NBER WORKING PAPER SERIES

MONETARY POLICY RISK: RULES VS. DISCRETION

David Backus
Mikhail Chernov
Stanley E. Zin
Irina Zviadadze

Working Paper 28983
http://www.nber.org/papers/w28983

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2021

Earlier versions of this work circulated with the title “Identifying monetary policy in macro-finance models.” We are grateful to the anonymous referees and editors, Andrew Karolyi and Itay Goldstein, for their thoughtful comments, which have helped us improve this work. We also thank Rich Clarida, John Cochrane, Patrick Feve, Mark Gertler, Valentin Haddad, Kei Kawai, Stacey Polloway, Guillaume Roussellet, Ken Singleton, Gregor Smith, Roman Sustek, Tommy Sveen, and Tao Zha for comments on earlier drafts, and participants in seminars at Aarhus University, ADEMU/CERGE-EI, Banque de France, BI Norwegian Business School, Bilkent University, the Euro Area Business Cycle Network, HEC Paris, INSEAD, New York University, Queen Mary University of London, Stony Brook University, the Swedish House of Finance, and the Toulouse School of Economics. Zviadadze gratefully acknowledges financial support from Investissements d'Avenir Labex (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Monetary Policy Risk: Rules vs. Discretion
David Backus, Mikhail Chernov, Stanley E. Zin, and Irina Zviadadze
NBER Working Paper No. 28983
July 2021
JEL No. E43,E52,G12

ABSTRACT

Long-run asset-pricing restrictions in a macro term-structure model identify discretionary monetary policy separately from a policy rule. We find that policy discretion is an important contributor to aggregate risk. In addition, discretionary easing coincides with good news about the macroeconomy in the form of lower inflation, higher output growth, and lower risk premiums on short-term nominal bonds. However, it also coincides with bad news about long-term financial conditions in the form of higher risk premiums on long-term nominal bonds. Shocks to the rule correlate with changes in the yield curve's level. Shocks to discretion correlate with changes in its slope.

David Backus
Stern School of Business
NYU
44 West 4th Street
New York, NY 10012-1126
N/A user is deceased

Mikhail Chernov
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza, Suite C-417
Los Angeles, CA 90095
and CEPR
and also NBER
mikhail.chernov@anderson.ucla.edu

Stanley E. Zin
Department of Economics
Leonard N. Stern School of Business
New York University
44 West 4th Street, Suite 7-91
New York, NY 10012-1126
and NBER
stan.zin@nyu.edu

Irina Zviadadze
Department of Finance
HEC Paris
1 rue de Liberation
78351, Jouy-en-Josas, France
Zviadadze@hec.fr

1 Introduction

When a central bank follows a monetary policy rule, it commits to a target interest rate that depends on the state of the macroeconomy. However, when policy is also subject to random departures from this rule – which we refer to as monetary policy *discretion* – the central bank also introduces a potential new source of risk.[1] The interest rate we observe, which we refer to as the policy rate, therefore, is the sum of the target rate and discretion. Our goal is to add to the traditional analysis of how monetary policy affects the macroeconomy by exploring the relationship between policy discretion and financial-market conditions.

Most studies to date explore the consequences of shocks to interest-rate policy in one of two ways. The most common method imposes sufficient structure on the behavior of discretionary policy to identify an exogenous policy shock (*e.g.*, Clarida, Galí, and Gertler, 1998, and Christiano, Eichenbaum, and Evans, 1999). Alternatively, using futures data to measure unexpected shocks to the policy rate requires less structure, but these shocks will necessarily combine both target-rate and discretionary shocks (*e.g.*, Kuttner, 2001, and Piazzesi and Swanson, 2008).

We use elements from both of these two approaches to explore different sources for monetary-policy risk. We use financial-market data to reveal the shock to the policy rate. We then impose just enough structure on the behavior of discretionary policy to separately identify its properties from those of the target rate. Since we don't want to wed ourselves to any particular theory of discretion, we impose only long-run neutrality restrictions. These apply across a broad class of both new-Keynesian and neoclassical models, and are likely to conform to most economists' prior beliefs. The features of the data we uncover, therefore, can serve as a benchmark for assessing the relative merits of competing candidate theories.

We specify a macro term-structure model that characterizes the dynamics of the pricing kernel as well as macro variables such as inflation and output growth. This is a convenient setting to explore questions of monetary policy through interest rate targeting. In theory, this framework can characterize the dynamic responses of both real and nominal variables to unobservable shocks to policy discretion, but in practice, this model alone is not enough. In the words of Joslin, Le, and Singleton

[1] Cochrane (2011), Sims and Zha (2006), among many others, refer to the difference between the actual policy rate and the target rate as a policy *disturbance*. Taylor (1999) uses the more value-loaded term of policy *mistake*.

(2013): "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." This echos the concerns of Cochrane (2011), "The crucial Taylor rule parameter is not identified in the new-Keynesian model."

The novelty in our solution to this identification problem is the addition of a long-run real asset-pricing restriction to the more customary long-run real output restriction. We assume that the discretionary monetary policy shock does not have a permanent effect on either the level of real output or the level of real asset values. But since this does not restrict short-run responses, our empirical model is still consistent with most reduced-form econometric models as well as most structural models used in macroeconomics and macro-finance.

We should be clearer about our use of "shocks" in this informal discussion to avoid confusion later. Here we use it to describe unexpected changes from any and all sources. We use long-run restrictions on the impact of unexpected changes to identify structural parameters. We do not use these restrictions to create an orthogonal structure to the exogenous innovations of our dynamic system. As a result, we will not know which shock – or if any shock – in our model can be interpreted as an exogenous change in monetary policy that does not also cause a change in macro and financial variables. Therefore, we cannot conduct the impulse response exercises emblematic of most empirical work in this area, such as structural vector autoregressions. But since we do identify a process for discretionary policy – just not an exogenous innovation to that policy – we are still able to see how it relates to key macroeconomic and financial variables.

In that context, we find that most macroeconomic and financial-market variables are related to a shock to the policy rate in much the same way as they are to a shock to just the target rate. What this masks, however, is that shocks to discretionary policy exhibit substantially different behavior especially with respect to output growth and risk premiums. For example, both discretionary easing and target-rate easing tend to coincide with good news on inflation. However, for output growth, discretionary easing tends to coincide with good news whereas easing through the target-rate coincides with bad news. This evidence is broadly consistent with patterns in the empirical macro literature that find that discretionary policy exhibits a preference for interest-rate smoothness or inertia, *e.g.*, Clarida, Galí, and Gertler (2000) and Rudebusch (2006), or slow adjustment of long-run inflation expectations, *e.g.*, Gürkaynak, Sack, and Swanson (2005a).

With respect to financial markets, we find that discretionary policy is an important contributor to both the mean and variance of risk premiums. We can attribute as much as 20% of the average forward premium on a 10-year discount bond to discretionary policy, which contrasts with its negligible contribution for very short-maturities. We can also attribute about 17% of the variance of forward premiums of all maturities to discretionary policy.

We find that easing through either the target rate or discretion tends to coincide with bad news about long-term financial-market conditions in the form of an unexpected increase in the term premium on long-maturity nominal bonds. However, discretionary and target-rate easing exhibit substantially different patterns in short-term nominal bond markets: discretionary easing tends to coincide with good news in the form of unexpected decreases in the risk premiums of short-term nominal bonds, whereas target-rate easing coincides with increases in these same risk premiums.

Finally, since the patterns for target-rate shocks are similar for both short- and long-term yields, movements in the target-rate are closely related to movements in the *level* of the yield curve. However, since the discretionary shock is more closely related to movements in short-term yields, movements in discretionary policy are closely related to movements in the *slope* of the yield curve.

Section 2 lays out the macro term-structure model that forms the basis for our empirical analysis. Section 3 introduces monetary policy in the form of a standard Taylor rule and frames the identification issue. Section 4 details our identification strategy. Section 5 presents our estimation method and empirical results. Section 6 explores the relationship between policy shocks and macroeconomic and financial-market conditions. Section 7 concludes and suggests future directions.

2 A macro term-structure model

2.1 Abritrage-free pricing

Affine term-structure models have 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), 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 state of the economy is an n-dimensional vector x_t that evolves over time according to the process

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

where A is stable, ε_t is independent across time, and $\varepsilon_t \sim \mathcal{N}(0, \Sigma)$, where Σ is positive definite and symmetric with a unique Cholesky decomposition, B. (All vectors and matrices conform in size to the dimension of x_t.) We will often replace ε_t with its equivalent form, Bw_t, where $w_t \sim iid\ \mathcal{N}(0, I)$ and I is the identity matrix. The unconditional covariance matrix of x_t, V_x, is the solution to $V_x = AV_xA^\top + \Sigma$. 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.

The pricing model starts with the specification of the log pricing kernel that will be used for valuing nominal cashflows,

$$-m^\$_{t+1} = a_0 + a^\top x_t + \lambda_t^\top \lambda_t/2 + \lambda_t^\top w_{t+1}, \tag{2}$$

where $\lambda_t = \lambda_0 + \lambda x_t$. The one-period continuously-compounded nominal interest rate, i_t, is then

$$i_t = -\log E_t \exp(m^\$_{t+1}) = a_0 + a^\top x_t.$$

Given the pricing kernel and the linear transition equation (1), the absence of arbitrage implies that the date-t price, $q_t^{(h)}$, of a default-free pure-discount bond that pays \$1, at date $t + h$, $h > 0$, is log-linear in the state:

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

See Appendix A for the standard derivation of $\mathcal{B}_0^{(h)}$ and $\mathcal{B}^{(h)}$ as functions of the model's parameters.

This implies that continuously compounded yields for $h > 0$ are linear in the state:

$$y_t^{(h)} \ = \ -\log q_t^{(h)}/h \ = \ \frac{1}{h}\left[\mathcal{B}_0^{(h)} + \mathcal{B}^{(h)} x_t\right],$$

and $i_t = y_t^{(1)}$. The forward interest rate for h periods in the future, $f_t^{(h)}$, is also linear in the state:

$$\begin{aligned} f_t^{(h)} &= \log q_t^{(h)} - \log q_t^{(h+1)} \\ &= a_0 + a^\top \mathcal{A}^{*(h)} A_0^* - a^\top \mathcal{A}^{*(h)} BB^\top \mathcal{A}^{*(h)\top} a/2 + a^\top A^{*h} x_t, \end{aligned}$$

as is the risk premium imbedded in this forward rate, *i.e.* the forward premium:

$$\begin{aligned} fp_t^{(h)} &= f_t^{(h)} - E_t i_{t+h} \\ &= a^\top \mathcal{A}^{*(h)} A_0^* - a^\top \mathcal{A}^{*(h)} B B^\top \mathcal{A}^{*(h)\top} a/2 + a^\top (A^{*h} - A^h) x_t. \end{aligned} \tag{3}$$

The *risk-neutral* dynamics of the state, $x_t = A_0^* + A^* x_{t-1} + B w_t$, are governed by parameters $A^* = A - B\lambda$ and $A_0^* = -B\lambda_0$ (see Appendix A), and $\mathcal{A}^{*(h)} = (I - A^*)^{-1}(I - A^{*h})$.

2.2 Term structure identification

At this level of generality, our term-structure model is under-identified. That is, the predictions of the model are invariant to linear transformations of the state variable, x_t. In particular, since multi-period yields depend only on A^*, we cannot distinguish among values for A and λ that leave $A - B\lambda$ unchanged. For example, consider a transformation of the state variable given by a matrix T that defines a new state variable $\hat{x}_t = T x_t$. The physical process for this new state variable is

$$\begin{aligned} \hat{x}_t &= TAx_{t-1} + TBw_t \\ &= (TAT^{-1})\hat{x}_{t-1} + (TB)w_t, \end{aligned}$$

and the nominal-risk-neutral process becomes

$$\begin{aligned} \hat{x}_t &= -(TB)\lambda_0 + T(A - B\lambda)T^{-1}\hat{x}_{t-1} + (TB)w_t \\ &= TA_0^* + (TA^*T^{-1})\hat{x}_{t-1} + (TB)w_t. \end{aligned}$$

Identification of this model, therefore, will be specific to the choice of the matrix T that defines the state variable. We adopt the canonical form used by Hamilton and Wu (2012), Joslin, Le, and Singleton (2013), and Joslin, Priebsch, and Singleton (2014), among many others, and choose a transformation that results in a diagonal matrix governing the dynamics under the risk-neutral distribution. We therefore set the columns of the matrix T^{-1} equal to the eigenvectors of the matrix $A - B\lambda$, which results in a diagonal matrix TA^*T^{-1}. It is important to note that this choice of a rotation affects only the interpretation of the state variable. It does not restrict the behavior of our model of the pricing kernel.

For notational simplicity, we drop the transformation given by the the matrix T. From here on we will refer to the parameters we identify as A, B, and a diagonal matrix A^*. But the reader should be aware that by assuming a diagonal A^*, we have chosen to work with a specific rotation of the abstract state space.

Finally, note that the conditional variance of bond yields, $\mathcal{B}^{(h)}BB^\top\mathcal{B}^{(h)\top} = a^\top(I - A^*)^{-1}(I - A^{*h})BB^\top(I - A^{*h})(I - A^*)^{-1}a$, cannot distinguish a from B. We therefore set a equal to a vector of ones so that the factors, x_t, inherit the same scale as the short rate, i_t. Note that we will continue to use the same notation, but from hereon, a will denote an n-dimensional vector of ones.

2.3 Macro variables and the real pricing kernel

Our empirical model will include some key macroeconomic variables along with the bond yields. Specifically, we will include the rate of inflation, $\pi_t = \log P_t - \log P_{t-1}$, where P_t is the price level, and the rate of growth of aggregate real output, $g_t = \log Y_t - \log Y_{t-1}$, where Y_t is the level of real output. Both are assumed to be linear functions of the vector of state variables:

$$\begin{aligned} \pi_t &= b_0 + b^\top x_t \\ g_t &= c_0 + c^\top x_t, \end{aligned} \tag{4}$$

hence, as is common in the macro-term structure models, *e.g.*, Ang and Piazzesi (2003), Bikbov and Chernov (2010), and Joslin, Le, and Singleton (2013), and many others, output growth and inflation are assumed to be stationary, though potentially highly persistent.[2] Unlike the factor loadings for bond yields, however, the parameters (b_0, b) and (c_0, c) are unrestricted by the model.

Implicit in this specification is a real (log) pricing kernel,

$$\begin{aligned} -m_{t+1} &= -m^{\$}_{t+1} - \pi_{t+1} \\ &= (a_0 - b_0) + (a^\top - b^\top A)x_t + \lambda_t^\top \lambda_t/2 + (\lambda_t^\top - b^\top B)w_{t+1}. \end{aligned} \tag{5}$$

Our empirical exercise does not use any data on real asset prices, nonetheless the real pricing kernel in equation (5) will play an important role in the identification of discretionary monetary policy.

[2] The reduced-form empirical evidence for nonstationary inflation, *e.g.*, Stock and Watson (2007), is weaker for the sample period we use below than for the the entire post-war period, although such evidence can be difficult to interpret in finite samples. Moreover, a nonstationary inflation process in the context of our model would result in a Taylor rule that is incompatible with a stationary nominal short rate, and would also imply a real yield curve that tends to $-\infty$ as maturity increases, which is inconsistent with the evidence from TIPS markets in Gürkaynak, Sack, and Wright (2010).

3 Monetary policy

There is nothing in the specification of our model that would allow us to attach any particular economic interpretation to the state variables, x_t, hence, we are not yet at a point where we can talk about *policy shocks* in a concrete way. To do this we must first introduce a specific policy rule, and explore how shocks to that rule evolve with the state variable, x_t.

3.1 A policy rule

Assume that the monetary authority uses the one-period bond market as its primary policy instrument, and trades whatever quantities of the bond necessary to achieve a target short-term interest rate, denoted i_t^R, according to a standard Taylor rule, as in Taylor (1993),

$$i_t^R = \tau_0 + \tau_\pi \pi_t + \tau_g g_t, \tag{6}$$

where τ_0, τ_π, and τ_g are policy parameters.[3] Implementation of this rule, however, is subject to discretion. The actual one-period interest rate that we observe is a combination of the target rate, i_t^R, and a policy discretion term, s_t,

$$i_t = i_t^R + s_t. \tag{7}$$

In keeping with our stationary log-linear model, we assume that discretion evolves with the state of the economy according to

$$s_t = d^\top x_t, \tag{8}$$

where d is a vector of unknown parameters, but is otherwise unrestricted. Policy discretion, s_t, could be *iid* noise, an independent AR(1) process, or it could depend on all the factors that drive the macroeconomy and financial markets, or just a subset of them. At this point, we are completely agnostic about the properties of s_t encoded in its factor loadings, d.

The necessity for a policy discretion term is self-evident since data on the short rate will never perfectly align with the data on inflation and output growth that set

[3]Taylor's original formulation of the rule used the deviation of output from a trend, *i.e.* potential output. We simplify this by using output growth itself rather than the de-trended level. This choice better aligns our model with the macro term-structure literature while still capturing the intent of Taylor's original rule.

the target rate. The fundamental determinants of this discretion, however, remain outside of our model. Conceptually, there are many possibilities that seem plausible. For example, it could reflect short-term political pressure that is at odds with the long-term economic goals of a central bank that lacks complete independence. Or it could reflect the market microstructure of interactions between the central bank and its network of private brokers when implementing a particular policy rule. Or real-time measurement error in inflation and output. Or a central bank's desire to keep the short rate positive. Or its response to a financial crisis. These all suggest that the central bank's interest rate policy will depend on more than just π_t and g_t through the target rate, i_t^R. "Even strong proponents of simple policy rules generally advise that they be used only as guidelines, not as substitutes for more complete policy analyses," in the words of Chairman Bernanke (2005).

