

Bond pricing and the macroeconomy

Gregory R. Duffee
Johns Hopkins University

This version July 2012

Abstract

This chapter reviews some of the academic literature that links nominal and real term structures with the macroeconomy. The main conclusion is that none of our models is consistent with basic properties of nominal yields. It is difficult to explain the average shape of the nominal yield curve, the variation of yields over time, and the predictability of excess bond returns. There are two overarching problems. First, much of the variation over time in economic activity is orthogonal to variation in nominal yields, and vice versa. Second, although mean excess returns to nominal Treasury bonds are positive, these returns do not appear to positively covary with risks that require compensation, at least according to standard asset-pricing models.

JEL Code: G12

Key words: Term structure, affine models, macro-finance, no-arbitrage

This is an update of a chapter prepared for Elsevier's *Handbook of the Economics of Finance*, edited by George Constantinides, Milt Harris, and Rene Stulz, forthcoming 2012. Relative to the handbook version, this version fixes some typos, corrects some table entries, and slightly alters some text for consistency with the corrected table entries. Some of the material in this chapter also appears in "Forecasting interest rates," a chapter prepared for Elsevier's *Handbook of Economic Forecasting*, Vol. 2., edited by Allan Timmermann and Graham Elliott, forthcoming 2012. Thanks to Michael Bauer, Jon Faust, Monika Piazzesi, Glenn Rudebusch, Jonathan Wright, and many seminar participants for helpful comments and conversations. Voice 410-516-8828, email duffee@jhu.edu. Address correspondence to 440 Mergenthaler Hall, 3400 N. Charles St., Baltimore, MD 21218.

1 Introduction

To many macroeconomists, the discussion of fixed income in the handbook chapter of Dai and Singleton (2003) appears to be written in some sort of dolphin language. Who are Feynman and Kac, and what magic is performed with an infinitesimal generator? What is Itô's Lemma, and why is it important in determining the price of a bond? More importantly, why is there no mention of inflation, monetary policy, the business cycle, supply and demand for bonds, or the Taylor (1993) rule? What is the point of describing the behavior of interest rates with a latent state vector for which there is no macroeconomic interpretation?

Since the appearance of Dai and Singleton's review, a flood of research has attempted to link bond prices to the macroeconomy. Almost all of this work steps back from the mathematical frontiers of term structure modeling. This retreat from the dolphin tongue helps link the intuition of finance to that of macroeconomics. Discussions of generalized Bessel processes are replaced with discussions of long-run inflation targets.

This chapter selectively and critically reviews this recent research. Its main conclusion is that much of the progress has been, in a sense, negative. We have a much better understanding of how difficult it is to reconcile the behavior of the term structure with our workhorse models of the macroeconomy and investor preferences. In particular, the shape of the nominal term structure varies substantially through time, yet this variation is depressingly difficult to align with variations in macroeconomic activity, inflation expectations, or plausible measures of risk premia. At a more basic level, our standard models are hard to reconcile with simple moments, such as a nominal yield curve that slopes up on average.

The next section lays out a general dynamic factor model with which we can study the joint behavior of bond yields and other macroeconomic variables. Section 3 follows the recent bond-pricing literature by imposing no-arbitrage restrictions on the framework. At this level of generality, no economically-motivated restrictions on risk premia are imposed. Section 4 uses U.S. data to describe the empirical links between nominal yields and macroeconomic activity. It discusses important tensions between properties of the data and properties implied

by a benchmark dynamic factor model.

Section 5 surveys consumption-based models of bond risk premia and compares their testable implications with U.S. data. Although many of the models are designed expressly to reproduce basic empirical properties of excess returns to nominal bonds, none of them can be viewed as a success. In each tested case, the mechanisms that generate the observed properties lack support in the data. Section 6 considers the restrictions that New Keynesian models impose on the joint dynamics of the term structure and the macroeconomy. Again, the evidence is less than encouraging. The final section concludes.

2 A factor model

What is a reasonable framework to use in estimating the joint dynamics of macroeconomic activity and inflation? An obvious modeling approach is to fit a vector autoregression (VAR) to the macroeconomic variables of interest. But both theory and empirical evidence tell us that the nominal term structure contains additional information about expected future inflation and real activity. Rational investors impound their information about future inflation and real rates into yields on nominal bonds. Standard models that incorporate a Taylor rule imply that both of these components of nominal yields are plausibly correlated with future economic activity. This theory is consistent with empirical evidence that short-term interest rates and the slope of the Treasury yield curve predict economic growth.¹

One way to incorporate information from the term structure is to add bond yields to the VAR. The macro-finance term structure literature takes a slightly different approach.

