (0.0014)	0	0
0.0376 (0.0512)	0.0201 (0.0750)	0.5889 (0.0344)	-0.2472 (0.0168)	0.0038 (0.0041)	-0.0063 (0.0022)	0.0030 (0.0007)	0
-0.0819 (0.0504)	-0.0563 (0.0434)	-0.0147 (0.0210)	0.6542 (0.0289)	-0.0012 (0.0012)	0.0010 (0.0006)	-0.0023 (0.0012)	0.0009 (0.0002)
<i>A*</i>				λ_0	<i>b</i>	<i>c</i>	d_2
0.9862 (0.0011)	0	0	0	-0.0442 (0.0097)	0.3238 (0.0896)	0.2933 (0.0081)	0.0490 (0.0692)
0	0.9250 (0.0026)	0	0	-0.0188 (0.0219)	0.2187 (0.0950)	0.2757 (0.0099)	0.3268 (0.0405)
0	0	0.8609 (0.0037)	0	-0.4221 (0.0297)	0.2126 (0.0058)	0.1736 (0.0484)	0.3832 (0.0469)
0	0	0	0.6775 (0.0010)	0.4805 (0.0109)	0.2446 (0.0132)	-0.0030 (0.0916)	0.3709 (0.0703)
τ_0	τ_1	τ_2		a_0	b_0	c_0	
-0.0106 (0.0014)	2.5771 (0.2867)	0.3972 (0.0807)		0.0099 (0.0041)	0.0068 (0.0020)	0.0070 (0.0006)	

Note: Just-identified GMM estimation of the model: $x_{t+1} = Ax_t + Bw_{t+1}$, $i_t = a_0 + a^\top x_t$, $\pi_t = b_0 + b^\top x_t$, $g_t = c_0 + c^\top x_t$, $hy_t^{(h)} = \mathcal{B}_0^{(h)} + \mathcal{B}^{(h)}x_t$, and $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1} (I - A^{*h})$, $\mathcal{B}_0^{(h)} = a_0 + \mathcal{B}_0^{(h-1)} - \mathcal{B}^{(h-1)}B\lambda_0 + \mathcal{B}^{(h-1)}BB^\top \mathcal{B}^{(h-1)\top} / 2$. The Taylor rule is $i_t = \tau_0 + \tau_1\pi_t + \tau_2g_t + d_2^\top x_t$. The state variable x_t is 4-dimensional, i_t is the short interest rate (1 quarter), $y_t^{(h)}$ is the yield on a discount bond of maturity $h = 8, 20, 28, 40$ (quarters), π_t is the inflation rate, g_t is the growth rate of real GDP, and $a^\top = [1 \ 1 \ 1 \ 1]$. The sample period is 1982Q3 to 2017Q2. Asymptotic standard errors are in parentheses.

Figure 1

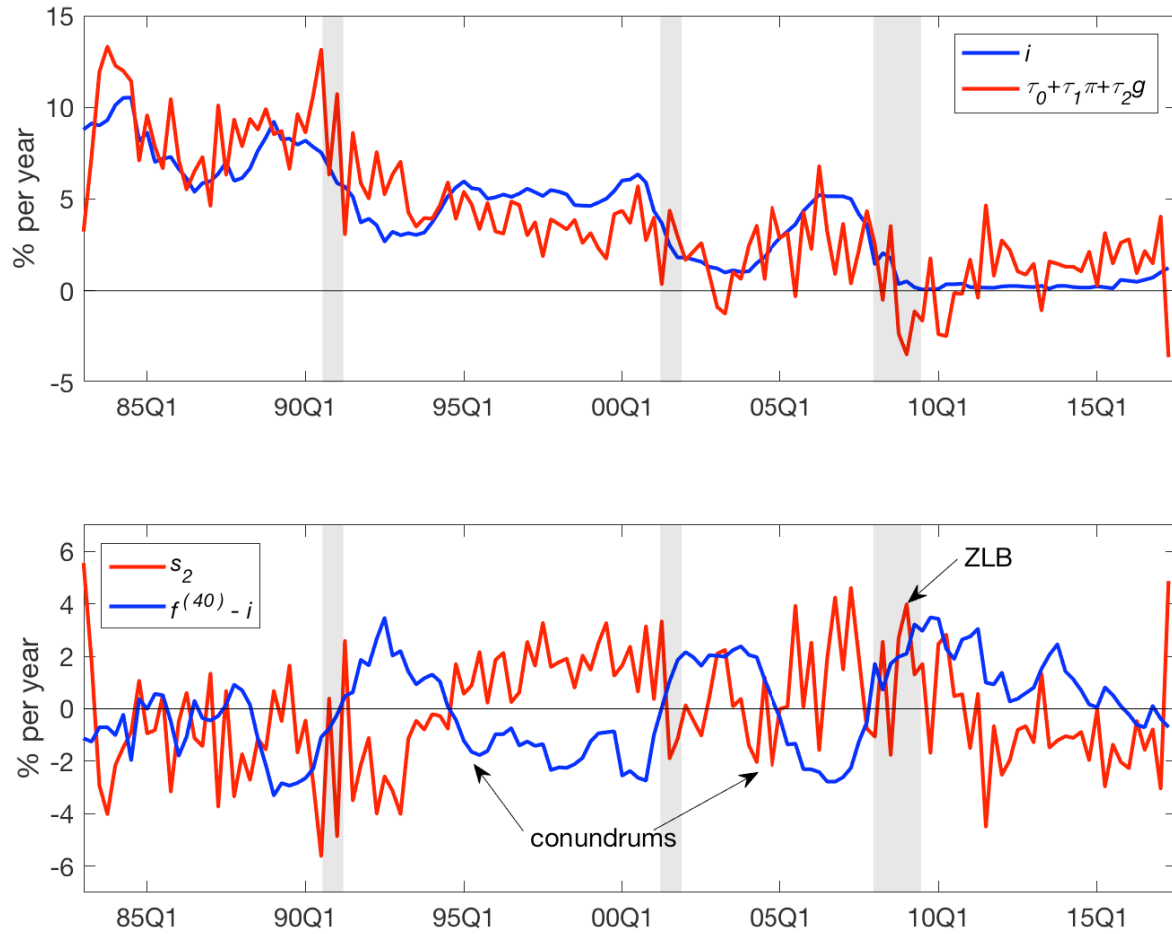
US GDP growth, CPI inflation, and select yields