The presence of the policy discretion term, s_t, in the specification of monetary policy is intended to capture these possibilities. However, each of these examples, either individually or in some combination, is likely to result in a process for the policy discretion term that depends on the state of the economy in a different way. In the empirical exercise below, we don't attempt to model s_t explicitly, rather we simply allow the data to determine the process for policy discretion. But as we show in the next section, since s_t is not directly observable, we must first place additional structure on the model to achieve identification.

3.2 Identification of monetary policy

Incorporating monetary policy into our term-structure model introduces an additional identification problem beyond the term-structure identification discussed above.[4] To keep this discussion separate, assume that we already know the values of all of the parameters of the macro term-structure model. That is, we know the factor loadings for inflation and output growth, *i.e.*, the parameter vectors b and c, as well as the intercept parameters a_0, b_0, and c_0. (Recall that the factor loadings for the interest rate have been scaled to equal a vector of ones.) The only parameters left to identify and estimate are those of the target rate, τ_0, τ_g, and τ_π, and policy discretion, d.

[4]See Backus, Chernov, and Zin (2015) for a more thorough discussion.

3.2.1 The problem

Since monetary policy discretion is only observed indirectly as the difference between the policy rate and the target rate, a necessary first step must be to identify the parameters that determine the target rate separately from the parameters that determine policy discretion. For example, if we see the policy rate increase when we see inflation and output increase, we could be tempted to conclude that this is a natural consequence of the policy rule. Likewise, if we see the policy rate increase when inflation and output are constant, we could be tempted to conclude that this must be the result of policy discretion. But since equilibrium inflation and output respond to both changes in the target rate and changes to policy discretion *simultaneously*, we can draw no such conclusions. Additional structure is needed to separately identify each of these effects.

To see this algebraically, we write each term in the monetary-policy equation (7) as a function of the state variable:

$$\underbrace{a_0 + a^\top x_t}_{=i_t} = \tau_0 + \tau_\pi(\underbrace{b_0 + b^\top x_t}_{=\pi_t}) + \tau_g(\underbrace{c_0 + c^\top x_t}_{=g_t}) + \underbrace{d^\top x_t}_{=s_t},$$

which implies $n + 1$ linear parameter restrictions

$$a = \tau_\pi b + \tau_g c + d$$
$$a_0 = \tau_0 + \tau_\pi b_0 + \tau_g c_0.$$

In other words, monetary policy added $n + 1$ new restrictions to our system of equations. But it also added $n + 3$ new parameters, τ_0, τ_π, τ_g, and d. Note also that there are no additional identifying restrictions provided by the equations for multi-period bonds yields. Identification of the equations for bond yields did not depend on monetary policy, and likewise, monetary policy does not depend on those bond yields. Even if we had exact prior knowledge of all the parameters of the pricing kernel, they would be of no help identifying the parameters of the monetary policy rule.

This lack of identification does not come as a surprise since our model is just a more general case of the well-known example in Cochrane (2011). Simplify our model by setting $a^\top = b^\top A$, so that the interest-rate equation is $i_t = r + E_t\pi_{t+1}$ (*i.e.*, a simple Fisher equation with a constant real rate, $r = a_0 - b_0$). Simplify the policy rule to depend only on inflation, *i.e.*, assume $\tau_g = 0$, then the parameter restrictions implied by the Taylor rule are

$$\begin{aligned} b^\top &= d^\top (A - \tau_\pi I)^{-1} \\ b_0 &= (a_0 - \tau_0)/\tau_\pi. \end{aligned} \tag{9}$$

Even if b_0, a_0, b, and A were known, we are still left with an under-identified system of equations: in this case $n+1$ equations in $n+2$ unknowns, τ_0, τ_π and d.

This example is simple, but telling. Since s_t is unobservable, we cannot estimate d directly. And measuring s_t as a residual in the policy rate equation requires prior knowledge of the other policy parameters, τ_0, τ_π, and τ_g, which we do not have. We need at least two additional restrictions on d to identify the Taylor rule in (6). Note that 2 restrictions will be enough to identify the 2 policy parameters τ_π and τ_g, so that τ_0 can be identified from the equation $\tau_0 = a_0 - \tau_\pi b_0 - \tau_t c_0$, which will then identify the policy discretion term, s_t, as a residual so that we can estimate the remaining n independent parameters of d. Additional restrictions on d may add information, but are not strictly necessary for identification.

It is worth highlighting that this lack of identification has nothing to do with the reduced-form nature of our arbitrage-free macro term-structure model and nothing to do with our definition of the policy rule. Placing additional economic structure on the pricing kernel, the macroeconomic variables, or the policy rule itself does not alter the monetary policy identification problem if that structure does not somehow restrict the policy parameters, τ_0, τ_π, τ_g, or d. (See Appendix B for examples.)

3.2.2 The role of instrumental variables

Finally, identification via an instrumental variables estimator presupposes the existence of at least two valid instruments, z_{it}, such that $E(z_{it}s_t) = 0$, for $i = 1, 2$. In the context of our state-space model, z_{it} is a function of the state variable, say $z_{it} = \beta_i^\top x_t$, for a vector of parameters β_i. Instrumental variables estimation, therefore, requires $E(z_{it}s_t) = E(\beta_i^\top x_t x_t^\top d) = \beta_i^\top V_x d = 0$. In other words, this identification is predicated on knowledge of at least two additional restrictions on the parameters of the policy rule, restrictions that are not part of the specification of the model. Without an economic theory of s_t that places additional structure on its behavior, the motivation and interpretation of any instrumental variables estimator is open to question. Unfortunately, macroeconomic theory is generally quite vague about the fundamental determinants of policy discretion even though a policy-discretion variable is a standard feature of most structural models of monetary policy.

The general feature of the identification problem highlighted in these simple examples extends to more complicated models: unless a model places explicit restrictions on the parameters of the policy rule it will be fundamentally under-identified. In the next section we introduce and justify a set of such restrictions and integrate them into our macro term-structure model.

4 The long-run neutrality of discretionary shocks

If we had exact prior knowledge of the values of τ_π and τ_g, then identification of the discretionary shock would require nothing more beyond the identification of the parameters of the macro term-structure model. To take a concrete example, if we were certain that $\tau_\pi = 1.5$ and $\tau_g = 0.5$, then given values for a, b, and c, the discretionary shock is simply $s_t = (a - 1.5b - 0.5c)^\top x_t$. Likewise, if we had exact prior knowledge of the values of d, then identification of the target-rate parameters would require nothing more beyond the identification of the parameters of the macro term-structure model. For example, if we were certain that the third factor in a three-factor macro term-structure model was an exogenous AR(1) process for discretionary policy, *i.e.*, $s_t = d_3 x_{3t}$, then we would also be certain that $d_1 = 0$ and $d_2 = 0$. The target-rate parameters, τ_π and τ_g, would then simply solve the linear equations $\tau_\pi b_1 + \tau_g c_1 = a_1$ and $\tau_\pi b_2 + \tau_g c_2 = a_2$. Unfortunately, we have neither a theoretical nor an empirical justification for assuming such exact prior knowledge either about the target-rate parameters or the discretionary policy parameters. And since identifying assumptions cannot be tested, relying on our personal intuition or sheer guesswork is unlikely to lead to convincing empirical conclusions.

Rather than assuming exact prior knowledge about specific parameter values governing monetary policy, perhaps we would have more confidence assuming exact prior knowledge of some properties of the joint distribution of variables in our model. For example, if we have exact prior knowledge that current and lagged values of the discretionary shock are uncorrelated with real output growth, *i.e.*, $E[g_t s_t] = 0$ and $E[g_t s_{t-1}] = 0$, then parameters of the macro term-structure model provide two linear equations, $c^\top V_x(a - \tau_\pi b - \tau_g c) = 0$ and $c^\top A V_x(a - \tau_\pi b - \tau_g c) = 0$, that we can solve for τ_π and τ_g. Unfortunately, such assumptions are also unlikely to lead to convincing empirical conclusions since they are inconsistent with both new-Keynesian sticky-price models and neoclassical models with frictions. In essence, we would be assuming away one of the most interesting and long-standing questions in macroeconomics – Does monetary policy have real effects? – ironically, for the sake of identifying monetary policy. Did we simply make an unfortunate choice using g_t? Why not something like a long-term bond yield, $y_t^{(n)}$, instead? Couldn't we just substitute $\mathcal{B}^{(n)}$ for c in those two linear restrictions to identify the target-rate parameters? In principle, yes of course. But if we want to allow for the possibility that real output growth affects the pricing kernel and, hence, bond valuations, while maintaining a channel for discretionary policy to affect real output growth, then we are right back in the same situation. We would have achieved identification by ruling out the most commonly used structural asset-pricing models, and as a result, our empirical conclusions would rightly be viewed with skepticism.

Where does that leave us? We could go on and on describing potential identifying restrictions, then questioning their usefulness as the basis for our empirical exercise. Once again, without specifying a complete structural model, such an exercise would amount to little more than speculation and subjective personal intuition. But there is no free lunch! To proceed we must necessarily restrict some features of our model and our empirical conclusions must necessarily depend entirely on those restrictions.

Ideally we want an approach that is flexible enough to accommodate as large a set of structural models as possible yet still restrict the parameters of monetary policy. To that end, we will focus exclusively on the long-run consequences of temporary shocks to discretionary policy. One feature shared across a very broad class of structural models, and which is likely to conform to most economists' prior beliefs, is that a temporary shock to discretionary monetary policy may affect the real economy, but that effect will *not be permanent*. We will now show how that seemingly weak and robust requirement can be used for identification, then how that identification translates to the empirical properties of our macro term-structure model.

4.1 A long-run quantity restriction

The first restriction we impose is a standard assumption in both new-Keynesian and neoclassical models: shocks to monetary policy discretion may have short-run consequences for the level of real output but they do not have a permanent impact. Note that in our macro term-structure model, we have implicitly assumed that the level of real output is nonstationary. Recall that g_{t+1} is the continuously compounded growth rate of real output, which we denote as Y_t. Since we have assumed that g_{t+1} is stationary, the process for $\log Y_t$ has a unit-root, $\log Y_{t+1} = \log Y_t + g_{t+1}$, which implies

$$\log Y_{t+n} = \log Y_t + g_{t+1} + g_{t+2} + \cdots + g_{t+n}.$$

In other words, even when n is arbitrarily large, the current shocks to the state of the economy, w_{t+1}, continue to have effects on $\log Y_{t+n}$ directly through g_{t+1} and indirectly through the conditional means of g_{t+j}, $j = 2, \ldots, n$.

What concerns us, however, isn't an arbitrary shock to the state of the economy, but rather the shock specifically to discretionary policy, $s_{t+1} = d^\top x_{t+1}$. What we would like to rule out is *that* specific linear function of the state affecting the conditional forecast of $\log Y_{t+n}$ when n is large. Note that the conditional covariance of g_{t+j} and s_{t+1} for $j \geq 1$ is

$$\mathrm{Cov}_t(g_{t+j}, s_{t+1}) = E_t[c^\top x_{t+j} x_{t+1}^\top d] = c^\top A^{j-1} B B^\top d,$$

which implies

$$\lim_{n \to \infty} \mathrm{Cov}_t(\log Y_{t+n}, s_{t+1}) = \sum_{j=0}^{\infty} c^\top A^j B B^\top d = c^\top (I - A)^{-1} B B^\top d.$$

The assumption that this long-run covariance is zero, therefore, implies a linear restriction on d. Recall that d is a function of the factor loadings from the macro term-structure model and the parameters of the policy rule, *i.e.*, $d = a - \tau_\pi b - \tau_g c$, so the restriction is

$$c^\top (I - A)^{-1} B B^\top (a - \tau_\pi b - \tau_g c) = 0, \tag{10}$$

that we can use to help identify the parameters of the monetary policy rule.

We could have derived this same restriction in a slightly different way. Following Beveridge and Nelson (1981) we can decompose the log of real output into a permanent component $\log Y_{t+1}^P$ that is a random walk, and a temporary component $\log Y_{t+1}^T$ that is stationary. Begin with the moving average representation for $\log Y_t$,

$$(1 - L) \log Y_{t+1} = c_0 + c^\top (I - AL)^{-1} B w_{t+1},$$

where L is the lag operator. The permanent component in the Beveridge-Nelson decomposition is given by the random walk with an innovation scaled by the moving average polynomial evaluated at $L = 1$,

$$(1 - L) \log Y_{t+1}^P = c^\top (I - A)^{-1} B w_{t+1},$$

and the transitory component is then $\log Y_{t+1}^T = \log Y_{t+1} - \log Y_{t+1}^P$. This is equivalent to decomposing the growth rate of real output, g_{t+1}, into additive permanent and transitory components, $g_{t+1} = g^P_{t+1} + g^T_{t+1}$, where $g^P_{t+1} = c^\top (I - A)^{-1} B w_{t+1}$. Therefore, if a discretionary monetary policy shock has no permanent effect on the level of real output, it's covariance with real output growth satisfies the restriction

$$E_t[g_{t+1}s_{t+1}] = E_t[(g^P_{t+1} + g^T_{t+1})s_{t+1}] = E_t[g^T_{t+1}s_{t+1}], \tag{11}$$

or simply $E_t[g^P_{t+1}s_{t+1}] = 0$. This can be written as

$$c^\top (I - A)^{-1} B B^\top (a - \tau_\pi b - \tau_g c) = 0,$$

which is the identical restriction to (10).

This alternative derivation doesn't add much value to the more intuitive direct method we originally used to derive equation (10). In the next section, however, we will need to apply similar arguments to the nonlinear process for the real pricing kernel, m_{t+1}, in which case the analog to the Beveridge-Nelson decomposition given in Alvarez and Jermann (2005), Hansen and Scheinkman (2009), and Hansen (2012), will prove much more useful. Nonetheless, providing the permanent-transitory decomposition for g_{t+1} may help clarify the parallel interpretation of our two identifying restrictions. It is also helpful to see that a model with the restriction $E_t[g_{t+1}s_{t+1}] = E_t[g^T_{t+1}s_{t+1}]$ still has ample (if not *total*) flexibility to match any short-run relationship between discretionary policy and real output evident in the data.

The restriction in equation (10) is analogous to the long-run restriction on monetary-policy shocks used in structural VAR models popularized by Blanchard and Quah (1989). In fact, if we could observe s_t, then it could be included in a VAR with g_t, and (10) would be the natural identifying restriction in the Blanchard-Quah methodology. In the context of our model however, s_t is unobservable. Therefore, we use this restriction to identify structural parameters of the policy rule rather than to orthogonalize the shocks in a structural VAR. The result of this identification will be a process for discretionary monetary policy, s_t, with the desired long-run neutrality property.

It is also important to note that imposing comparable restrictions that shocks to other variables in our model have no impact on the level of long-run real output, do not restrict the parameters of discretionary policy, d, or the target-rate, τ_0, τ_π, or τ_g. For example, assuming that the shock to inflation has no permanent effect on the level of real output implies the constraint $c^\top (I - A)^{-1} B B^\top b = 0$, or similarly the shock to nominal interest rates, $c^\top (I - A)^{-1} B B^\top a = 0$, or the shock to any long-term bond yield, $c^\top(I-A)^{-1}BB^\top\mathcal{B}^{(n)}=0$. None of these constraints involve the parameters of discretionary policy, d, or the target-rate, τ_0, τ_π, or τ_g. Adding such restrictions will over-identify the parameters of the macro term-structure model and may be of interest for other reasons, but they will not help identify the unknown policy parameters.

Alternatively, we could consider adding other observable real quantities to our macro term-structure model, such as consumption and investment, and impose similar restrictions on the long-run responses of their levels to discretionary policy shocks as additional identifying restrictions. These would indeed involve d and, in principle, could help with identification. However, given the evidence that the levels of other real variables are cointegrated with the level of real output, *e.g.*, Engle and Granger (1991), such restrictions are unlikely to add much new information beyond (10), and would at best provide a very weak identification. Instead, we make use of our pricing-kernel model and explore real asset prices as the source for additional – and as yet unexplored – identifying restrictions.

4.2 A long-run asset-pricing restriction

The second restriction we impose is a natural analog to (10) applied to real asset prices: shocks to discretionary policy may have short-run consequences for the level of real asset values but they do not have a permanent impact. Since we have already assumed that the long-run level of real quantities are unaffected by shocks to policy discretion, what this assumption adds is a restriction to the process for real marginal valuations, denoted $\mathcal{M}_t$. As with real output, this is a unit root process in logs with increments determined by the real pricing kernel, $\log\mathcal{M}_{t+1}=\log\mathcal{M}_t+m_{t+1}$, which implies

$$\log\mathcal{M}_{t+n}=\log\mathcal{M}_t+m_{t+1}+m_{t+2}+\cdots+m_{t+n}.$$

In other words, even when n is arbitrarily large, current shocks to the state of the economy, w_{t+1}, continue to have effects on real marginal valuations, $\log\mathcal{M}_{t+n}$, directly through m_{t+1} and indirectly through the conditional means of m_{t+j}, $j=2,\ldots,n$. Unlike our model of real quantities, however, the real pricing kernel has an additional channel for w_{t+1} to affect long-run real valuations: directly through the price of risk in m_{t+2}, and indirectly through the conditional means of the prices of risk in m_{t+j}, $j=3,\ldots,n$.