2.1 A bare-bones framework

Dynamics of short-term interest rates, bond yields, and other macroeconomic variables such as economic activity and inflation are determined by a p -vector of factors. Denote the vector

¹Early references are Estrella and Hardouvelis (1991) and Bernanke and Blinder (1992).

by x_t . Begin by assuming that it has Gaussian linear dynamics given by

$$x_{t+1} = \mu + Kx_t + \Sigma\epsilon_{t+1}, \quad \epsilon_{t+1} \sim MVN(0, I). \quad (1)$$

In the language of the Kalman filter, (1) is a transition equation. A measurement equation links observables to the state vector. Stack time- t observations in a vector \tilde{z}_t . This vector contains bond yields as well as other data. The measurement equation is

$$\tilde{z}_t = \mathbf{A} + \mathbf{B}x_t + \eta_t, \quad \eta_t \sim MVN(0, \Omega). \quad (2)$$

Deviations from an exact factor model are captured by the η_t vector, which is unpredictable prior to t .

The usual macro-finance interpretation of the measurement equation is that the state vector captures the common variation in the observables, while the η_t vector captures idiosyncratic deviations that we can think of as cross-sectional errors. This description is borrowed from the latent-factor, no-arbitrage term structure literature, in which p factors determine prices of all bonds. Deviations from factor-model pricing are explained by imperfections such as measurement error and bid-ask bounce.

However, this decomposition into common components and cross-sectional error components is not required by (1) and (2). According to the equations, the state vector simply captures the persistent components of the observables and η_t captures white-noise components, which have no persistence. In fact, these equations are compatible with observables that are mutually independent. An example highlights how the measurement equation can be interpreted in a macro-finance setting, foreshadowing some of the empirical analysis in Section 4.2.

In this example, the observables are nominal bond yields, inflation, and real consumption growth. Assume that real consumption growth is the sum of a slow-moving conditional mean and a white-noise shock. Then the state vector captures the conditional mean of

real consumption growth, while the white-noise shock shows up in η_t . If inflation can also be described as the sum of a conditional mean and a white-noise shock, then the state vector also includes the determinants of its conditional mean. Nothing in this setup requires that the factors that determine the conditional mean of real consumption growth are the same factors that determine the conditional mean of inflation. In addition, nothing requires that the white-noise shock to consumption growth is orthogonal to the white-noise shock to inflation.

In typical consumption-based utility models, nominal yields depend on the union of the factors driving conditional expectations of consumption growth and inflation. They do not depend on the components of consumption growth and inflation that appear in η_t . The state vector will also include any factors that drive yields which are independent of consumption growth and inflation. Thus it makes sense to describe the state vector as the source of common shocks to bond yields, while η_t picks up idiosyncratic components of yields such as measurement error. More generally, we can label the cross-sectional deviations for bond yields as “noise.” For this example, the label is an inappropriate description of the cross-sectional deviations of the macro variables. Section 4.4 discusses practical modeling issues that hinge on the proper interpretation of cross-sectional deviations.

We return to the distinction between common factors and persistent factors in the empirical analysis of Section 4.2. The final contribution of this discussion is to introduce notation and terminology for the observed data less the white-noise shocks η_t . Observed data are denoted with tildes, while the non- η_t components are denoted without them. Using this notation we can rewrite (2) as

$$\tilde{z}_t = z_t + \eta_t, \tag{3}$$

$$z_t \equiv \mathbf{A} + \mathbf{B}x_t. \tag{4}$$

To avoid putting perhaps unwarranted structure on the state vector, I refer to z_t as the persistent component of the observables rather than the common component. For example,

we can distinguish between observed inflation $\tilde{\pi}_t$ and the persistent component of inflation implied by the factor model, denoted π_t .

2.2 Implications and alternatives

There are three general restrictions embedded in this model. The first is that p variables determine the conditional expectations of the observables. The second is the linear Markov structure of the state dynamics, and the third is the Gaussian description of the innovations.