Note: The time period is 1982Q3 to 2017Q2. Real GDP growth is from the NIPA and CPI inflation is from the BLS, both downloaded from FRED. Yields are from Gürkaynak, Sack, and Wright (2007).

Figure 2

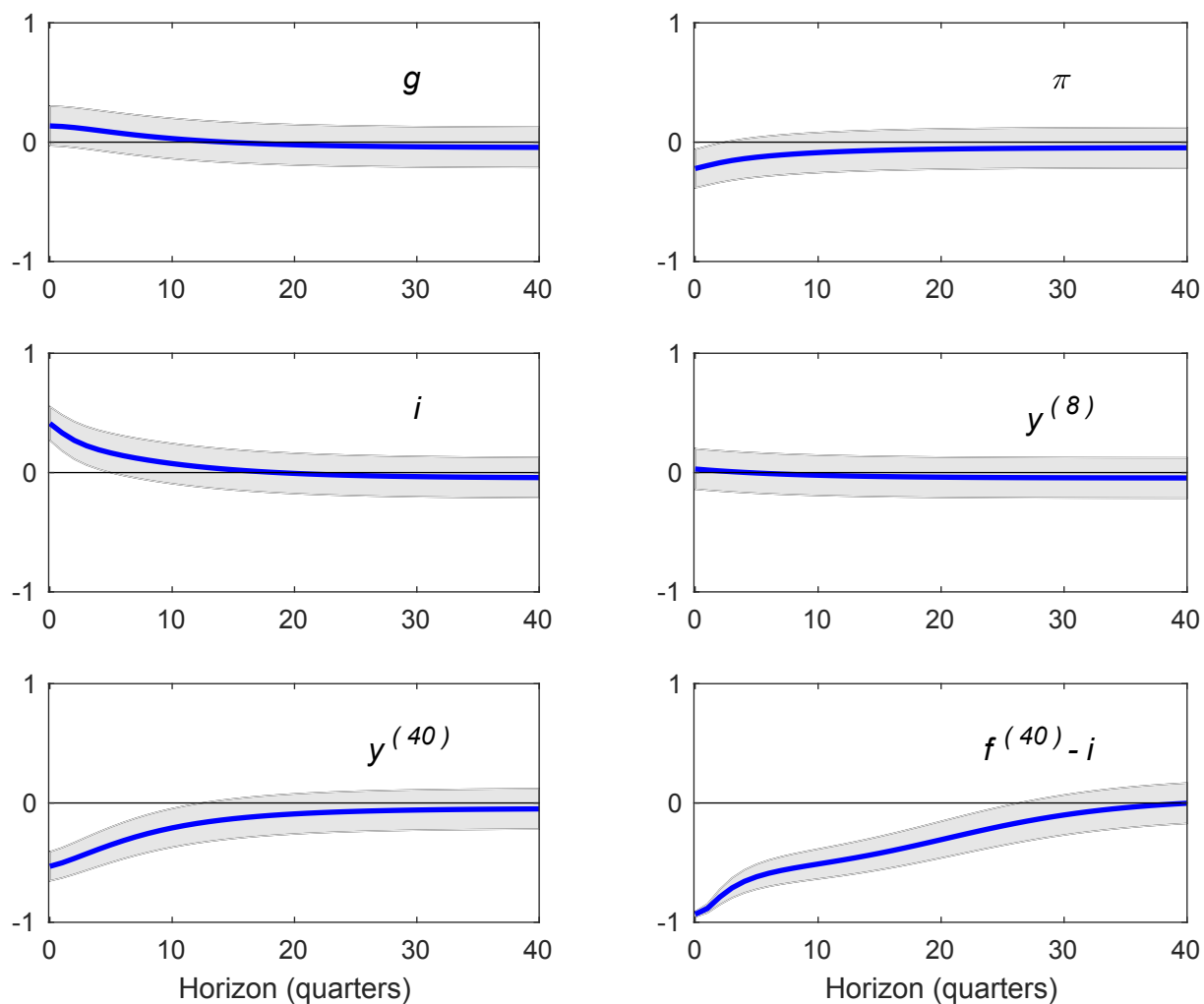
Taylor rule shock



Note: The top panel plots the Taylor rule including the shock, *i.e.*, the short rate, and the Taylor rule excluding the shock, $\tau_0 + \tau_1 \pi_t + \tau_2 g_t$. The difference is the value of the shock, s_{2t} , plotted in the lower panel along with the difference between the ten-year forward rate minus the current interest rate, $f_t^{(40)} - i_t$, *i.e.*, the forward premium.

Figure 3

Dynamic correlations with Taylor rule shock



Note: Each panel plots the conditional correlation of a variable at horizons $t + j$ for $j = 0, 1 \dots 40$ (quarters), with the Taylor rule shock, s_{2t} . The variables are real GDP growth, g_t , inflation, π_t , the short interest rate, i_t , the two-year discount bond yield, $y_t^{(8)}$, the ten-year discount bond yield, $y_t^{(40)}$, and the different between the ten-year forward rate and the short rate, $f_t^{(40)} - i_t$. The shaded area represent a 95% confidence interval.