Again, what concerns us isn't how an arbitrary shock to the state of the economy affects long-run real valuations, but rather the shock specifically to discretionary policy, $s_{t+1} = d^\top x_{t+1}$. What we would like to rule out is *that* specific linear function of the state affecting the conditional forecast of $\log \mathcal{M}_{t+n}$ when n is large.

In parallel to our earlier discussion of the decomposition of real output growth, we follow Alvarez and Jermann (2005), Hansen and Scheinkman (2009), and Hansen (2012), and decompose the real pricing kernel into multiplicative permanent and transitory components, which will then be additive in logs, $m_{t+1} = m^P_{t+1} + m^T_{t+1}$. If a discretionary monetary policy shock has no permanent impact on the real asset values, its covariance with the real pricing kernel must satisfy the restriction analogous to equation (11),

$$E_t[m_{t+1}s_{t+1}] = E_t[(m^P_{t+1} + m^T_{t+1})s_{t+1}] = E_t[m^T_{t+1}s_{t+1}], \tag{12}$$

or simply $E_t[m^P_{t+1}s_{t+1}] = 0$.

It is worth emphasizing that this restriction is *not* a requirement that the prices or returns of long-term assets such as long-maturity bonds, be unaffected by the shocks to monetary policy either through the target rate or discretion. On the contrary, the model still has ample (if not *total*) freedom to match whatever features of these relationships are evident in the data. The restriction simply connects the price of the real risk embedded in the policy discretion term to its covariance with the return on a real consol bond. Let $r^{(\infty)}_{t+1}$ denote the one-period log return on a zero-coupon real consol, then the real pricing kernel decomposition above implies that $-m^T_{t+1} = r^{(\infty)}_{t+1}$. Therefore, this restriction can be written as

$$E_t[m_{t+1}s_{t+1}] = -E_t[r^{(\infty)}_{t+1}s_{t+1}].$$

That is, the real risk price of a shock to policy discretion can be measured with its covariance with the return on a real consol bond.

Following Hansen and Scheinkman (2009), we extract the permanent component using the dominant eigenvalue, ρ, of $e^{m_{t+1}}$, and its corresponding eigenfunction, ϕ_t. The eigenfunction and eigenvalue satisfy the equation

$$E_t\left[e^{m_{t+1}}\phi_{t+1}\right] = e^{\rho}\phi_t,$$

which implies that the log of the permanent component is defined by

$$m^P_{t+1} = m_{t+1} - \rho + \log\phi_{t+1} - \log\phi_t.$$

Given the affine structure of the real pricing kernel in (5), the eigenfunction will be log-linear, $\log \phi_t = k^\top x_t$, where $k^\top = (b^\top A^* - a^\top)(I - A^*)^{-1}$, and the permanent component of the real pricing kernel is given by

$$m^P_{t+1} = -(b+k)^\top BB^\top (b+k)/2 - (b+k)^\top A^*_0 - \lambda_t^\top \lambda_t/2 \\ + [b^\top A - a^\top - k^\top (I - A)]x_t + (b^\top B - \lambda_t^\top + k^\top B)w_{t+1}.$$

The covariance restriction in (12) is then

$$E_t[m^P_{t+1} s_{t+1}] = [(b^\top - a^\top)(I - A^*)^{-1} - \lambda_t^\top B^{-1}]BB^\top d = 0. \tag{13}$$

The first term in this expression is analogous to the permanent component of the Beveridge-Nelson decomposition we saw previously. But here it is applied to the process for the conditional mean of the *risk-neutral* distribution of the real pricing kernel. The second term recognizes the fact that shocks to the pricing kernel may also have a separate permanent effect through the price of risk. Therefore, we must restrict the combined correlation through both channels.[5]

This provides us with a set of linear restrictions on d, one for each value of x_t. Since we need only one more restriction for our monetary policy identification exercise, and since we are primarily concerned with restricting long-run behavior, we assume that this conditional moment restriction holds when x_t is equal to the mean of its long-run distribution, $Ex_t = 0$, which implies

$$[(b^\top - a^\top)(I - A^*)^{-1} - \lambda_0^\top B^{-1}]BB^\top (a - \tau_\pi b - \tau_g c) = 0. \tag{14}$$

Note that given the dependence of this restriction on the price-of-risk parameters, λ_0 and λ (recall $A^* = A - B\lambda$), a well-specified pricing-kernel model is essential to the construction of this second identifying restriction.[6]

[5] Following Hansen (2012), we can use m^P_{t+1} in our affine setup to form a "twisted" process for the state variable analogous to the risk-neutral distribution, $x_{t+1} = A^P_0 + A^P x_t + Bw_{t+1}$, with the property that the permanent component of the valuation of a random payoff $\exp\{g^\top x_{t+1}\}$, given by $E_t \exp\{m^P_{t+1} + g^\top x_{t+1}\}$ is equal to the expected value of that payoff under the twisted process, $E^P_t \exp\{g^\top x_{t+1}\}$, where $A^P = A^*$ and $A^P_0 = (b-a)^\top (I - A^*)^{-1} BB^\top + A^*_0$. In that context, the restriction in (13) is equivalent to the restriction $E^P_t s_{t+1} = E_t s_{t+1}$. In other words, the discretionary shock does not contribute any long-run real risk.

[6] The restriction in (14) still has content when the real asset pricing kernel corresponds to a *literally* risk-neutral model, $-m_{t+1} = r_t$. In that case, $A^* = A$ and $\lambda_0 = b^\top B$, so that $\lambda_t - b^\top B = 0$. The real marginal utility of wealth becomes a simple linear process with innovations given by the real interest rate, so that its permanent component is given by the Beveridge-Nelson decomposition. The identifying restriction in (14) reduces to $(a^\top - b^\top A)(I - A)^{-1} BB^\top d = 0$, which still says that the discretionary policy shock is uncorrelated with the permanent component in the real marginal utility of wealth.

Intuitively, the value-added of this asset-pricing-based restriction relative to the quantity-based restriction in equation (10), is in the way it restricts the long-run response of real risk prices to a discretionary monetary policy shock. To see this, consider the log real pricing kernel in a typical new-Keynesian macro model like the one discussed in Section 3.2, $m_{t+1} = \delta_0 + \delta g_{t+1}$, which is based on a representative agent with CRRA expected utility (we're now interpreting g_{t+1} as real consumption growth as in an endowment economy). The real risk prices in this model are constant so that risk-neutral persistence is the same as physical persistence, *i.e.*, $A^* = A$. In addition, from equation (17) we know that $\lambda_0^\top B^{-1} = \delta c^\top$ and $(b - a)^\top = \delta c^\top A$. In this case, equation (14) reduces to $c^\top (I - A)^{-1} BB^\top d = 0$. This is identical to the restriction in equation (10). A simpler way to see this, of course, is just to note that the moving-average coefficients in this linear time-series model of the log real pricing kernel are equal to δ times the moving-average coefficients of g_{t+1}, hence, it's permanent component is proportional to that of g_{t+1}. And obviously, any model that shares this feature with CRRA expected utility will likewise produce a redundant restriction. Since this is not the case for our model with state-dependent risk prices, equation (14) will add an additional restriction that we can use to identify the parameters of monetary policy.

5 Empirical Results

Our empirical exercise proceeds sequentially. In the first step we estimate the parameters of the macro term-structure model detailed in Section 2. Note that the restrictions in Section 4 that will ultimately identify monetary policy play no role in this step. By design, the identification assumptions of the factor model outlined in Section 2.2 do not restrict either the temporary or permanent components of the pricing kernel. And the same is true for the factor loadings for macro variables in equation (4) which are completely unrestricted. In other words, any equilibrium implications of monetary policy or the restrictions in Section 4 will be captured in the estimates of the macro term-structure model through the observed behavior of interest rates, inflation, output growth, and multi-period bond yields.

The second step is to substitute the parameter values from this macro term-structure estimation into the two restrictions that identify the monetary policy parameters given in equations (10) and (14), and solve for estimates of τ_g and τ_π. And those two values will then imply a value for $d = a - \tau_\pi b - \tau_g c$. In that sense, the parameters of our monetary policy model will be *just-identified* even when the macro term-structure model is *over-identified*.

5.1 Macro term-structure estimation

We estimate the affine term-structure model as outlined in Section 2 using quarterly US data from 1980Q3 to 2019Q4 and Generalized Method of Moments based on the conditional-moment restrictions of our model. (See the Appendix D for details.) The sample period begins one year after Volcker's ascendance as Fed chair to allow his monetary regime to establish credibility.[7] For the quarterly interest rate we use the Fama-Bliss data (available from CRSP), and for longer-maturities we use continuously compounded default-free pure-discount bond yields as measured by Gürkaynak, Sack, and Wright (2007). All yields are measured on the last day of the quarter. Real GDP growth rates are from the National Income and Product Accounts, and core CPI inflation from the Bureau of Labor Statistics. Inflation is measured as the quarter-to-quarter change in the average monthly CPI. Growth rates are continuously compounded at annual rates. Figures 1 displays these standard data for our sample period.

The results of GMM estimation are summarized in Table 1. The estimated parameter values have a number of noteworthy features. The risk-neutral dynamics encoded in the non-zero elements of A^*, are substantially more persistent than the actual dynamics of the state space in A. The absolute values of the 4 eigenvalues of A, which govern the persistence in the process for x_t, are 0.9775, 0.6705, 0.6705, and 0.3180 (the middle pair correspond to complex conjugates). The diagonals of A^* are 0.9961, 0.8814, 0.8348, and 0.3638, and they are very precisely estimated – a consequence of the additional information in the cross-equation restrictions implied by the absence of arbitrage – and are 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 the model's dynamics below.

[7] Based on our understanding of the events of that time as summarized in Goofriend and King (2005), it seemed prudent to avoid the temporary disruption caused by the imposition of credit controls at the beginning of 1980, which the Fed was able to effectively work around by the summer of that year. Since our focus will be on a stable Taylor rule we avoided using pre-Volcker data. The end of our sample was dictated by the available data when we undertook the estimation.

The inflation rate has significant loadings (the values of b) for all four factors. On the other hand, real GDP growth has significant 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 with the fourth latent factor.

The average price of risk, λ_0, is negative for all four latent factors and appears to be different from zero for three of the four factors. For the third factor – the third most persistent under the risk-neutral distribution – it is both small in absolute value and statistically insignificant. The average price of risk for the first factor – the most persistent factor under the risk-neutral distribution – is negative and close to zero, but it is statistically significant. The least persistent factor has an average prices of risk that is relatively larger and is statistically significant.

Finally, since we will be basing our monetary-policy identification on the decomposition of the real pricing kernel into permanent and transitory components, it would be reassuring if those components were consistent with asset-pricing restrictions beyond the term-structure moments used in estimation. Alvarez and Jermann (2005) propose using an entropy bound for this purpose. (See Appendix C for closed-form expression for various measures of entropy for our model.) At our point estimates, the ratio of unconditional entropy of the permanent component of the real pricing kernel to the unconditional entropy of the real pricing kernel itself is 1.0267, which is consistent with the estimates of lower bounds provided in Alvarez and Jermann (2005).

5.2 Monetary policy estimation

Given estimates of the parameters of the macro term-structure model and the two additional restrictions in equations (10) and (14), we can identify the target rate parameters, τ_0, τ_π, and τ_g, as well as the factor loadings for policy discretion, d. Note that once we have identified τ_π and τ_g, identification of τ_0 and d follows from equation (6): $\tau_0 = a_0 - \tau_\pi b_0 - \tau_g c_0$ and $d = a - \tau_\pi b - \tau_g c$. Estimates of these parameters are presented in the bottom panel of Table 1. Asymptotic standard errors are calculated using the delta method. (See the Appendix for details.)

It is both reassuring and surprising that although we have adopted a novel asset-pricing approach for identification, the estimates for the Taylor rule parameters are quite conventional. The coefficient on π_t is 1.6230, which safely satisfies the Taylor-principle stability condition, $\tau_\pi > 1$. It is larger than Taylor's original specification of $\tau_\pi = 1.5$, suggesting a somewhat more aggressive inflation policy. Our estimate of τ_g is 0.6532, which is quite close to Taylor's value of $\tau_g = 0.5$, however, the units are difficult to compare directly as we use output growth rather than deviations from a potential-output trend.

The estimates of the factor loadings for the policy discretion term, d, are significant for only two of the four factors. The loading on the first factor – the most persistent factor – is small and insignificant, which stands in contrast to real GDP growth which has a small but significant loading on that factor. But similar to GDP growth, the policy shock does not appear to depend on the fourth factor. If we interpret that factor as a purely nominal factor, then this suggests that policy discretion is not its source. On the other hand, s_t depends significantly on the second and third factors, but in the opposite direction to GDP growth, which suggests a smoothing role for discretion. This two-factor structure is consistent with the findings of Gürkaynak, Sack, and Swanson (2005b), who exploit high-frequency data around FOMC announcements to measure changes in s_t directly, which they then relate to high-frequency movements in asset-pricing factors.

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 discretionary shock s_t. Notice how much smoothness policy discretion imparts to the short rate relative to the target rate. In addition, we observe large and persistent discretionary tightening prior to the last three recessions, followed by large and persistent discretionary easing after those recessions. The zero lower bound shows up as large positive discretionary tightening in 2009-2010, followed by persistent discretionary easing over the subsequent post-crisis years of our sample.

How exactly does our identification strategy separate the observable effects of i_t^R and s_t? Our numerical estimates may help clarify this natural question. Consider two models, the first has no policy discretion so that $i_t = \hat{\tau}_0 + \hat{\tau}_\pi \pi_t + \hat{\tau}_g g_t$. The implied values are $\hat{\tau}_\pi = 2.5590$ and $\hat{\tau}_g = 0.8371$ (using d_2 and d_3), which are not unreasonable but imply a more aggressive policy than most central banker's would likely admit to. The second model has discretion policy, but it looks just like the policy rule itself, $s_t = d_\pi \pi_t + d_g g_t$, so that $i_t = \tau_0 + (\tau_\pi + d_\pi)\pi_t + (\tau_g + d_g)g_t$. Obviously, without further structure, the two models are identical. To separate the effect of d_π and d_g from τ_π and τ_g, we also assume that the policy maker wants the nominal interest-rate risk created by their discretion to be uncorrelated with the levels of the long-run real economy. That means we need to impose the two linear restrictions given by equations (10) and (14) on this second model. This results in parameters for this theory of policy discretion of $d_\pi = 0.9360$ and $d_g = 0.1838$, which in turn imply our estimates for the policy rule, $\tau_\pi = 1.6230$ and $\tau_g = 0.6532$. The policy maker's discretion amplifies the response to inflation and output growth of a less aggressive policy rule, but in a limited way that insures that the long-run risk for the levels of the real economy are unaffected.

Is this simple example a good structural model of policy discretion? Almost certainly not since it implies that the short rate is spanned by inflation and output growth, which is easy to reject. But more generally, we don't take a stand on those kinds of questions. Building deeper structural models of policy discretion is beyond the scope of this paper, as is building a deeper structural model of our reduced-form parameters for the pricing kernel, output growth, and inflation. Nonetheless, by identifying the reduced-form parameters of policy discretion using restrictions that are consistent with a wide range of different theories, we have provided an important first step in this broader research program. We know that to be consistent with term-structure evidence, any structural model of preferences must result in a marginal rate of intertemporal substitution that looks like our estimated affine pricing kernel. Now we also know that to be consistent with the evidence, any structural model of policy discretion must have a reduced-form that looks like our estimate of d.

6 Implications of monetary policy shocks

Since we have identified the parameters of the target rate and the policy disturbance as part of a dynamic macro term-structure model, we can use that model to get a better understanding of how policy shocks are related to the rest of the economy. We consider both unconditional moments and dynamic correlations. We also use our identification of the discretionary shock to shed light on some historical *policy conundrums*.

6.1 Unconditional moments

Equation (3) provides us a simple way to decompose the nominal term premiums into a target-rate component and a discretionary policy component. Note that the factor loading for the short rate satisfies the restriction $a = a^R + d$, where $a^R = \tau_\pi b + \tau_g c$

is the policy rate factor loading attributable to the target rate, i_t^R. Therefore, the expected value of the forward premium can be written as

$$\begin{aligned} E[fp_t^{(h)}] = {} & (a^R)^\top \mathcal{A}^{*(h)} A_0^* - (a^R)^\top \mathcal{A}^{*(h)} BB^\top \mathcal{A}^{*(h)\top} (a^R)/2 \\ & + d^\top \mathcal{A}^{*(h)} A_0^* - d^\top \mathcal{A}^{*(h)} BB^\top \mathcal{A}^{*(h)\top} d/2 \\ & - (a^R)^\top \mathcal{A}^{*(h)} BB^\top \mathcal{A}^{*(h)\top} d. \end{aligned} \tag{15}$$

The first line in equation (15) is the part of the average forward premium that is attributable to the target rate, the second line is attributable to policy discretion, and the third line captures their interaction through the "convexity" term.[8]

The blue line in Figure 3 is the fraction of the expected value of the forward premium that we can attribute to policy discretion. You can see that it is very small for short-maturity bonds. Almost all of the average premium for short bonds seems to be associated with the the target rate. This share rises steadily, however, before settling down at approximately 20% for long maturity bonds.

We can do a similar exercise for the unconditional variance of the forward premium. Equation (3) implies that the variance of the forward premium can be written as

$$\begin{aligned} \text{Var}[fp_t^{(h)}] = {} & (a^R)^\top (A^{*h} - A^h) V_x (A^{*h} - A^h)^\top (a^R) \\ & + d^\top (A^{*h} - A^h) V_x (A^{*h} - A^h)^\top d \\ & + 2(a^R)^\top (A^{*h} - A^h) V_x (A^{*h} - A^h)^\top d, \end{aligned} \tag{16}$$

where once again we interpret the first term in equation (16) as the part of the variance of the forward premium attributable to the target rate, the second term is the part attributable to policy discretion, and the third term captures the covariance between i_t^R and s_t.