It is worth discussing why a setting with obviously counterfactual assumptions plays a central role in the literature. It is easy to find evidence of nonlinear, non-Gaussian dynamics in bond yields. For example, Gray (1996) concludes that a model of time-varying mean reversion and time-varying GARCH effects fits the dynamics of the short-term interest rate. Unfortunately, tractability must sometimes trump truth. Section 3 imposes no-arbitrage on this framework to produce a dynamic term structure model. Dynamic term structure models require (a) a method for computing bond prices that satisfy no-arbitrage; and (b) transition densities of prices from t to $t + \tau$. Researchers naturally restrict their attention to models for which (a) and (b) are computationally feasible.

The best-known class of models with tractable bond pricing is the affine class of Duffie and Kan (1996). This class includes both homoskedastic (Gaussian) and heteroskedastic models. Dai and Singleton (2000) and Duffee (2002) combine this affine class with linear dynamics of the underlying state vector to produce the “completely affine” and “essentially affine” classes respectively. One of the conclusions in Duffee (2002) is that only the Gaussian models in this class are sufficiently flexible to generate plausible forecasts of future yields.

Recent research has attempted to find alternatives to the Gaussian class. Cheridito, Filipović, and Kimmel (2007) extend the essentially affine class to give non-Gaussian versions greater flexibility. A nonlinear tweak to the completely affine class is introduced by Duarte (2004). The quadratic class, which has nonlinear dynamics, is developed by Leippold and Wu (2002) and Ahn, Dittmar, and Gallant (2002). Dai, Singleton, and Yang (2007) and Ang,

Bekaert, and Wei (2008) construct models with regime switches along certain dimensions. A fairly general nonlinear framework with affine pricing is developed by Le, Singleton, and Dai (2010). All of these approaches show promise, but none has gained traction in the applied literature. This literature generally focuses on predicting yields and macroeconomic variables rather than constructing conditional second moments. Linear Gaussian no-arbitrage models are easy to understand and use, and can generate complicated yield dynamics.

Within the linear Gaussian class, there is some flexibility to deviate from the framework used here. For example, we could use a first-order autoregression. A VAR(1) differs in two ways from the factor model of (1) and (2). The first difference is that with the factor model, the length of the common vector need not equal the number of observed variables. In most term structure applications the latter exceeds the former. In this typical case, the factor model is a more parsimonious description of dynamics than a VAR(1). The factor model allows this inequality to be reversed. An example with more factors than observables appears in Section 6, but almost all models in this chapter fall into the typical case. The second difference is the presence of the cross-sectional errors, which affect the autocovariances of the observed data. First-order vector autoregressions generate smooth autocovariance functions. With the factor model, covariances drop quickly from lag zero to lag one, then die out more slowly.

Additional modifications are possible that generate more complicated autocovariance functions. For example, Piazzesi and Schneider (2007) link the state-vector shocks in (1) to the cross-sectional shocks in (2). Ang and Piazzesi (2003) add additional lags of the state vector to the dynamics of (1), much like using a higher-order vector autoregression.

2.3 What are the factors?

What are the fundamental determinants of the dynamic behavior of the term structure? Although this is an important question, it is vacuous in the context of a simple factor model. The question is effectively asking how we should interpret elements of the state vector. But

the state vector is arbitrary. An observationally equivalent model is produced by scaling, rotating, and translating the state vector. Associated with each rotation is a different set of parameters of the transition equation (1) and the measurement equation (2).

Define such a transformation as

$$x_t^* = \underbrace{\Gamma_0}_{p \times 1} + \underbrace{\Gamma_1}_{p \times p} x_t \quad (5)$$

where Γ_1 is nonsingular. An observationally equivalent model replaces x_t with x_t^* , replaces the parameters of (1) with

$$K^* = \Gamma_1 K \Gamma_1^{-1}, \quad \mu^* = \Gamma_1 \mu + (I - K^*) \Gamma_0, \quad \Sigma^* = \Gamma_1 \Sigma,$$

and replaces the parameters of (2) with

$$\mathbf{B}^* = \mathbf{B} \Gamma_1^{-1}, \quad \mathbf{A}^* = \mathbf{A} - \mathbf{B}^* \Gamma_0.$$

Examples of these transformations help illustrate the indeterminacy.

An obvious transformation of the form (5) starts with the persistent components of the observables defined by (4). When the dimension of the observed data is at least as large as the dimension of the state vector, we can choose p elements of the vector in (4) as factors. The only requirement is that the corresponding $p \times p$ submatrix of \mathbf{B} in (4) is nonsingular.

For example, consider a model with a length-two state vector that determines the persistent components of inflation, output growth, and yields on two bonds. As long as the invertibility requirement is satisfied, we can choose the state vector such that the factors are the persistent components of inflation and output growth. This does not mean that inflation and output growth are the fundamental determinants of bond yields, any more than does a rotation of the factors into the bond yields imply that yields are the fundamental determinants of inflation and output growth. The model lacks sufficient structure to allow for such

claims. All the model claims is that two factors drive conditional expectations among the observables.

2.4 Taylor rule stories

Other examples of this indeterminacy use the Taylor rule as a starting point. This discussion is inspired by the models examined in Ang, Dong, and Piazzesi (2007). A baseline Taylor (1993) rule is that the nominal short rate depends on the output gap, inflation, and an unobserved monetary policy component. Denote the persistent components of the nominal short rate, the output gap, and inflation by $r_t^{\$}$, g_t , and π_t respectively. Observed values are denoted with tildes. The Taylor rule is

$$r_t^{\$} = \delta_0 + \delta_{1g}g_t + \delta_{1\pi}\pi_t + \phi_t \quad (6)$$

where the residual ϕ_t is unobserved monetary policy.

This equation can be embedded in a factor model by defining the state vector

$$x_t = \begin{pmatrix} g_t & \pi_t & \phi_t \end{pmatrix}' \quad (7)$$

and specifying the dynamics of the state as (1). An example of a measurement equation is

$$\begin{pmatrix} \tilde{g}_t \\ \tilde{\pi}_t \\ \tilde{r}_t^{\$} \\ \text{other observables}_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \delta_0 \\ A \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \delta_{1g} & \delta_{1\pi} & 1 \\ B_1 & B_2 & B_3 \end{pmatrix} x_t + \eta_t,$$

where the category “other observables” can include any other variable that we believe is determined by the same state vector.

The first example of indeterminacy is that without additional restrictions, the monetary

policy factor is unidentified. An observationally equivalent model is generated with

$$x_t^* = \begin{pmatrix} g_t \\ \pi_t \\ \phi_t^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \Gamma_{11} & \Gamma_{12} & 1 \end{pmatrix} x_t, \quad (8)$$

where Γ_{11} and Γ_{12} are arbitrary. One normalization that pins down these parameters is the requirement that the unconditional covariances of unobserved monetary policy with inflation and output are zero. This is equivalent to treating the Taylor rule equation (6) as a regression equation. An alternative normalization is that future values of the monetary policy variable cannot be predicted with either current inflation or the current output gap, except to the extent that these variables are correlated with the current value of the monetary policy variable. Formally, we can impose the normalization

$$E(\phi_{t+1}^* | x_t^*) = E(\phi_{t+1}^* | \phi_t^*).$$

which defines the parameters of the rotation (8).

Another example of indeterminacy is that the state vector can be transformed to replace the output gap and inflation with their one-step-ahead forecasts. Use the transformation

$$x_t^* = \begin{pmatrix} E_t(g_{t+1}) \\ E_t(\pi_{t+1}) \\ \phi_t \end{pmatrix} = \begin{pmatrix} [\mu + Kx_t]_{(1:2,:)} \\ \phi_t \end{pmatrix},$$

which is an affine transformation of x_t . This transformation allows us to rewrite the Taylor rule in forward-looking form

$$r_t^{\$} = \delta_0 + \delta_{1g} E_t(g_{t+1}) + \delta_{1\pi} E_t(\pi_{t+1}) + \phi_t. \quad (9)$$

This forward-looking rule is observationally equivalent to the baseline rule (6).

Yet another state vector for this model is

$$x_t^* = \left(g_t \quad \pi_t \quad r_t^{\$} \right)' \tag{10}$$

In this version monetary policy does not appear explicitly, but there are no observable implications of a model written in terms of this state vector that differ from those of a model written in terms of the state vector (7).

The main conclusion to draw from this discussion is that the general factor model of Section 2.1 is consistent with many reasonable interpretations of the determinants of macroeconomic activity, interest rates, and inflation. Absent additional restrictions, there is no way to reject one of these interpretations in favor of another. Section 6 discusses some possible restrictions based on New Keynesian models. First, however, we turn to restrictions that tie together yields on bonds with different maturities.

3 No-arbitrage restrictions

In the macro-finance interpretation of Section 2.1’s factor model, shocks to p factors determine common innovations among observed bond yields. Thus there are p common shocks to observed bond prices. The intuition of Black and Scholes suggests that the compensation investors receive for facing of each of these shocks should be the same across all bonds. This no-arbitrage intuition underlies the literature of continuous-time fixed-income pricing that begins with Vasicek (1977) and Cox, Ingersoll, and Ross (1985), both of which are special cases of the affine class described by Duffie and Kan (1996).