The red line in Figure 3 plots the share of the variance of the forward premium that we can attribute to policy discretion. It is at more than 15% for all maturities and rises to approximately 20% around one-year maturities. In other words, a nontrivial fraction of the volatility we see in nominal risk premiums is associated with volatility in discretionary policy.

[8] Following Campbell and Ammer (1993), we apportion the conditional covariance that appears in the "convexity" term of the unconditional mean of the forward premium equally across the two variables. We do the same for the unconditional variance of the forward premium in equation (16).

6.2 Dynamic properties

Our model has a VAR structure, so it is tempting to consider impulse response functions as a way to track marginal dynamic responses to specific shocks. However, our identifying assumptions do not create an orthogonal system of shocks. Rather, the identified policy shocks are potentially affected by the entire vector of innovations, w_t, as are the other variables in the model.

Therefore, we measure the average dynamic response of an endogenous variable, say z_t, with a factor loading of, say β, *i.e.*, $z_t = \beta^\top x_t$, to a shock to monetary policy, $i_{t+1} - E_t i_{t+1}$, as the dynamic covariance:

$$\mathrm{Cov}_t(z_{t+j}, i_{t+1} - E_t i_{t+1}) = \beta^\top A^{j-1} B B^\top a,$$

for $j \geq 1$. In fact, given our separate identification of the target rate, i_t^R, from the policy discretion term, s_t, we can say even more. To measure the average dynamic response of an endogenous variable to a shock to the target rate, $i_{t+1}^R - E_t i_{t+1}^R$, we can calculate the dynamic covariance:

$$\mathrm{Cov}_t(z_{t+j}, i_{t+1}^R - E_t i_{t+1}^R) = \beta^\top A^{j-1} B B^\top (a^R).$$

Likewise, to measure the average dynamic response of an endogenous variable to a shock to policy discretion, $s_{t+1} - E_t s_{t+1}$, we can calculate the dynamic covariance:

$$\mathrm{Cov}_t(z_{t+j}, s_{t+1} - E_t s_{t+1}) = \beta^\top A^{j-1} B B^\top d.$$

By substituting the appropriate factor loadings for output growth, inflation, bond yields, *etc.*, in place of β in these formulas, we can get a sense of how the different sources of monetary policy shocks are related to different aspects of the economy. To control for scale we report these covariances as correlation coefficients. And again we caution the reader not to interpret these as impulse responses. Even though they are similar in appearance, as bivariate correlations, they contain different information.

In Figure 4, we plot these dynamic correlations for the two macro variables in our empirical model, inflation and real output growth. The top panels display the combined effect of a policy shock. The lower panels decompose that shock into its two components, with the blue line representing a shock to the target rate and the red line representing a shock to policy discretion. (Dashed lines represent 2-standard deviation confidence bounds.)

We find that an unexpected shock to the policy rate is positively correlated with unexpected inflation. The correlation is large and persistent remaining significantly different from zero out to about 3 years. That is, inflation today is still correlated with policy shocks from three years ago. The two components of the policy shock have similar effects, but policy discretion is less correlated and that correlation is less persistent. Our model fails to find much of a relationship between unexpected shocks to real output growth and unexpected shocks to the policy rate. However, that appears to be the result of offsetting effects of the two components of that policy shock. A shock to the target rate is positively correlated with shocks to output growth, whereas an unexpected shock to policy discretion is negatively correlated. Both of these are significant, but neither is very persistent becoming indistinguishable from zero after 2 or 3 quarters.

To summarize, unexpected easing through the target rate tends to coincide with good macroeconomic news on inflation and bad news on growth. This is almost a necessary feature of the model given the structure of the Taylor rule. Discretionary easing also tends to coincide with good news on inflation, but in contrast to the target rate, it coincides with good news on growth. And as discussed in Section 4, although our identifying assumptions necessarily places a restriction across the parameters of the macro term-structure model and the discretionary shock, they do not dictate either the size or the sign of these correlations. Rather these correlations are implications of the data.

In Figure 5 we repeat this exercise for the short rate, *i.e.*, the policy rate, and the five-year bond yield. The response of the short rate to a policy shock is obviously perfectly correlated with itself, but it is also persistently correlated with future interest rates remaining significant beyond 3 years. And echoing the behavior we saw for inflation, the response of the short rate is similar for shocks coming from either source, with the correlation of policy discretion being both smaller and less persistent.

On the right side of Figure 5 we see that an unanticipated shock to the policy rate is positively correlated with shocks to the long end of the yield curve measured here with the 5-year bond yield. And that correlation is also quite persistent. This is consistent with evidence that motivates the search for additional sources of monetary policy shocks as in Boyarchenko, Haddad, and Plosser (2017), Gürkaynak, Sack, and Swanson (2005a), and Hanson and Stein (2015). What we see from the lower panel, however, tells a very different story depending on the source of the policy shock. Shocks to the long bond yield are not correlated with discretionary policy shocks. The relationship between a policy-rate shock and shocks to long bond yields is attributable entirely to the correlation with the shock to the target rate. Once again, this is not a feature of the model or the way we identify the policy discretion term. In principle our model has the freedom to take on virtually any correlation pattern with any variable. This outcome is strictly a feature of monetary policy as reflected through our data.

We can also think of these patterns in terms of the customary level and slope factors of empirical term-structure models. A monetary policy easing originating in a shock to the target rate tends to coincide with an unexpected decrease in the level of the yield curve since its behavior is similar at both the long and short ends of the curve. Discretionary policy easing, however, tends to coincide with an unexpected increase in the slope of the yield curve since it is only related to movements at the short end of the curve. These correlations are depicted in Figure 6.

Long-term bond yields embed forecasts of future interest rates as well as risk premiums. We have already observed that policy-rate shocks from both sources have persistent positive correlations with future interest rates. But what about risk premiums? Nominal forward premiums given in equation (3) provide a clean decomposition of these two effects. Figure 7 plots dynamic correlations of forward premiums shocks with policy-rate shocks for short-maturity forward rates of 3 months and long-maturity forward rates of maturity 5 years. The top panels suggest that nominal risk premium shocks may have a small and short-lived negative correlation with policy-rate shocks. However, once again, this is masking two very different and offsetting effects for each of the sources of the policy-rate shock. The correlation of risk premiums at the very short end of the maturity structure with a shock to policy discretion is positive but short-lived. The correlation of risk premium on long-maturity bonds, on the other hand, are negative and more persistent (significant out to about one year). In contrast, the correlation with shocks to the target rate are negative for both long and short maturities, but the lagged correlation becomes significantly positive for long maturities. In other words, the risk characteristics of monetary policy shocks depends very much on the source of those shocks.

To summarize, an unanticipated easing originating with the target rate tends to be associated with bad news about financial conditions summarized by higher-than-expected risk premiums on nominal bonds of both long and short maturities. On the other hand, an unanticipated easing of discretionary policy tends to coincide with good news in the form of lower-than-expected risk premiums on short-term nominal bonds, but bad news in the form of higher-than-expected risk premiums on long-term nominal bonds.

6.3 Policy conundrums

Without a structural model of the fundamental source for policy discretion we can only speculate on the cause of the correlations we find. However, these results do cast some light on the so-called conundrum that has puzzled policy makers in the past: long nominal bond yields often move in ways that appear disconnected from discretionary policy as depicted in Figure 5. 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 our results and Figure 2 suggest is that the policy tightening through the tightening of the target rate was larger than the overall increase in the policy rate and that discretionary easing continued until 1995. As we've seen, target rate shocks tends to coincide with increases in the level of long bond yields, whereas discretionary easing is unrelated to those yields. In other words, this particular combination of target-rate and discretionary policies tends to coincide with both an increase in the level of the yield curve *and* and increase in its slope, which seems to be what Chairman Greenspan found puzzling.

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 our findings and Figure 2 suggest is that much of the tightening of the target rate in the years prior to this episode was undone through the discretionary part of policy. Discretionary easing had averaged about 200 basis points in the previous three years. Discretionary tightening then averages more than 200 basis points in the subsequent three years. In this case, this particular combination of target-rate and discretionary policies tends to coincide with not just an increase in the level of the yield curve as Chairman Greenspan expected, but also a decrease in its slope.

To get a visual image of these correlations, the lower panel of Figure 2 plots the discretionary shock, s_t, along with a standard measure of the slope of the yield curve, the 5-year forward spread, $f_t^{(20)} - i_t$. The negative correlation between shocks to discretionary monetary policy and the slope of the yield seems to be a fairly consistent pattern throughout this time period, not just during Greenspan's conundrums.

7 Conclusion

We estimate a macro 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 arrive at a novel identification of shocks to discretionary monetary policy. We explore the properties of discretionary shocks through their dynamic correlations with shocks to macroeconomic and financial market conditions. To arrive at a deeper understanding of the causes and consequences of the empirical facts we uncover, the challenge we now face is to develop plausible structural models of the fundamental sources of discretionary shocks that are capable of accounting for the strong connection we see in the data between these shocks and financial-market conditions.

The value we derive from integrating asset-pricing models with time-varying risk premiums into the identification and estimation of monetary policy was foreshadowed by Backus and Wright (2007). They concluded that analyzing monetary policy through the narrow lens of constant risk premiums was problematic: "We follow a long line of work in suggesting that expectations-hypothesis intuition, based on constant term premiums, is likely to be misleading." They saw a need for research that connected monetary policy to fundamental shocks and ultimately to risk premiums embedded in interest rates: "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 believe the findings summarized in this paper demonstrate the value of this insight, and that future work in this area will continue to benefit from the integration of models of macroeconomic policy with models of asset pricing.

References

Abel, Andrew B., 1999, "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics* 43, 3-33.

Alvarez, Fernando, and Urban J. Jermann, 2005, "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica*, 73, 1977-2016.

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., Mikhail Chernov, and Stanley E. Zin, 2015, "Identifying Taylor rules in macro-finance models," NBER Working Paper 19360.

Backus, David K., and Jonathan H. Wright, 2007, "Cracking the conundrum," Finance and Economics Discussion Series 2007-46, Federal Reserve Board.

Bernanke, Ben S., 2005, "Monetary policy and the housing bubble," Speech made at the Annual Meeting of the American Economic Association, January 3.

Bernanke, Ben S., 2013, " Long-term interest rates," Remarks at the Federal Reserve Bank of San Francisco, March 1.

Beveridge, Stephen, and Charles Nelson, 1982, "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary Economics* 7, 151-174.

Bikbov, Ruslan, and Mikhail Chernov, 2010,"No-arbitrage macroeconomic determinants of yield curves," *Journal of Econometrics* 159, 166-182.

Blanchard, Olivier Jean, and Danny Quah, 1989, "The dynamic effects of aggregate demand and supply disturbances," *American Economic Review* 79, 655-673.

Boyarchenko, Nina, Valentin Haddad, and Matthew Plosser, 2017, "The federal reserve and market confidence," Federal Reserve Bank of New York Staff Report No. 773.

Campbell, J. Y., and J. Ammer, 1993, "What moves the stock and bond markets? A variance decomposition for long-term asset returns," *Journal of Finance* 48, 3-37.

Campbell, John Y., and Robert J. Shiller, 1988, "Stock prices, earnings, and expected dividends," *Journal of Finance* 43, 661-676.

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

Clarida, Richard, Jordi Galí, and Mark Gertler, 1998, "Monetary policy rules in practice: Some international evidence," *European Economic Review* 42, 1033-1067.

Clarida, Richard, Jordi Galí, and Mark Gertler, 2000, "Monetary policy rules and macroeconomic stability: Evidence and some theory," *Quarterly Journal of Economics* 115, 147-180.

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

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.

Engle, Robert, and Clive Granger, 1991, "Long-run economic relationships: Readings in cointegration," Oxford University Press.

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.

Goodfriend, M., and R. G. King, 2005, "The incredible Volcker disinflation," *Journal of Monetary Economics* 52, 981-1015.

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., B. Sack, and E. Swanson, 2005a, "The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models," *American Economic Review* 95, 425-436.

Gürkaynak, Refet S., B. Sack, and E. Swanson, 2005b, " Do actions speak louder than words? The response of asset prices to monetary policy actions and statements," *International Journal of Central Banking* 1, 55-93.

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.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2010, "The TIPS yield curve and inflation compensation," *American Economic Jouranl: Macroeconomics* 2, 70-92.

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 Jose Scheinkman, 2009, "Long term risk: An operator approach," *Econometrica* 77, 177-234.

Hanson, S. G. and J. C. Stein, 2015, "Monetary policy and long-term real rates," *Journal of Financial Economics* 115, 429-448.

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

Kuttner, Kenneth N., 2001, "Monetary policy surprises and interest rates: Evidence from the Fed funds futures market," *Journal of Monetary Economics* 47, 523–544.

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.

Newey, Whitney K., 1985, "Generalized method of moments specification testing," *Journal of Econometrics*, 29, 229-256.

Piazzesi, Monika, and Eric Swanson, 2008, "Futures prices as risk-adjusted forecasts of monetary policy," *Journal of Monetary Economics* 55, 677-691.

Rudebusch, Glenn D., 2006, "Monetary policy inertia: Fact or fiction?" *International Journal of Central Banking* 2, 85-135.

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.

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

Stock, James H., and Mark W. Watson, 2007, "Why has US inflation become harder to forecast?" *Journal of Money, Credit and Banking* 39, 3-33.

Taylor, John B., 1993, "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.

Taylor, John. B., 1999, "A historical analysis of monetary policy rules," in Monetary Policy Rules, ed. John B. Taylor, National Bureau of Economic Research.

Table 1. GMM estimation

I. State dynamics

A				B			
0.9169	0.2825	0.2617	0.6601	0.0022	0	0	0
(0.0400)	(0.0459)	(0.0588)	(0.1311)	(0.0002)			
0.4680	0.5501	-0.0713	-0.9401	-0.0056	0.0110	0	0
(0.1885)	(0.0623)	(0.0354)	(0.4700)	(0.0030)	(0.0048)		
-0.4394	0.1510	0.7531	0.6428	0.0042	-0.0111	0.0022	0
(0.1676)	(0.0766)	(0.0686)	(0.4458)	(0.0031)	(0.0051)	(0.0002)	
0.0028	-0.0668	-0.0570	0.4147	-0.0004	0.0010	-0.0010	0.0014
(0.0036)	(0.0215)	(0.0287)	(0.0722)	(0.0003)	(0.0005)	(0.0003)	(0.0001)

II. Term-structure model

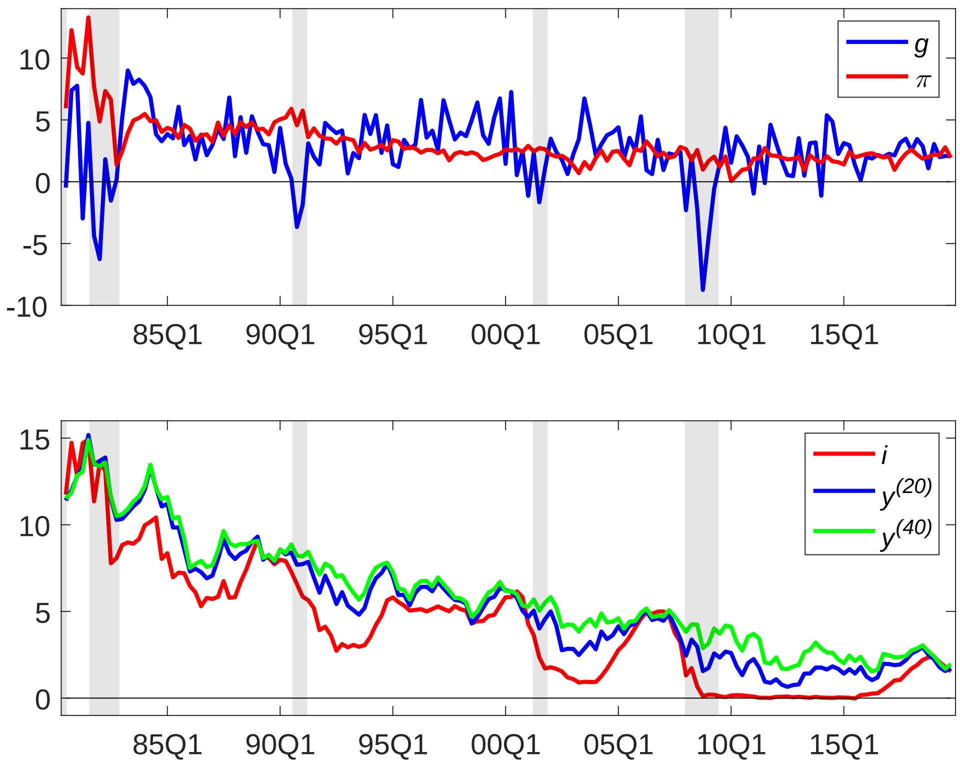
A^*				λ_0	b	c
0.9961	0	0	0	-0.0625	0.4947	0.0158
(0.0006)				(0.0053)	(0.0060)	(0.0071)
0	0.8814	0	0	-0.1421	0.4243	-0.1025
	(0.0020)			(0.0191)	(0.0126)	(0.0379)
0	0	0.8348	0	-0.0128	0.4751	-0.2578
		(0.0035)		(0.0355)	(0.0169)	(0.0438)
0	0	0	0.3638	-0.5803	0.5927	0.0311
			(0.0102)	(0.0777)	(0.0181)	(0.1221)