Perhaps surprisingly, this no-arbitrage logic does not carry over to the discrete-time framework of Section 2.1. The framework specifies that bond yields, not bond prices, are affine functions of the state vector. The log transformation implies that the price of any one bond is not perfectly conditionally correlated with the value of a portfolio of other bonds.

However, following Backus and Zin (1994), we can circumvent this difficulty by modeling stochastic discount factors as conditionally log-normal processes.

3.1 Stochastic discount factors

By the requirement of no-arbitrage, the real price of an arbitrary financial instrument that does not pay dividends at $t + 1$ satisfies the law of one price

$$P_t = E_t(M_{t+1}P_{t+1}), \quad (11)$$

where M_{t+1} is the strictly positive real stochastic discount factor (SDF).

The real SDF is one of the two building blocks of the nominal SDF. The other is the price level Π_t , which is expressed as dollars per unit of consumption. Continuously-compounded inflation is therefore

$$\tilde{\pi}_t \equiv \log \frac{\Pi_t}{\Pi_{t-1}}. \quad (12)$$

Recall that $\tilde{\pi}_t$ denotes the sum of the persistent and non-persistent components of inflation.

Denote the nominal price of a financial instrument as $P_t^\$$. Multiply (11) by the time- t price level to produce

$$P_t^\$ = E_t(M_{t+1}e^{-\tilde{\pi}_{t+1}}P_{t+1}^\$).$$

Therefore the stochastic discount factor for nominal payoffs is

$$M_{t+1}^\$ = M_{t+1}e^{-\tilde{\pi}_{t+1}}. \quad (13)$$

We begin to add structure to these SDFs by assuming they are conditionally log-normally distributed. This implies inflation is conditionally normally distributed. Denote the logs of the SDFs with lowercase m 's. Then the relationship between the continuously-compounded

one-period real interest rate and the real SDF is

$$r_t = -E_t(m_{t+1}) - \frac{1}{2}\text{Var}_t(m_{t+1}). \quad (14)$$

Similarly, the one-period nominal interest rate is

$$r_t^{\$} = r_t + E_t(\tilde{\pi}_{t+1}) + \text{Cov}_t(m_{t+1}, \tilde{\pi}_{t+1}) - \frac{1}{2}\text{Var}_t(\tilde{\pi}_{t+1}). \quad (15)$$

The nominal short rate deviates from the Fisher equation through the third and four terms on the right side of (15). The third term recognizes that a nominally safe investment has a stochastic real value. If inflation tends to be unexpectedly high when the marginal value of consumption is high, then investors require additional compensation to hold nominally safe investments. The fourth term accounts for Jensen's inequality. The expected real value of a nominally safe investment is higher when inflation is more volatile.

The next assumption is that a state vector with Gaussian dynamics (1) captures all variation in the real one-period interest rate, the log of the real SDF, and the conditional mean of inflation. The log of the real SDF has the form

$$m_{t+1} = -r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\epsilon_{t+1}, \quad (16)$$

which satisfies the no-arbitrage requirement (14). Element i of the p -vector Λ_t is the compensation investors require to face the risk that $\epsilon_{i,t+1}$ is unexpectedly high. If element i of Λ_t is positive, a positive shock to $\epsilon_{i,t+1}$ corresponds to an unexpectedly low marginal value of consumption. In this case, assets with values that increase in $\epsilon_{i,t+1}$ earn positive risk premia.

The description of the real SDF is completed by specifying the determinants of both the real short rate and risk compensation Λ_t . The minimum set of restrictions we can impose that allow us to price multi-period bonds are that both the short rate and risk compensation

are arbitrary affine functions of the state vector. Formally,

$$r_t = \delta_0 + \delta_1' x_t, \quad (17)$$

$$\Lambda_t = \Sigma^{-1} (\lambda_0 + \lambda_1 x_t). \quad (18)$$

The parameters of the short-rate equation (the scalar δ_0 and the vector δ_1), as well as the parameters of risk dynamics (the vector λ_0 and the matrix λ_1) are unconstrained by the requirement of no-arbitrage. The form of (17) implies that the real short rate varies over time, but does not link the variation to any fundamental story. For example, the short rate is not linked to expected consumption growth or investment. Similarly, the form of (18) implies that risk compensation varies over time, but does not link this variation to any utility-based story. This risk compensation specification is the discrete-time counterpart of the Gaussian essentially affine model of Duffee (2002).

To derive the form of the nominal SDF, assume that inflation has the form

$$\tilde{\pi}_t = A_\pi + B_\pi' x_t + \eta_{\pi,t}, \quad \eta_{\pi,t} \sim N(0, \Omega_\pi). \quad (19)$$

Then some algebra reveals the log of the nominal SDF is

$$m_{t+1}^\$ = -r_t^\$ - \frac{1}{2} \Lambda_t^{\$'} \Lambda_t^\$ - \frac{1}{2} \Omega_\pi - \Lambda_t^{\$'} \epsilon_{t+1} - \eta_{\pi,t+1}. \quad (20)$$

The nominal short rate is

$$r_t^\$ = \delta_0^\$ + \delta_1^{\$'} x_t, \quad (21)$$

$$\delta_0^\$ = \delta_0 + A_\pi + B_\pi' (\mu - \lambda_0) - \frac{1}{2} B_\pi' \Sigma \Sigma' B_\pi - \frac{1}{2} \Omega_\pi, \quad (22)$$

$$\delta_1^{\$'} = \delta_1' + B_\pi' (K - \lambda_0). \quad (23)$$

The compensation for bearing nominal risks is

$$\Lambda_t^{\$} = \Sigma^{-1} (\lambda_0^{\$} + \lambda_1^{\$} x_t), \quad (24)$$

$$\lambda_0^{\$} = \lambda_0 + \Sigma \Sigma' B_{\pi}, \quad (25)$$

$$\lambda_1^{\$} = \lambda_1. \quad (26)$$

To summarize, the real SDF is completely specified by the parameters of the real short rate (17), the state dynamics (1), and real compensation for facing state risk (18). The nominal SDF is completely specified by the parameters of the nominal short rate (21), the state dynamics, and nominal compensation for facing state risk (24). The two SDFs are linked by relation between inflation and the state vector (19).

Finally, note that the maximum conditional Sharpe ratios implied by these two SDFs vary over time. The relevant formulas are

$$\text{max conditional Sharpe for real payoffs} = \sqrt{e^{\Lambda_t' \Lambda_t} - 1} \quad (27)$$

and

$$\text{max conditional Sharpe for nominal payoffs} = \sqrt{e^{\Lambda_t^{\$}' \Lambda_t^{\$}} - 1}. \quad (28)$$

3.2 Bond pricing

The real and nominal SDFs described in Section 3.1 are parameterized symmetrically. Therefore the bond-pricing formulas are also symmetric. The following derivation, which follows Ang and Piazzesi (2003), is for real bonds. The addition of a dollar superscript produces nominal bond prices.

The form of the SDFs imply that zero-coupon bond prices are exponential affine in the

state vector. To verify this claim, assume that bond prices can be written as

$$P_t^{(n)} = \exp(\mathbb{A}(n) + \mathbb{B}(n)'x_t). \quad (29)$$

Then bond yields are affine functions of the state vector,

$$y_t^{(n)} = -\frac{1}{n}(\mathbb{A}(n) + \mathbb{B}(n)'x_t). \quad (30)$$

Plug (16), (17), (1), (18) and (29) into (11) to solve for $A(n)$ and $B(n)$. The result is

$$\begin{aligned} \mathbb{A}(n) + \mathbb{B}(n)'x_t = & \mathbb{A}(n-1) + \mathbb{B}(n-1)'(\mu + Kx_t) - \delta_0 - \delta_1x_t \\ & + \frac{1}{2}\mathbb{B}(n-1)'\Sigma\Sigma'\mathbb{B}(n-1) - \mathbb{B}(n-1)'(\lambda_0 + \lambda_1x_t). \end{aligned}$$

Matching coefficients in x_t produces the difference equation

$$\mathbb{B}(n)' = \mathbb{B}(n-1)'(K - \lambda_1) - \delta_1'. \quad (31)$$

Matching constant terms produces the difference equation

$$\mathbb{A}(n) = \mathbb{A}(n-1) + \mathbb{B}(n-1)'(\mu - \lambda_0) - \delta_0 - \frac{1}{2}\mathbb{B}(n-1)'\Sigma\Sigma'\mathbb{B}(n-1). \quad (32)$$