III. Taylor rule

τ_π	τ_g	$d^\top$			
1.6230	0.6532	0.1868	0.3783	0.3973	0.0177
(0.3122)	(0.2164)	(0.1549)	(0.1471)	(0.1939)	(0.1732)

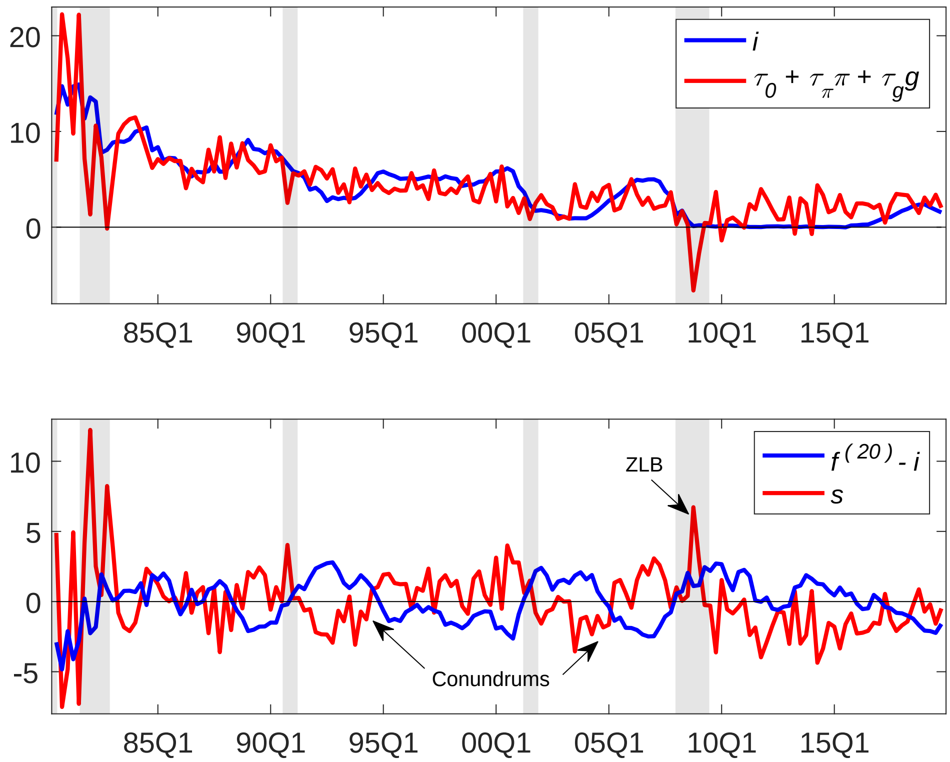
Note. Based on moment restrictions in equations (22) and (23) for a sample period 1980Q3 to 2019Q4. Asymptotic standard errors are in parentheses. State dynamics: $x_{t+1} = Ax_t + Bw_{t+1}$. Macro term-structure model: $i_t = a_0 + a^\top x_t$, $\pi_t = b_0 + b^\top x_t$, $g_t = c_0 + c^\top x_t$, $hy_t^{(h)} = \mathcal{B}_0^{(h)} + \mathcal{B}^{(h)} x_t$, and $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1}(I - A^{*h})$, $\mathcal{B}_0^{(h)} = a_0 + \mathcal{B}_0^{(h-1)} - \mathcal{B}^{(h-1)} B\lambda_0 - \mathcal{B}^{(h-1)} BB^\top \mathcal{B}^{(h-1)\top}/2$. Policy: $i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d^\top x_t$. 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 $h = 4$, 12, 20, 40 (quarters), π_t is the inflation rate, g_t is the growth rate of real GDP, and $a^\top = [1\ 1\ 1\ 1]$. Values for intercepts are fixed at their sample means. The absolute value of the eigenvalues of A are 0.9775, 0.6705, 0.6705, and 0.3180.

Figure 1. US GDP growth, CPI inflation, and yields



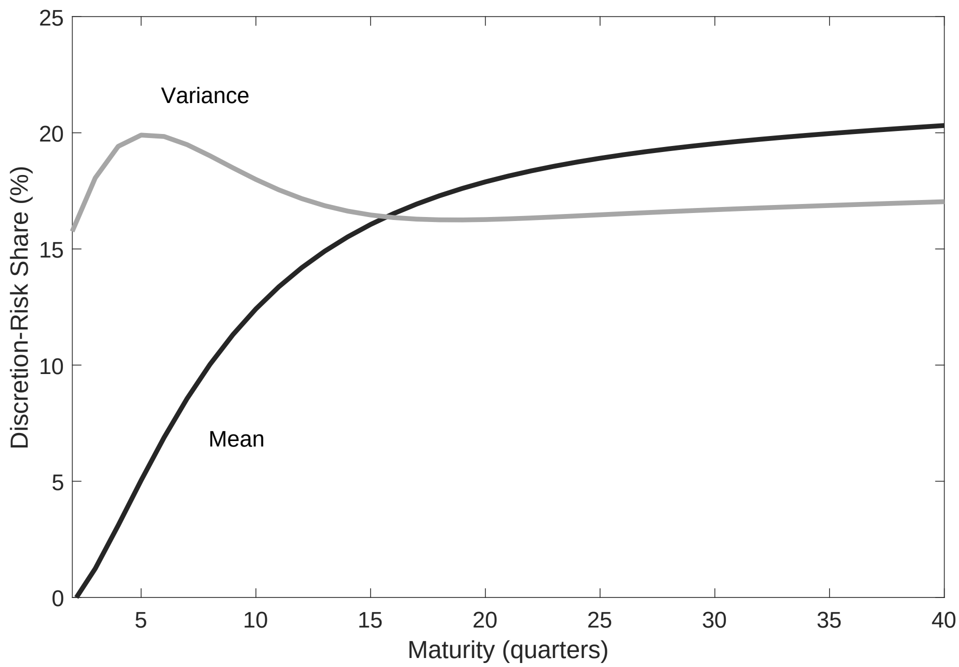
Note: The time period is 1980Q3 to 2019Q4. Real GDP growth is from the NIPA and CPI inflation is from the BLS, both downloaded from FRED. The short rate is from Fama and Bliss (available from CRSP), and yields are from Gürkaynak, Sack, and Wright (2007). (Along with these 3 maturities, we also used a 12-quarter yield in our estimation, which is not plotted to avoid making the graph too dense.) All variables are percentages, continuously compounded at annual rates.

Figure 2. Target-rate and discretionary policy shocks



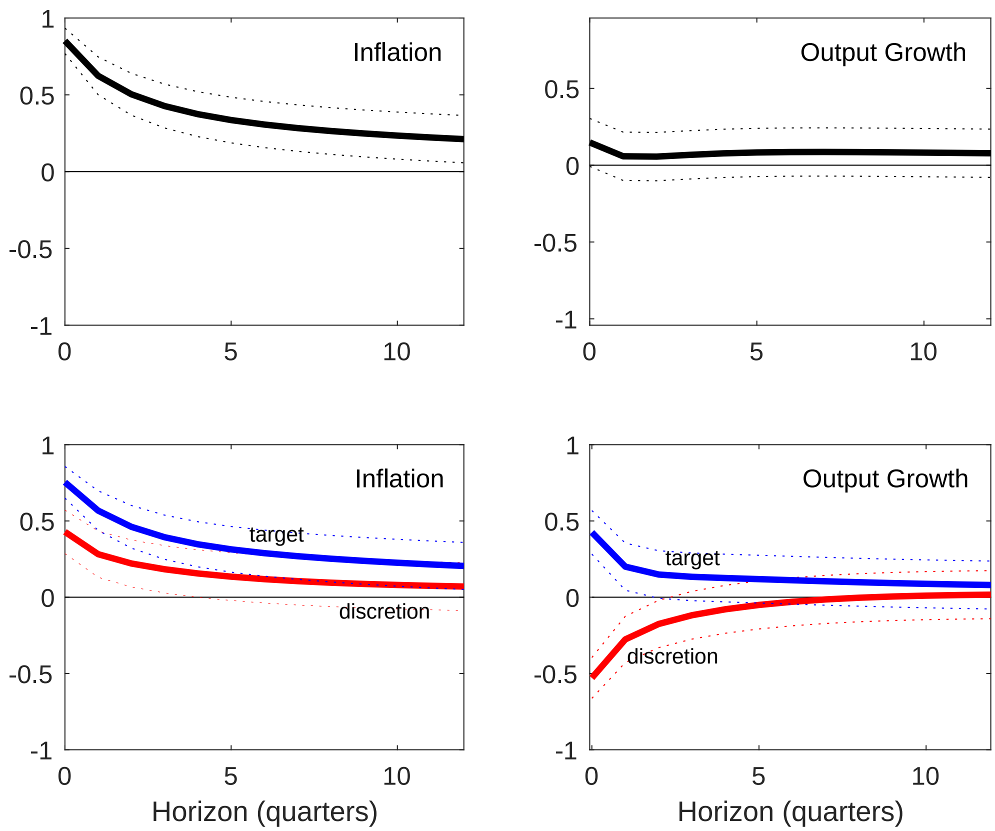
Note: The top panel plots the policy rate including the discretionary shock, *i.e.*, the short rate, and excluding the shock, *i.e.* the target rate, $i_t^R = \tau_0 + \tau_\pi \pi_t + \tau_g g_t$ using estimated parameter values for τ_0, τ_π, and τ_g from Table 1. The difference is the value of the shock, s_t, plotted in the lower panel along with the 5-year forward spread, $f_t^{(20)} - i_t$, *i.e.*, a long forward rate minus the current interest rate.

Figure 3. Forward premium share of discretionary policy



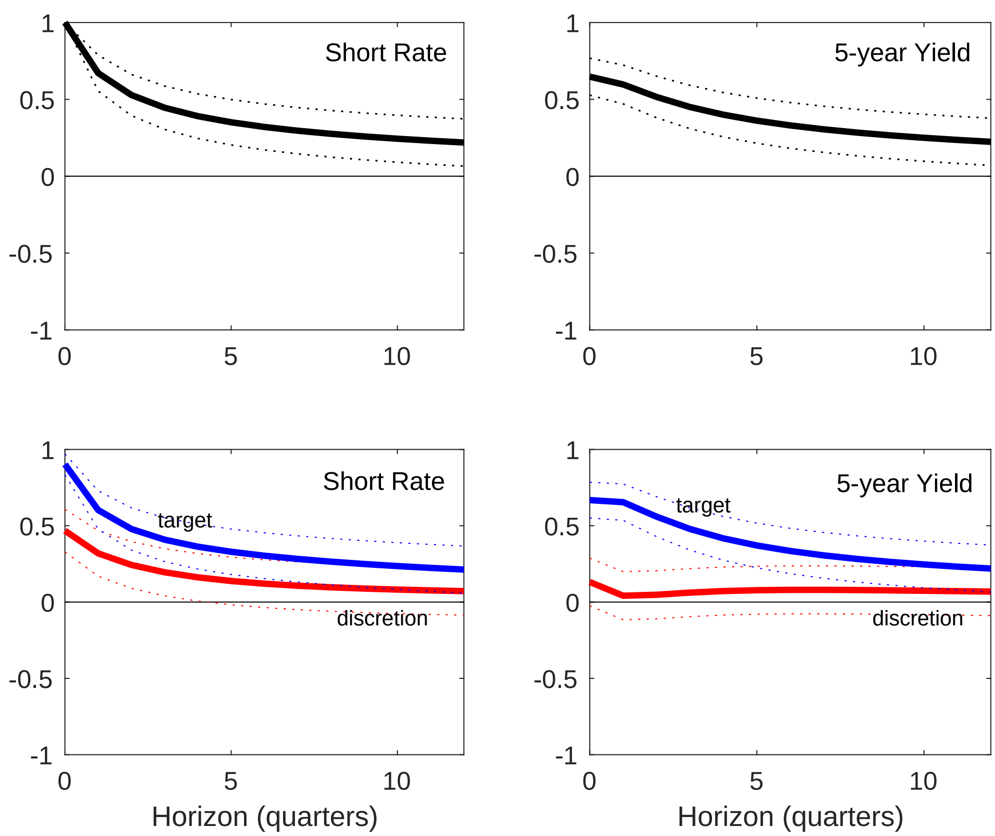
Note: The blue line plots the fraction of the unconditional mean of the forward premium that is attributable to the discretionary policy shock. The red line plots the fraction of the unconditional variance of the forward premium that is attributable to the discretionary policy shock.

Figure 4. Dynamic correlations: Macro variables



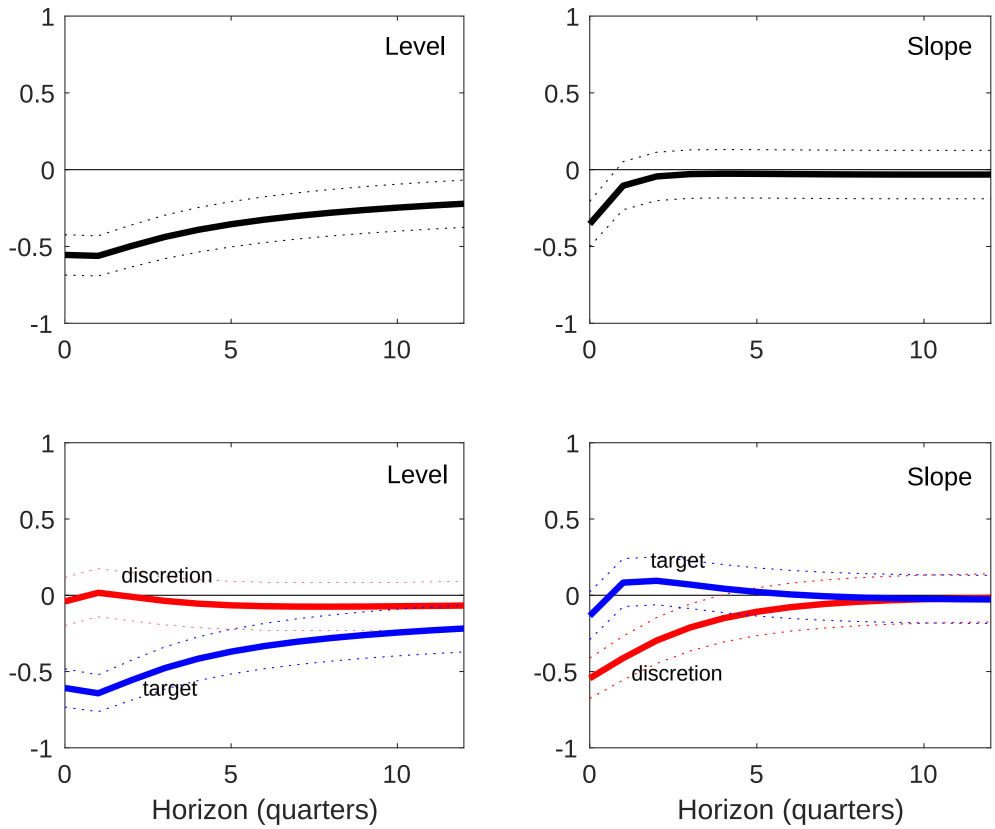
Note: The top panels plots the dynamic correlation with a shock to the policy rate for inflation and real output growth. The lower panels decompose this into the two sources for a policy shock, with the blue line representing a shock to the target rate and the red line representing a shock to policy discretion. Dashed lines represent 2-standard-deviation confidence bounds.

Figure 5. Dynamic correlations: Interest rates



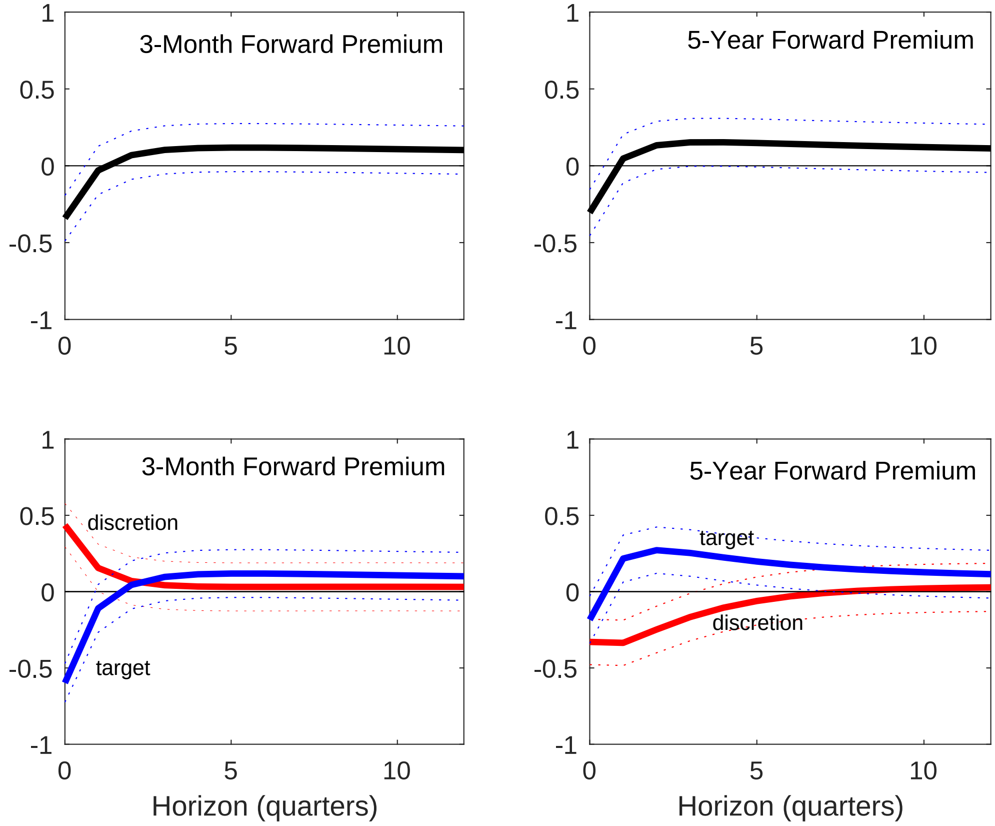
Note: The top panels plots the dynamic correlation with a shock to the policy rate for the short rate and the 5-year bond yield. The lower panels decompose this into the two sources for a policy shock, with the blue line representing a shock to the target rate and the red line representing a shock to policy discretion. Dashed lines represent 2-standard-deviation confidence bounds.