It is easy to show that these difference equations imply that bond prices are also given by the pseudo-risk-neutral valuation,

$$P_t^{(n)} = e^{-r_t} E_t^Q \left(P_{t+1}^{(n-1)} \right) \quad (33)$$

where the real equivalent martingale dynamics of the state vector are

$$x_{t+1} = \mu^q + K^q x_t + \Sigma \epsilon_{t+1}^q, \quad (34)$$

$$\mu^q = \mu - \lambda_0, \quad K^q = K - \lambda_1. \quad (35)$$

The recursions (31) and (32) begin with the requirement that a maturity-zero bond has a price of one for all x_t , or $\mathbb{A}(0) = 0, \mathbb{B}(0) = 0$. The solution to (31) is

$$\mathbb{B}(n)' = -\delta_1' (I - K^q)^{-1} (I - (K^q)^n) \quad (36)$$

There is no closed-form solution to (32).

The vector λ_0 and the matrix λ_1 determine risk compensation. The log expected return to an n -period bond from t to $t + 1$ is

$$\log E_t \left(R_{t+1}^{(n)} \right) = r_t + \mathbb{B}(n - 1)' (\lambda_0 + \lambda_1 x_t). \quad (37)$$

A close look at this equation is warranted. The vector $\lambda_0 + \lambda_1 x_t$ is the sensitivity of the marginal value of a unit of consumption to state-vector shocks. If, say, element i of the vector is 0.02, then a positive one standard deviation shock to $x_{i,t+1}$ lowers the marginal value of a unit of consumption at $t + 1$ by two percent (a log change of -0.02). The vector $\mathbb{B}(n - 1)$ is the sensitivity of an n -maturity real bond's real price to state shocks at $t + 1$. The log expected excess return is the product of the bond-price sensitivity to the state and the sensitivity of the marginal value of consumption to the state. Aside from a superscript, the same equation applies to the nominal SDF and nominal prices of nominal bonds.

3.3 Implications of no-arbitrage restrictions

No-arbitrage restrictions are inherently cross-sectional. They link the time- t price of one financial instrument to time- t prices of other instruments. This feature of the restrictions has two implications relevant to empirical analysis. First, the empirical content of the restrictions depends on the number of yields observed at t . Second, the restrictions have only an indirect effect on estimates of a model's dynamics.

To clarify these points, compare the general factor model of Section 2.1 to the no-arbitrage model of Section 3.2. Assume that we observe a panel of nominal bond yields and other macroeconomic variables. Stack the observed yields in the d -vector $\tilde{y}_t^\$$ and the other variables in the c -vector \tilde{f}_t . The general factor model's measurement equation is

$$\begin{pmatrix} \tilde{y}_t^\$ \\ \tilde{f}_t \end{pmatrix} = \begin{pmatrix} A_y^\$ \\ A_f \end{pmatrix} + \begin{pmatrix} B_y^\$ \\ B_f \end{pmatrix} x_t + \eta_t. \quad (38)$$

Recall that the state vector can be arbitrarily transformed using (5). The parameters of the term structure model are also transformed following the rotations described by Dai and Singleton (2000). It is convenient to translate the state vector such that the parameters of (1) satisfy $\mu = 0$, K diagonal, and Σ lower triangular with ones along the diagonal. These restrictions exactly identify the state vector without imposing any normalizations on the measurement equation (38).

With the general factor model, the vector $A_y^\$$ and $B_y^\$$ are unrestricted. They contain a total of $d(1+p)$ free parameters. Similarly, A_f and B_f are unrestricted. This general model nests the model with no-arbitrage restrictions. By imposing no-arbitrage, $A_y^\$$ and $B_y^\$$ are determined by the $(1+p)^2$ parameters of (21) and (24). Therefore no-arbitrage imposes $(d-p-1)(1+p)$ overidentifying restrictions on $A_y^\$$ and $B_y^\$$.

The intuition behind the number of restrictions is straightforward. Expected returns to bonds are determined by the dynamic properties of the short rate and p prices of risk. A panel of $p+1$ bond yields exactly identifies these dynamic properties. Each additional yield adds $p+1$ restrictions, corresponding to the bond's mean excess return and the sensitivity of the bond's conditional mean excess return to the p -dimensional state vector. For example, a four-factor model fit to a panel of six bond yields has five overidentifying parametric restrictions.