Figure 6. Dynamic correlations: Level and slope factors



Note: The top panels plots the dynamic correlation with a shock to the policy rate for the the level and slope factors calculated from the first two principal components of the variance of a vector of forward rates for maturities 1 to 5 years. The lower panels decompose this into the two sources for a policy shock, with the blue line representing a shock to the target rate and the red line representing a shock to the policy disturbance. Dashed lines represent 2-standard-deviation confidence bounds.

Figure 7. Dynamic correlations: Risk premiums



Note: The top panels plots the dynamic correlation with a shock to the policy rate for the forward risk premiums at horizons of 3-month and 5-years. The lower panels decompose this into the two sources for a policy shock, with the blue line representing a shock to the target rate and the red line representing a shock to the policy disturbance. Dashed lines represent 2-standard-deviation confidence bounds.

Appendix

A Equilibrium term structure

Given the nominal pricing kernel in equation (2) and the linear transition equation (1), the absence of arbitrage implies that the date-t price, $q_t^{(h)}$, of a default-free pure-discount bond that pays \$1, at date $t+h$, $h>0$, is log-linear in the state:

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

where

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

and $q_t^{(0)} = 1$ implies the initial conditions $\mathcal{B}_0^{(0)} = 0$ and $\mathcal{B}^{(0)} = 0$. A_0^* is referred to as the *risk-neutral* mean of x_t, and A^* is referred to as the *risk-neutral* persistence of x_t associated with the pricing kernel $m_{t+1}^{\$}$. That is, if the dynamics of the state variable were given by the *risk-neutral* process $x_t = A_0^* + A^* x_{t-1} + Bw_t$, then the date-$t$ price of any arbitrary random future nominal payoff, say $F(x_{t+1})$, would be given by $e^{-i_t} E_t^*[F(x_{t+1})]$, where E^* denotes the expected value using the risk-neutral process.

B Lack of identification in macro models

Cochrane's example showing the lack of identification of Taylor rule parameters extends to models with more macroeconomic structure. Demonstrating this for a number of standard models is instructive both for understanding the general identification problem and our particular identification strategy.

B.1 A structural real pricing kernel

Consider a model in which the real interest rate is endogenous and correlated with real output growth as in Gallmeyer, Hollifield, Palomino, and Zin (2007). A simplified version of the log of the real pricing kernel that is commonly used in structural macro models is

$$-m_{t+1} = \delta_0 + \delta g_{t+1},$$

where δ_0 and δ are structural parameters. Assume, for the sake of simplicity, that δ_0 and δ are known. we maintain the usual notation for the state variable's dynamics, $x_t = Ax_{t-1} + Bw_t$, and $g_{t+1} = c_0 + c^\top x_{t+1}$, and assume that c_0, c, A, and B are also known. The log of the nominal pricing kernel is given by in (2),

$$\begin{aligned} -m^\$_{t+1} &= \delta_0 + \delta g_{t+t} + \pi_{t+1} \\ &= \delta_0 + \delta c_0 + b_0 + (\delta c + b)^\top A x_t + (\delta c + b)^\top B w_{t+1}. \end{aligned}$$

This model is equivalent to the more general affine model in (2) with the added parameter restrictions

$$\begin{aligned} a^\top &= (\delta c + b)^\top A \\ \lambda &= 0 \\ a_0 &= \delta_0 + \delta c_0 + b_0 - (\delta c + b)^\top B B^\top (\delta c + b)/2 \\ \lambda_0^\top &= (\delta c + b)^\top B. \end{aligned} \tag{17}$$

None of these restrictions involve τ_0, τ_π, τ_g, or d, hence, even though they serve to over-identify the parameters of the macro term-structure model, they play no role in identifying the Taylor rule. The system of equations in (9) is now

$$\begin{aligned} b^\top &= [d^\top + c^\top(\tau_g I - \delta A)](A - \tau_\pi I)^{-1} \\ b_0 &= [\tau_0 + (\tau_g - \delta)c_0 - \delta_0 + (\delta c + b)^\top B B^\top (\delta c + b)/2]/(1 - \tau_\pi), \end{aligned} \tag{18}$$

which is a system of $n + 1$ equations in $n + 3$ unknown policy parameters, τ_0, τ_π, τ_g, and d. Even with this additional economic structure, the policy rule is still under-identified.

B.2 A model with a Phillips curve

A similar result holds for models that introduce a correlation between inflation and real output growth through a Phillips curve, as in Gallmeyer, Hollifield, and Zin

(2005), and many others. For example, consider adding to Cochrane's pricing kernel a simple Phillips curve given by

$$\pi_t = \kappa_\pi E_t \pi_{t+1} + \kappa_g g_t,$$

where κ_π and κ_g are structural parameters. Assume again, for the sake of simplicity, that κ_π and κ_g are also known. We now have a two-dimensional rational expectations model in the forward-looking endogenous variables π_t and g_t. Maintaining the same notation, we can denote the solutions for these variables as $\pi_t = b_0 + b^\top x_t$ and $g_t = c_0 + c^\top x_t$. The Phillips curve imposes a restriction across these two processes

$$\begin{aligned} c^\top &= b^\top (I - \kappa_\pi A)/\kappa_g \\ c_0 &= b_0(1 - \kappa_\pi)/\kappa_g. \end{aligned}$$

which then alters the interest rate solution

$$\begin{aligned} b^\top &= d^\top[(1 + \kappa_\pi \tau_g/\kappa_g)A - (\tau_\pi + \tau_g/\kappa_g)I]^{-1} \\ b_0 &= (\tau_0 - r)/[1 - \tau_\pi - \tau_g(1 - \kappa_\pi)/\kappa_g]. \end{aligned} \tag{19}$$

The restrictions in (19) are different in form than those in (9) or (18), but they share the same unavoidable under-identification: $n + 1$ equations with $n + 3$ unknowns. Unless additional restrictions that involve the policy parameters are imposed on the model, monetary policy remain unidentified. And this is in the best-case scenario in which there are no identification issues with any other parameters of the model.

B.3 A forward-looking Taylor rule

Many specifications of the Taylor rule include a forward-looking expected inflation term, such as

$$i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d_e E_t \pi_{t+1} + s_t,$$

where d_e is a constant parameter. There are two ways to view this addition. In our terminology, the expected-inflation component could be considered part of the "rule" or it could be considered policy "discretion". This distinction is irrelevant for the solution of the model, but we will return to it later so that we can better align this example with the approach of our empirical model.

Assume that the log of the real pricing kernel is given by an AR(1) as in the first example above,

$$-m_{t+1} = \delta_0 + \delta c_0 + \delta c^\top A x_t + \delta c^\top B w_{t+1}.$$

Given the policy rule and the real pricing kernel, we can solve for the inflation process consistent with an equilibrium nominal short rate. Once again, we guess a linear solution,

$$\pi_t = b_0 + b^\top x_t,$$

so that expected future inflation is given by

$$E_t\pi_{t+1} = b_0 + b^\top Ax_t.$$

Once again, the log of the nominal pricing kernel combines the log of the real kernel with the inflation rate:

$$\begin{aligned}-m^{\$}_{t+1} &= -m_{t+1} + \pi_{t+1}\\ &= \delta_0 + \delta c_0 + b_0 + (\delta c + b)^\top Ax_t + (\delta c + b)^\top Bw_{t+1},\end{aligned}$$

which implies a nominal interest rate,

$$\begin{aligned}i_t &= -\log E_t e^{m^{\$}_{t+1}}\\ &= a_0 + (\delta c + b)^\top Ax_t.\end{aligned}$$

Therefore, the inflation process consistent with an equilibrium short rate must satisfy the equation

$$a_0 + (\delta c + b)^\top Ax_t = \tau_0 + \tau_\pi\pi_t + \tau_g g_t + d_e E_t\pi_{t+1} + s_t.$$

The factor loading, therefore, must satisfy

$$(\delta c + b)^\top A = \tau_\pi b^\top + \tau_g c^\top + d_e b^\top A + d^\top,$$

which implies

$$b^\top = [d^\top + c^\top(\tau_g I - \delta A)][(1 - d_e)A - \tau_\pi I]^{-1},$$

which reduces to (18) when $d_e = 0$. There are three things to note about this equilibrium. Firstly, the parameter d_e affects the factor loadings for inflation and the nominal interest rate and, hence, their means and volatilities. But d_e does not affect the dynamics of either process: their autocorrelation functions are still determined solely by the matrix A. Secondly, augmenting the Taylor rule to include a forward-looking expected inflation term did not change the basic identification problem. Even if d_e were known with certainty, τ_π and τ_g are still unidentified. And if d_e is also unknown, the lack of identification is even worse: we would need yet another restriction on the model to separately identify d_e beyond the restrictions that we need to identify τ_π and τ_g. Finally, note that our empirical analysis is consistent with this interest-rate rule. The policy discretion term absorbs the effect of forward-looking policy and the factor loadings for discretion are $[d_e b^\top A + d^\top]$.

B.4 A Taylor rule with a lagged interest rate

Popular applications of the Taylor rule often include a lagged interest-rate term, presumably to capture the central bank's desire to smooth interest-rate changes. To see how this interest-rate smoothing behavior is captured by the policy discretion term in our abstract state-space representation, it is instructive to work out a simple example.

We continue with an AR(1) representation for the log of the real pricing kernel:

$$-m_{t+1} = \delta_0 + \delta c_0 + \delta c^\top x_{t+1}.$$

The Taylor rule now includes a lagged interest rate to capture discretionary smoothing:

$$i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d_s i_{t-1} + s_t,$$

where d_s is a constant parameter. Evidently, for this equation to hold, the equilibrium nominal interest-rate process must be an infinite-order distributed lag in π_t, g_t and s_t, which suggests that to solve for the equilibrium, we need to work with a state space that defines the infinite-order MA representation of x_t,

$$\{Bw_t, Bw_{t-1}, Bw_{t-2}, \ldots\}.$$

We then guess a linear process for the equilibrium inflation rate that can accommodate this infinite-order distributed lag structure,

$$\pi_t = b_0 + \gamma_0^\top Bw_t + \gamma_1^\top Bw_{t-1} + \gamma_2^\top Bw_{t-2} + \ldots,$$

where b_0 and γ_j, $j = 0, 1, 2, \ldots$ are to be found. The log of the nominal pricing kernel combines the log of the real kernel with the inflation rate as before:

$$\begin{aligned} -m_{t+1}^\$ &= -m_{t+1} + \pi_{t+1} \\ &= \delta_0 + \delta c_0 + b_0 + \sum_{j=0}^{\infty} (\delta c^\top A^j + \gamma_j^\top) Bw_{t+1-j}, \end{aligned}$$

which implies a nominal interest rate,

$$i_t = a_0 + \sum_{j=0}^{\infty} (\delta c^\top A^{j+1} + \gamma_{j+1}^\top) Bw_{t-j}.$$

Therefore, the inflation process consistent with an equilibrium short rate must satisfy the equation

$$
\begin{aligned}
a_0 + \sum_{j=0}^{\infty} (\delta c^\top A^{j+1} + \gamma_{j+1}^\top) B w_{t-j} &= \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d_s i_{t-1} + s_t \\
&= \tau_0 + \tau_\pi b_0 + \tau_g c_0 + d_s a_0 \\
&\quad + \tau_\pi \sum_{j=0}^{\infty} \gamma_j^\top B w_{t-j} + \tau_g \sum_{j=0}^{\infty} c^\top A^j B w_{t-j} \\
&\quad + d_s \sum_{j=0}^{\infty} (\delta c^\top A^{j+1} + \gamma_{j+1}^\top) B w_{t-1-j} + \sum_{j=0}^{\infty} d^\top A^j B w_{t-j}.
\end{aligned}
$$

Aligning terms we have

$$
\begin{aligned}
\gamma_0^\top &= [c^\top(\delta A - \tau_g I) - d^\top]/\tau_\pi + \gamma_1^\top/\tau_\pi \\
\gamma_j^\top &= [c^\top(\delta A - (\tau_g + d_s\delta)I) - d^\top]A^j/(\tau_\pi + d_s) + \gamma_{j+1}^\top/(\tau_\pi + d_s), \quad j \geq 1.
\end{aligned}
$$

If $A/(\tau_\pi + d_s)$ is a stable matrix, we can solve these recursions to obtain

$$
\begin{aligned}
\gamma_0^\top &= [d^\top + c^\top(\tau_g I - \delta A)][A - (\tau_\pi + d_s)I]^{-1}(\tau_\pi + d_s)/\tau_\pi + d_s \delta c^\top A[A - (\tau_\pi + d_s)I]^{-1}/\tau_\pi \\
\gamma_1^\top &= [d^\top + c^\top((\tau_g + d_s\delta)I - \delta A)][A - (\tau_\pi + d_s)I]^{-1}A \\
\gamma_{j+1}^\top &= \gamma_j^\top A, \quad j \geq 1,
\end{aligned}
$$

which implies that inflation is a vector ARMA(1,1) process. This in turn, implies that the nominal pricing kernel also follows a vector ARMA(1,1) process. On the other hand, since

$$
\delta c^\top A^{j+1} + \gamma_{j+1}^\top = (\delta c^\top A + \gamma_1^\top)A^j, \quad j \geq 0,
$$

the nominal interest rate, i_t, is still an AR(1) with d_s affecting only its factor loadings through γ_1. In other words, the smoothing parameter d_s affects the mean and variance of the interest rate, but perhaps counter-intuitively, *not* its autocorrelation function which is still governed solely by A.

There are a two things to note about this equilibrium. Firstly, the parameter d_s clearly affects the dynamics of the model: when $d_s = 0$, inflation and the log of the nominal pricing kernel exhibit the same AR(1) dynamics as the original state variable, but when $d_s \neq 0$, both of these processes become ARMA(1,1). In both cases, the nominal interest rate has the same AR(1) dynamics as the original state variable, hence, the dynamics of interest rates do not identify the value of d_s. Secondly, augmenting the Taylor rule to include an interest-rate smoothing term did not change the basic identification problem. Even if d_s were known with certainty, τ_π and τ_g are still unidentified. And if d_s is also unknown, the lack of identification is even worse: we would need yet another restriction on the model to separately identify d_s beyond the restrictions that we need to identify τ_π and τ_g.

In the context of our basic state-space model, the presence of a lagged endogenous variable in the Taylor rule via interest-rate smoothing is reflected in a higher-dimensional AR(1) state variable – the abstract state variable would now need to have a dimension $2n$ where n is the dimension of x_t. For example, we could define the abstract state variable as

$$\hat{x}_t = \begin{bmatrix} Bw_t \\ \sum_{j=0}^{\infty} A^j B w_{t-1-j} \end{bmatrix}$$

which will have a VAR(1) representation

$$\hat{x}_t = \hat{A}\hat{x}_{t-1} + \hat{B}\hat{w}_t, \quad \text{where } \hat{A} = \begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix},$$

and $\hat{w}_t = [w_t \;\; 0]^\top$. Expressed as linear functions of this new state variable, the factor loadings for the nominal interest rate, inflation, and the discretionary policy shock are given by

$$\begin{aligned} \hat{a}^\top &= [\delta c^\top A + \gamma_1^\top \quad (\delta c^\top A + \gamma_1^\top)A] \\ \hat{b}^\top &= [\gamma_0^\top \quad \gamma_1^\top] \\ \hat{d}^\top &= [d^\top \quad d^\top A + d_s(\delta c^\top A + \gamma_1^\top)]. \end{aligned}$$

C Entropy of the real pricing kernel

The log real pricing kernel is given by

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

The conditional entropy of the real pricing kernel is

$$\begin{aligned} L_t\big(\exp\{m_{t+1}\}\big) &= \log E_t\big(\exp\{m_{t+1}\}\big) - E_t\big(m_{t+1}\big) \\ &= b^\top B B^\top b/2 + b^\top A_0^* + \lambda_0^\top \lambda_0/2 + [b^\top(A^* - A) + \lambda_0^\top \lambda]x_t + x_t^\top \lambda^\top \lambda x_t/2, \end{aligned}$$

which has an unconditional expected value

$$EL_t\big(\exp\{m_{t+1}\}\big) = b^\top BB^\top b/2 + b^\top A_0^* + \lambda_0^\top\lambda_0/2 + \mathrm{tr}(\lambda V_x\lambda^\top)/2,$$

where we've used the result $E[x_t^\top\lambda^\top\lambda x_t] = \mathrm{tr}[\mathrm{Var}(\lambda x_t)] = \mathrm{tr}(\lambda V_x\lambda^\top)$.

The conditional mean of the real pricing kernel is

$$E_t\exp\{m_{t+1}\} = \exp\{(b_0 - a_0) + b^\top A_0^* + b^\top BB^\top b/2 + (b^\top A^* - a^\top)x_t\},$$

which has an unconditional entropy of

$$\begin{aligned} L\big(E_t\exp\{m_{t+1}\}\big) &= \log E\big(E_t\exp\{m_{t+1}\}\big) - E\log\big(E_t\exp\{m_{t+1}\}\big) \\ &= (b^\top A^* - a^\top)V_x(A^*b - a)/2. \end{aligned}$$

Combine these two components to calculate the unconditional entropy of the real pricing kernel

$$\begin{aligned} L\big(\exp\{m_{t+1}\}\big) &= EL_t\big(\exp\{m_{t+1}\}\big) + L\big(E_t\exp\{m_{t+1}\}\big) \\ &= b^\top BB^\top b/2 + b^\top A_0^* + (b^\top A^* - a^\top)V_x(A^*b - a)/2 \\ &\quad + \lambda_0^\top\lambda_0/2 + \mathrm{tr}(\lambda V_x\lambda^\top)/2. \end{aligned}$$

We can do comparable calculations for the the permanent component of the real pricing kernel whose log is given by

$$\begin{aligned} m_{t+1}^P &= -(b+k)^\top BB^\top(b+k)/2 - (b+k)^\top A_0^* - \lambda_t^\top\lambda_t/2 \\ &\quad + [b^\top A - a^\top - k^\top(I - A)]x_t + (b^\top B - \lambda_t^\top + k^\top B)w_{t+1}. \end{aligned}$$

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

The conditional entropy is given by

$$\begin{aligned} L_t\big(\exp\{m_{t+1}^P\}\big) &= (b^\top + k^\top)BB^\top(b+k)/2 - (b+k)^\top B\lambda_0 + \lambda_0^\top\lambda_0/2 \\ &\quad + [-(b+k)^\top B^\top\lambda + \lambda_0^\top\lambda]x_t + x_t^\top\lambda^\top\lambda x_t/2, \end{aligned}$$

which has an unconditional expected value

$$\begin{aligned} EL_t\big(\exp\{m_{t+1}^P\}\big) &= (b^\top + k^\top)BB^\top(b+k)/2 - (b+k)^\top B\lambda_0 + \lambda_0^\top\lambda_0/2 + E[x_t^\top\lambda^\top\lambda x_t]/2 \\ &= (b^* - a^*)^\top BB^\top(b^* - a^*)/2 + (b^* - a^*)A_0^* + \lambda_0^\top\lambda_0/2 + \mathrm{tr}(\lambda V_x\lambda^\top)/2, \end{aligned}$$

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

The conditional mean of the permanent component is equal to 1 by construction, so it's unconditional entropy is 0 by construction. Therefore, the unconditional entropy of the permanent component is simply the unconditional mean of its conditional entropy

$$L\big(\exp\{m_{t+1}^P\}\big) = (b^* - a^*)^\top BB^\top(b^* - a^*)/2 + (b^* - a^*)^\top A_0^* + \lambda_0^\top\lambda_0/2 + \mathrm{tr}(\lambda V_x\lambda^\top)/2.$$

D Estimation details

D.1 Term-structure moment restrictions

We assume that four bond yields for $h = 1, 4, 12, 20$, where maturity is measured in quarters, can be used to identify the $n = 4$ dimensions of the state. Stack these 4 variables into the vector $z_{1t} = [i_t \;\; y_t^{(4)} \;\; y_t^{(12)} \;\; y_t^{(20)}]^\top$. It will be important to keep in mind when interpreting our empirical results that we are *not* replacing our original state variable, x_t, with this vector of yields, z_{1t}. Rather we are exploiting the equilibrium relationship between these yields and the latent factors to solve for each factor as a particular linear combination of yields as implied by our arbitrage-free model. Therefore, factor loadings and risk prices will retain their original interpretation as responses to the latent state variable x_t. In that spirit, we can rewrite the dynamics for the state as

$$\begin{aligned} z_{1t} = R_0 + Rx_t &= R_0 + R(Ax_{t-1} + Bw_t) \\ &= \tilde{R}_0 + \tilde{A}z_{1t-1} + \tilde{B}w_t, \end{aligned} \tag{20}$$

where $R_0 = [a_0 \;\; \mathcal{B}_0^{(4)}/4 \;\; \mathcal{B}_0^{(12)}/12 \;\; \mathcal{B}_0^{(20)}/20]$, $R = [a \;\; \mathcal{B}^{(4)}/4 \;\; \mathcal{B}^{(12)}/12 \;\; \mathcal{B}^{(20)}/20]^\top$, $\tilde{A} = RAR^{-1}$, $\tilde{R}_0 = (I - \tilde{A})R_0$, and $\tilde{B} = RB$. Recall that the parameters $\mathcal{B}_0^{(h)}$ and $\mathcal{B}^{(h)}$ are functions of the parameters A^*, A_0^* and B. 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, $z_{2t} = [\pi_t \;\; g_t \;\; y_t^{(40)}]^\top$, is also linear in the state variable:

$$\begin{aligned} z_{2t} = &\quad G_0 + Gx_t + u_t \\ = &\quad \tilde{G}_0 + \tilde{G}z_{1t} + u_t, \end{aligned} \tag{21}$$

where $u_t \sim iid \; \mathcal{N}(0, \sigma_u^2 I)$ is interpreted as measurement error with variance σ_u^2. The parameters before the rotation of the state space are $G_0 = [b_0 \;\; c_0 \;\; \mathcal{B}_0^{(40)}/40]^\top$, $G = [b^\top \;\; c^\top \;\; \mathcal{B}^{(40)\top}/40]^\top$, and the parameters $\mathcal{B}_0^{(40)}$ and $\mathcal{B}^{(40)}$ are functions of 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:

$$z_{t+1} | z_t \sim \mathcal{N}(C_0 + Cz_t, DD^\top),$$

where $z_t = [z_{1t} \;\; z_{2t}]^\top$, $C_0 = [\tilde{R}_0^\top \;\; \tilde{G}_0 + \tilde{G}\tilde{R}_0]^\top$, and

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

This 7-variable system has 45 parameters: 16 in A, 10 in B, and 4 each in A^*, b, c, and λ_0, plus the unrestricted intercept parameters, a_0, b_0, and c_0. We begin by estimating these parameters using a just-identified GMM estimator that exactly matches 45 moment restrictions implied by equations (20) and (21):

$$\begin{aligned} E(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1}) &= 0 \\ E(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1}) z_{1t-1}^\top &= 0 \\ E(z_{2t} - \tilde{G}_0 - \tilde{G} z_{1t}) &= 0 \\ E(z_{2t} - \tilde{G}_0 - \tilde{G} z_{1t}) z_{1t}^\top &= 0 \\ E(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1})(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1})^\top &= \tilde{B}\tilde{B}^\top. \end{aligned} \tag{22}$$

Estimation of asymptotic standard errors uses the parameter estimates from these restrictions and the correlation structure implied by the model which we detail below. To improve the efficiency of this estimator, we then impose additional over-identifying restrictions based on the time-series independence of the shocks, w_t,

$$\begin{aligned} E(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1}) z_{1t-2}^\top &= 0 \\ E(z_{1t} - \tilde{R}_0 - \tilde{A} z_{1t-1}) z_{1t-3}^\top &= 0, \end{aligned} \tag{23}$$

which adds 32 more moment restrictions. To avoid the numerical issues associated with solving a system of 77 nonlinear equations in 45 parameters, we adopt a one-step-efficient GMM estimator as in Newey (1985) detailed in the next section.

D.2 GMM estimation

Stack the p true values of the parameters of the model in the vector θ_0, and write the N moment restrictions in equations (22) and (23) as

$$\eta_T(\theta_0) = \frac{1}{T} \sum_{t=1}^{T} \eta(z_t, \theta_0).$$

The structure of the model implies a central limit theorem so that as $T \to \infty$,

$$V_T^{-1/2} T^{1/2} \eta_T(\theta_0) \sim^a N(0, I),$$

where $\sim_a$ denotes the asymptotic sampling distribution and

$$V_T = \text{Var}\big(T^{1/2} \eta_T(\theta_0)\big) = \frac{1}{T} \text{Var}\Big(\sum_{t=1}^{T} \eta(z_t, \theta_0)\Big),$$

which simplifies for the *iid* case, $V_T = V$.

For the over-identified case, $N > p$, the GMM estimator, $\hat{\theta}_T$, is the minimum of a quadratic form of the moment restrictions and a $N \times N$ positive-definite matrix, Ω_T,

$$\eta_T(\theta)^\top \Omega_T \eta_T(\theta),$$

which implies

$$\eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T \eta_T(\hat{\theta}_T) = 0,$$

where $\eta_{\theta T}(\theta)$ is the matrix derivative of η_T with respect to θ. Note that a linear combination of the mean-value expansion of $\eta_T(\hat{\theta}_T)$ around θ_0 is equal to zero,

$$0 = \eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T \eta_T(\hat{\theta}_T) \quad = \eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T \eta_T(\theta_0) + \eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T \eta_{\theta T}(\bar{\theta}_T)(\hat{\theta}_T - \theta_0),$$

where $\bar{\theta}_T$ is between $\hat{\theta}_T$ and θ_0. This implies

$$T^{1/2}(\hat{\theta}_T - \theta_0) = -[\eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T \eta_{\theta T}(\bar{\theta}_T)]^{-1} \eta_{\theta T}(\hat{\theta}_T)^\top \Omega_T V_T^{1/2} V_T^{-1/2} T^{1/2} \eta_T(\theta_0).$$

As before, combine these results to show that

$$T^{1/2}(\hat{\theta}_T - \theta_0) \sim_a N\left(0, [\eta_\theta(\theta_0)^\top \Omega \eta_\theta(\theta_0)]^{-1} \eta_\theta(\theta_0)^\top \Omega V \Omega \eta_\theta(\theta_0) [\eta_\theta(\theta_0)^\top \Omega \eta_\theta(\theta_0)]^{-1}\right),$$

where Ω is the limit of Ω_T. The estimator with the smallest asymptotic covariance matrix, $\hat{\theta}_T^*$, uses a consistent estimator, V_T, of that matrix V, as the weighting matrix, *i.e.*, $\Omega_T = V_T$, which implies

$$T^{1/2}(\hat{\theta}_T^* - \theta_0) \sim_a N\left(0, [\eta_\theta(\theta_0)^\top V \eta_\theta(\theta_0)]^{-1}\right).$$

Newey (1985, Lemma 4) shows that if the minimization of $\eta_T(\theta)^\top V_T^{-1} \eta_T(\theta)$ begins at a consistent estimator, say $\tilde{\theta}_T$, then the first step of a Gauss-Newton algorithm has the same asymptotic distribution as the minimizer. That is, the estimator

$$\tilde{\theta}_T^* = \tilde{\theta}_T - \alpha_T [\eta_{\theta T}(\tilde{\theta}_T)^\top V_T \eta_{\theta T}(\tilde{\theta}_T)]^{-1} \eta_{\theta T}(\tilde{\theta}_T)^\top V_T \eta_T(\tilde{\theta}_T),$$

is asymptotically equivalent to $\hat{\theta}_T^*$, so that the asymptotic sampling distribution of $\tilde{\theta}_T^*$ is

$$T^{1/2}(\tilde{\theta}_T^* - \theta_0) \sim_a N\left(0, [\eta_\theta(\theta_0)^\top V \eta_\theta(\theta_0)]^{-1}\right),$$

for an appropriate choice of step size, α_T, that converges to 1 as $T \to \infty$.

Our estimation starts with a just-identified estimator, $\tilde{\theta}_T$, that is the solution to $\eta_T(\tilde{\theta}_T) = 0$ where $\eta_T(\theta)$ is defined in equations (22). We also use $\tilde{\theta}_T$ to estimate the matrix V. Values of the estimates at this round are tabulated at the end of this Appendix.

We have parameters γ_0 that are continuous and differentiable functions of the other parameters of the model, $\gamma_0 = f(\theta_0)$, where f_θ has full rank. The continuous mapping theorem implies that $\hat{\gamma}_T = f(\hat{\theta}_T)$ will be a consistent estimator when $\hat{\theta}_T$ is a consistent estimator. The asymptotic distribution of $\hat{\gamma}_T$ follows from a mean-value expansion of $f(\hat{\theta}_T)$ around θ_0,

$$f(\hat{\theta}_T) = f(\theta_0) + f_\theta(\bar{\theta}_T)(\hat{\theta}_T - \theta_0),$$

which implies

$$f_\theta(\bar{\theta}_T)^{-1}(\hat{\gamma}_T - \gamma_0) = \hat{\theta}_T - \theta_0.$$

Given the asymptotic sampling distribution for $\hat{\theta}_T$, we have

$$T^{1/2}(\hat{\gamma}_T - \gamma_0) \sim_a N\left(0, f_\theta(\theta_0)[\eta_\theta(\theta_0)^\top V^{-1}\eta_\theta(\theta_0)]^{-1} f_\theta(\theta_0)^\top\right).$$

D.3 Covariance-matrix estimation

It is convenient to partition the moment restrictions in equations (22) and (23) into 7 blocks with their dimensions listed on the right:

$$\begin{aligned}
\eta_{1t} &= z_{1t} - \tilde{R}_0 - \tilde{A}z_{1t-1} = \tilde{B}w_t && n \times 1\\
\eta_{2t} &= z_{1t-1} \otimes \eta_{1t} = z_{1t-1} \otimes \tilde{B}w_t && n^2 \times 1\\
\eta_{3t} &= z_{1t-2} \otimes \eta_{1t} = z_{1t-2} \otimes \tilde{B}w_t && n^2 \times 1\\
\eta_{4t} &= z_{1t-3} \otimes \eta_{1t} = z_{1t-3} \otimes \tilde{B}w_t && n^2 \times 1\\
\eta_{5t} &= z_{2t} - \tilde{G}_0 - \tilde{G}z_{1t} = u_t && (m+1) \times 1\\
\eta_{6t} &= z_{1t} \otimes \eta_{5t} = z_{1t} \otimes u_t && (m+1)n \times 1\\
\eta_{7t} &= \mathcal{L}\text{vec}\left(\eta_{1t}\eta_{1t}^\top - \tilde{B}\tilde{B}^\top\right) = \mathcal{L}\text{vec}\left(\tilde{B}w_t w_t^\top \tilde{B}^\top - \tilde{B}\tilde{B}^\top\right) && n(n+1)/2 \times 1,
\end{aligned}$$

where $n = 4$, $m = 2$, and $\mathcal{L}$ is an *elimination matrix* of zeros and ones that selects the $n(n+1)/2$ unique elements of the vectorized symmetric covariance matrix.

We need to solve for the $N \times N$ covariance matrix, $\text{Var}(\frac{1}{T}\sum_{t=1}^{T}\eta_t)$, where T is the sample size and $E[\eta_t] = 0$ are the $N \times 1$ moment restrictions. We begin by finding the $N \times N$ matrix $V = \text{Var}(\eta_t) = E[\eta_t\eta_t^\top]$. Recall $z_{1t} = R_0 + Rx_t$, so $Ez_{1t} = R_0$ and $Ez_{1t}z_{1t}^\top = R_0R_0^\top + RV_xR^\top$, $Ez_{1t}z_{1t-1}^\top = R_0R_0^\top + RAV_xR^\top$, and $Ez_{1t}z_{1t-2}^\top = R_0R_0^\top + RA^2V_xR^\top$, where V_x solves $V_x = AV_xA^\top + BB^\top$. Apply the restrictions of the model to the variances and covariances of the 7 blocks of η_t defined above:

$$\begin{aligned}
\mathrm{Var}(\eta_{1t}) &= E[\eta_{1t}\eta_{1t}^\top] = E[\tilde{B}w_t w_t^\top \tilde{B}^\top] = \tilde{B}\tilde{B}^\top \\
\mathrm{Cov}(\eta_{2t}, \eta_{1t}) &= E[z_{1t-1} \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] = E[z_{1t-1}] \otimes \tilde{B}\tilde{B}^\top = R_0 \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{1t}, \eta_{2t})^\top \\
\mathrm{Cov}(\eta_{3t}, \eta_{1t}) &= E[\eta_{3t}\eta_{1t}^\top] = E[(z_{1t-2} \otimes \tilde{B}w_t)w_t^\top \tilde{B}^\top] \\
&= E[z_{1t-2} \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] = E[z_{1t-2}] \otimes \tilde{B}\tilde{B}^\top = R_0 \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{1t}, \eta_{3t})^\top \\
\mathrm{Cov}(\eta_{4t}, \eta_{1t}) &= E[\eta_{4t}\eta_{1t}^\top] = E[(z_{1t-3} \otimes \tilde{B}w_t)w_t^\top \tilde{B}^\top] \\
&= E[z_{1t-3} \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] = E[z_{1t-3}] \otimes \tilde{B}\tilde{B}^\top = R_0 \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{1t}, \eta_{4t})^\top \\
\mathrm{Cov}(\eta_{5t}, \eta_{1t}) &= E[\eta_{5t}\eta_{1t}^\top] = E[u_t w_t^\top \tilde{B}^\top] = 0 \\
&= \mathrm{Cov}(\eta_{1t}, \eta_{5t})^\top \\
\mathrm{Cov}(\eta_{6t}, \eta_{1t}) &= E[\eta_{6t}\eta_{1t}^\top] = E[(z_{1t} \otimes u_t)w_t^\top \tilde{B}^\top] = E[z_{1t} \otimes u_t w_t^\top \tilde{B}^\top] = 0 \\
&= \mathrm{Cov}(\eta_{1t}, \eta_{6t})^\top \\
\mathrm{Var}(\eta_{2t}) &= E[\eta_{2t}\eta_{2t}^\top] = E[(z_{1t-1} \otimes \tilde{B}w_t)(z_{1t-1}^\top \otimes w_t^\top \tilde{B}^\top)] \\
&= E[z_{1t-1}z_{1t-1}^\top \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] \\
&= [R_0 R_0^\top + RV_x R^\top] \otimes \tilde{B}\tilde{B}^\top
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Cov}(\eta_{2t},\eta_{3t}) &= E[\eta_{2t}\eta_{3t}^\top] = E[(z_{1t-1}\otimes \tilde{B}w_t)(z_{1t-2}^\top \otimes w_t^\top \tilde{B}^\top)] \\
&= E[z_{1t-1}z_{1t-2}^\top \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] \\
&= [R_0R_0^\top + RAV_xR^\top] \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{3t},\eta_{2t})^\top \\
\mathrm{Cov}(\eta_{2t},\eta_{4t}) &= E[\eta_{2t}\eta_{4t}^\top] = E[(z_{1t-1}\otimes \tilde{B}w_t)(z_{1t-3}^\top \otimes w_t^\top \tilde{B}^\top)] \\
&= E[z_{1t-1}z_{1t-3}^\top \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] \\
&= [R_0R_0^\top + RA^2V_xR^\top] \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{4t},\eta_{2t})^\top \\
\mathrm{Cov}(\eta_{2t},\eta_{5t}) &= E[\eta_{2t}\eta_{5t}^\top] = E[(z_{1t-1}\otimes \tilde{B}w_t)u_t^\top] = 0 \\
&= \mathrm{Cov}(\eta_{5t},\eta_{2t})^\top \\
\mathrm{Cov}(\eta_{2t},\eta_{6t}) &= E[\eta_{2t}\eta_{6t}^\top] = E[(z_{1t-1}\otimes \tilde{B}w_t)(z_{1t}^\top \otimes u_t^\top)] \\
&= E[z_{1t-1}z_{1t}^\top \otimes \tilde{B}w_t u_t^\top] = 0 \\
&= \mathrm{Cov}(\eta_{6t},\eta_{2t})^\top \\
\mathrm{Var}(\eta_{3t}) &= E[\eta_{3t}\eta_{3t}^\top] = E[(z_{1t-2}\otimes \tilde{B}w_t)(z_{1t-2}^\top \otimes w_t^\top \tilde{B}^\top)] \\
&= E[z_{1t-2}z_{1t-2}^\top \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] \\
&= [R_0R_0^\top + RV_xR^\top] \otimes \tilde{B}\tilde{B}^\top \\
\mathrm{Cov}(\eta_{3t},\eta_{4t}) &= E[\eta_{3t}\eta_{4t}^\top] = E[(z_{1t-2}\otimes \tilde{B}w_t)(z_{1t-3}\otimes \tilde{B}w_t)^\top] \\
&= E[z_{1t-2}z_{1t-3}^\top \otimes \tilde{B}w_t w_t^\top \tilde{B}^\top] \\
&= [R_0R_0^\top + RAV_xR^\top] \otimes \tilde{B}\tilde{B}^\top \\
&= \mathrm{Cov}(\eta_{4t},\eta_{3t})^\top
\end{aligned}
$$

$$\begin{aligned}
\mathrm{Cov}(\eta_{3t},\eta_{5t}) &= E[\eta_{3t}\eta_{5t}^\top] = E[(z_{1t-2}\otimes \tilde{B}w_t)u_t^\top] = 0\\
&= \mathrm{Cov}(\eta_{5t},\eta_{3t})^\top\\
\mathrm{Cov}(\eta_{3t},\eta_{6t}) &= E[\eta_{3t}\eta_{6t}^\top] = E[(z_{1t-2}\otimes \tilde{B}w_t)(z_{1t}^\top\otimes u_t^\top)]\\
&= E[z_{1t-2}z_{1t}^\top\otimes \tilde{B}w_t u_t^\top] = 0\\
&= \mathrm{Cov}(\eta_{6t},\eta_{3t})^\top\\
\mathrm{Var}(\eta_{4t}) &= E[\eta_{4t}\eta_{4t}^\top] = E[(z_{1t-3}\otimes \tilde{B}w_t)(z_{1t-3}^\top\otimes w_t^\top\tilde{B}^\top)]\\
&= E[z_{1t-3}z_{1t-3}^\top\otimes \tilde{B}w_t w_t^\top\tilde{B}^\top]\\
&= [R_0R_0^\top + RV_xR^\top]\otimes \tilde{B}\tilde{B}^\top\\
\mathrm{Cov}(\eta_{4t},\eta_{5t}) &= E[\eta_{4t}\eta_{5t}^\top] = E[(z_{1t-3}\otimes \tilde{B}w_t)u_t^\top] = 0\\
&= \mathrm{Cov}(\eta_{5t},\eta_{4t})^\top\\
\mathrm{Cov}(\eta_{4t},\eta_{6t}) &= E[\eta_{4t}\eta_{6t}^\top] = E[(z_{1t-3}\otimes \tilde{B}w_t)(z_{1t}^\top\otimes u_t^\top)]\\
&= E[z_{1t-3}z_{1t}^\top\otimes \tilde{B}w_t u_t^\top] = 0\\
&= \mathrm{Cov}(\eta_{6t},\eta_{4t})^\top\\
\mathrm{Var}(\eta_{5t}) &= E[\eta_{5t}\eta_{5t}^\top] = E[u_tu_t^\top] = \sigma_u^2 I_{(3\times 3)}\\
\mathrm{Cov}(\eta_{6t},\eta_{5t}) &= E[\eta_{6t}\eta_{5t}^\top] = E[(z_{1t}\otimes u_t)u_t^\top]\\
&= E[z_{1t}\otimes u_tu_t^\top] = E[z_{1t}]\otimes\sigma_u^2 I_{(3\times 3)} = R_0\otimes\sigma_u^2 I_{(3\times 3)}\\
&= \mathrm{Cov}(\eta_{5t},\eta_{6t})^\top
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(\eta_{6t}) &= E[\eta_{6t}\eta_{6t}^\top] = E[(z_{1t}\otimes u_t)(z_{1t}^\top\otimes u_t^\top)]\\
&= E[z_{1t}z_{1t}^\top\otimes u_tu_t^\top] = [R_0R_0^\top + RV_xR^\top]\otimes\sigma_u^2 I_{(3\times 3)}.
\end{aligned}$$

For the remaining variances and covariances involving η_{7t}, we need to expand the $\mathrm{vec}(\cdot)$ operator:

$$\begin{aligned}
\mathrm{vec}\big(\eta_{1t}\eta_{1t}^\top - \tilde{B}\tilde{B}^\top\big) &= \mathrm{vec}\big(\tilde{B}w_tw_t^\top\tilde{B}^\top - \tilde{B}\tilde{B}^\top\big)\\
&= \mathrm{vec}\big(\tilde{B}w_tw_t^\top\tilde{B}^\top\big) - \mathrm{vec}\big(\tilde{B}\tilde{B}^\top\big)\\
&= (\tilde{B}\otimes\tilde{B})\mathrm{vec}(w_tw_t^\top) - \mathrm{vec}\big(\tilde{B}\tilde{B}^\top\big)\\
&= CW_t - D,
\end{aligned}$$

where $C = \tilde{B}\otimes\tilde{B}$, $D = \mathrm{vec}\big(\tilde{B}\tilde{B}^\top\big)$, and $W_t = \mathrm{vec}(w_tw_t^\top)$. We can rewrite $\eta_{7t} = LCW_t - LD$, so that

$$\begin{aligned}
\mathrm{Var}(\eta_{7t}) &= E[\eta_{7t}\eta_{7t}^\top] = E\big[(\mathcal{L}CW_t - \mathcal{L}D)(W_t^\top C^\top\mathcal{L}^\top - D^\top\mathcal{L}^\top)\big]\\
&= \mathcal{L}CE[W_tW_t^\top]C^\top\mathcal{L}^\top + \mathcal{L}DD^\top\mathcal{L}^\top - \mathcal{L}DE[W_t^\top]C^\top\mathcal{L}^\top - \mathcal{L}CE[W_t]D^\top\mathcal{L}^\top,
\end{aligned}$$

which requires us to first calculate the first two moments of the vector W_t: $\bar{W} = E[W_t]$ and $H = E[W_t W_t^\top]$. We can simplify these expectations using independence and properties of the standard normal, $Ew_{jt} = 0$, $Ew_{jt}^2 = 1$, $Ew_{jt}^3 = 0$, and $Ew_{jt}^4 = 3$, to get exact expressions for $\bar{W}$ and H – appropriate arrangements of the values 0, 1, and 3 – which we can then use to calculate the variance of η_{7t} as

$$\begin{aligned}\mathrm{Var}(\eta_{7t}) &= \mathcal{L}(\tilde{B} \otimes \tilde{B})H(\tilde{B} \otimes \tilde{B})^\top \mathcal{L}^\top + \mathcal{L}\mathrm{vec}(\tilde{B}\tilde{B}^\top)\mathrm{vec}(\tilde{B}\tilde{B}^\top)^\top \mathcal{L}^\top \\ &\quad - \mathcal{L}\mathrm{vec}(\tilde{B}\tilde{B}^\top)\bar{W}^\top(\tilde{B} \otimes \tilde{B})^\top \mathcal{L}^\top - \mathcal{L}(\tilde{B} \otimes \tilde{B})\bar{W}\mathrm{vec}(\tilde{B}\tilde{B}^\top)^\top \mathcal{L}^\top.\end{aligned}$$

Note that since $E[W_t w_t^\top] = 0$, and since w_t and u_t are independent, there are no other non-zero covariances:

$$\mathrm{Cov}(\eta_{7t}, \eta_{it}) = \mathrm{Cov}(\eta_{it}, \eta_{7t})^\top = 0,$$

for $1 = 1, 2, \ldots, 6$.

Finally, given the independence of w_t and u_t across time, there are no nonzero autocovariances, $E[\eta_t \eta_{t-j}^\top] = 0$ for $j > 0$. Therefore, we have

$$\mathrm{Var}\big(\frac{1}{T}\sum_{t=1}^{T}\eta_t\big) = \frac{1}{T^2}\sum_{t=1}^{T}\mathrm{Var}\big(\eta_t\big) = \frac{1}{T}V.$$

The covariance matrix of the moment equations for the just-identified model is calculated in a similar fashion.

E Additional empirical results

Just-identified GMM estimation (4-Factor Model)

I. State dynamics

A				B			
0.8373	0.4066	0.3323	0.8918	0.0031	0	0	0
(0.0936)	(0.0820)	(0.0847)	(0.2484)	(0.0004)			
0.5795	0.4529	-0.0231	-1.1433	-0.0059	0.0094	0	0
(0.8142)	(0.2622)	(0.2658)	(2.0181)	(0.0094)	(0.0169)		
-0.4688	0.1350	0.6377	0.6642	0.0033	-0.0092	0.0021	0
(0.8023)	(0.2569)	(0.3328)	(1.9049)	(0.0096)	(0.0173)	(0.0003)	
-0.0029	-0.0716	-0.0532	0.3696	-0.0002	0.0007	-0.0010	0.0014
(0.0287)	(0.0558)	(0.0910)	(0.1058)	(0.0004)	(0.0003)	(0.0006)	(0.0001)

II. Term-structure model

A^*				λ_0	b	c
0.9961	0	0	0	-0.0759	0.4945	0.0171
(0.0011)				(0.0561)	(0.0060)	(0.0101)
0	0.8817	0	0	-0.0865	0.4244	-0.1016
	(0.0337)			(0.2693)	(0.0337)	(0.1218)
0	0	0.8338	0	-0.0477	0.4758	-0.2592
		(0.0546)		(0.2726)	(0.0614)	(0.1578)
0	0	0	0.3634	-0.5820	0.5939	0.0316
			(0.1062)	(0.2796)	(0.0196)	(0.1696)

III. Taylor rule

τ_π	τ_g	$d^\top$			
2.1219	0.4570	-0.0571	0.1460	0.1089	-0.2747
(0.5947)	(0.3107)	(0.2926)	(0.2763)	(0.3684)	(0.3447)

Note. Based on moment restrictions in equations (22) for a sample period 1980Q3 to 2019Q4. Asymptotic standard errors are in parentheses. State dynamics: $x_{t+1} = Ax_t + Bw_{t+1}$. Macro term-structure model: $i_t = a_0 + a^\top x_t$, $\pi_t = b_0 + b^\top x_t$, $g_t = c_0 + c^\top x_t$, $hy_t^{(h)} = \mathcal{B}_0^{(h)} + \mathcal{B}^{(h)} x_t$, and $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1}(I - A^{*h})$, $\mathcal{B}_0^{(h)} = a_0 + \mathcal{B}_0^{(h-1)} - \mathcal{B}^{(h-1)} B\lambda_0 - \mathcal{B}^{(h-1)} BB^\top \mathcal{B}^{(h-1)\top}/2$. Policy: $i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d^\top x_t$. 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 $h = 4$, 12, 20, 40 (quarters), π_t is the inflation rate, g_t is the growth rate of real GDP, and $a^\top = [1\ 1\ 1\ 1]$. Values for intercepts are fixed at their sample means. The eigenvalues of A are 0.9729, 0.6258, 0.5820, and 0.1167.

Just-Identified GMM Estimation (3-Factor Model)

I. Term-Structure Model

A^*			λ_0	b	c
0.9899	0	0	-0.0962	0.4491	0.1240
(0.0010)			(0.0266)	(0.0215)	(0.0532)
0	0.8941	0	-0.1103	0.4068	-0.0384
	(0.0097)		(0.0852)	(0.0488)	(0.1211)
0	0	0.4380	-0.5066	0.5876	-0.0410
		(0.0955)	(0.2113)	(0.1207)	(0.2880)

A			B		
1.0123	0.1221	0.2914	0.0021	0	0
(0.0183)	(0.0402)	(0.1171)	(0.0002)		
-0.0032	0.8862	0.8862	-0.0016	0.0021	0
(0.0275)	(0.0555)	(0.1628)	(0.0002)	(0.0002)	
-0.0627	-0.0942	0.3894	0.0002	-0.0011	0.0014
(0.0180)	(0.0004)	(0.0824)	(0.0002)	(0.0022)	(0.0001)

II. Taylor Rule

τ_0	τ_g	τ_π	$d^\top$		
-0.0076	0.5510	1.8932	0.0815	0.2510	-0.0899
(0.0078)	(1.3558)	(0.3759)	(0.1548)	(0.1398)	(0.0934)

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 $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1}(I - A^{*h})$, $\mathcal{B}_0^{(h)} = a_0 + \mathcal{B}_0^{(h-1)} - \mathcal{B}^{(h-1)} B\lambda_0 + \mathcal{B}^{(h-1)} BB^\top \mathcal{B}^{(h-1)\top}/2$. Policy: $i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d^\top x_t$. The state variable x_t is 3-dimensional, i_t is the short interest rate (1 quarter), $y_t^{(h)}$ is the yield on a discount bond of maturity h = 4, 12, 20, 40 (quarters), π_t is the inflation rate, g_t is the growth rate of real GDP, and $a^\top = [1\ 1\ 1\ 1]$. Values for a_0, b_0, and c_0 are fixed at their sample means. The sample period is 1980Q3 to 2019Q4. Asymptotic standard errors are in parentheses.

GMM Estimation (3-Factor Model)

I. Term-Structure Model

A^*			λ_0	b	c
0.9899	0	0	-0.0969	0.1241	0.4491
(0.0001)			(0.0024)	(0.0006)	(0.0004)
0	0.8942	0	-0.1028	-0.0383	0.4068
	(0.0006)		(0.0102)	(0.0012)	(0.0008)
0	0	0.4381	-0.4779	-0.0409	0.5876
		(0.0054)	(0.0299)	(0.0032)	(0.0028)

A			B		
1.0166	0.1154	0.2996	0.0021	0	0
(0.0182)	(0.0409)	(0.0929)	(0.0001)		
0.0054	0.9004	0.0792	-0.0015	0.0023	0
(0.0202)	(0.0521)	(0.1145)	(0.0002)	(0.0001)	
-0.0484	-0.0910	0.4389	0.0002	-0.0010	0.0015
(0.0158)	(0.0350)	(0.0791)	(0.0001)	(0.0001)	(0.0001)

II. Taylor Rule

τ_0	τ_g	τ_π	$d^\top$		
-0.0078	0.5827	1.8918	0.0782	0.2528	-0.0877
(0.0064)	(0.8765)	(0.0723)	(0.1406)	(0.0072)	(0.0102)

Note: Over-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 $\mathcal{B}^{(h)} = a^\top (I - A^*)^{-1}(I - A^{*h})$, $\mathcal{B}_0^{(h)} = a_0 + \mathcal{B}_0^{(h-1)} - \mathcal{B}^{(h-1)} B\lambda_0 + \mathcal{B}^{(h-1)} BB^\top \mathcal{B}^{(h-1)\top}/2$. Policy: $i_t = \tau_0 + \tau_\pi \pi_t + \tau_g g_t + d^\top x_t$. The state variable x_t is 3-dimensional, i_t is the short interest rate (1 quarter), $y_t^{(h)}$ is the yield on a discount bond of maturity h = 4, 20, 40 (quarters), π_t is the inflation rate, g_t is the growth rate of real GDP, and $a^\top = [1\ 1\ 1\ 1]$. Values for a_0, b_0, and c_0 are fixed at their sample means. The eigenvalues of A are 0.9892, 0.8861, and 0.4806. The sample period is 1980Q3 to 2019Q4. Asymptotic standard errors are in parentheses.