No-arbitrage imposes no restrictions on the dynamics of the state vector. It also imposes no restrictions on A_f and B_f . This is true even if one of the elements of f_t is inflation. The

nominal SDF, combined with the parameters of the inflation mapping (19), implies the real SDF. But since the measurement equation (38) does not include observations of real bond yields, there are no restrictions imposed on the inflation process.

Cross-sectional restrictions have an effect on estimates of a model's dynamics to the extent they alter how the estimation procedure infers dynamics from observables. The measurement equation (38) allows the observations to have non-factor components. These drive a wedge between the dynamics of observables and dynamics of the state vector. The information in any one observed variable about the dynamics of the state depends on the variance of that variable's cross-sectional error. The no-arbitrage model imposes cross-sectional restrictions, and therefore alters estimates of cross-sectional errors relative to the general factor model.

Therefore the effect of no-arbitrage restrictions on estimated dynamics critically depends on the way in which the cross-sectional errors are modeled. For example, if p linear combinations of observed variables are assumed to be observed without error, Joslin, Singleton, and Zhu (2011) show that no-arbitrage restrictions are irrelevant to estimates of state dynamics. Joslin, Le, and Singleton (2011) note that empirically, the covariance matrix of the shocks in (38) is not well-described by a diagonal matrix. They conclude that when the covariance matrix is specified more accurately, there is very little scope for no-arbitrage restrictions to affect estimates of factor dynamics of macro-finance models.

4 The variation of yields with the macroeconomy: U.S. evidence

The main objective of a macro-finance model is to link yields to the macroeconomy. But what part of the macroeconomy? Naturally, inflation should be an important determinant of nominal yields. In addition, the logic of the Taylor rule also points to the output gap, or more generally, the cyclic component of macroeconomic activity. Measures of economic growth are also plausibly related to real and nominal yields. For example, standard representative-

agent models imply that risk-free real rates are affected by expected aggregate consumption growth.

The empirical macro-finance literature uses a variety of measures of inflation, cyclical economic activity, and economic growth. This section discusses these measures. It then examines the strength of the links between macroeconomic variables and bond yields. Two strong and important conclusions are that much of the variation in economic activity over time is not accompanied by variations in nominal bond yields, and vice versa.

4.1 Macroeconomic data

Inflation is almost always constructed from price indexes such as the CPI, the CPI excluding food and energy, the GDP price index, or a consumption price index. For most purposes the choice of inflation measure is unimportant because they are highly correlated over horizons examined in the macro-finance literature, such as quarterly and annually. For example, over the period 1952Q2 through 2010Q4, quarter-to-quarter inflation measures constructed from the CPI and from the GDP price index have a correlation of 0.8.²

There is no unambiguous measure of the output gap. Five quarterly measures are studied. One is detrended real GDP, defined as the residual of a regression of log real GDP on a constant and a linear time trend. Another is detrended industrial production, defined similarly. The level of industrial production in a given quarter is defined as the value for the quarter's final month. Two other measures use the Hodrick-Prescott filter to remove the non-cyclic components of log real GDP and log industrial production, respectively. The final measure is demeaned industrial capacity utilization, used by Rudebusch and Wu (2008). As with industrial production, the level in a given quarter is the value for the quarter's final month. These output gap measures are highly persistent, with quarterly autocorrelations between 0.83 (HP-filtered industrial production) and 0.99 (regression-detrended real GDP).

Table 1 reports correlations among the measures. Correlations between cyclic components

²In this chapter, quarterly inflation based on monthly price indexes is calculated using log-changes from the final month of the previous quarter to the final month of the current quarter.

of real GDP and industrial production are high (0.9 to 0.95) when the non-cyclic component is removed using a common technique. But when the techniques differ, correlations between the cyclic measures are around 0.45. Correlations between capacity utilization and the other measures are around 0.7.

Table 1 also reports correlations involving five quarterly measures of economic growth. Three are quarter-to-quarter changes in the log of per capita real consumption of nondurables and services, real GDP, and industrial production. Another is the Federal Reserve Bank of Chicago's National Activity Index (CFNAI). The quarterly value is the average of the quarter's three monthly values. The final measure of economic growth is the first principal component of a large panel of data related to macroeconomic activity and financial market conditions. This measure is constructed by Ludvigson and Ng (2010). The sign of this component is chosen such that an increase in the measure corresponds to higher growth. The level of this principal component in a given quarter is the value for the quarter's final month. These measures of economic growth are less persistent than measures of the output gap, with quarterly autocorrelations ranging from 0.37 (consumption growth) to 0.74 (CFNAI).

According to Table 1, the measures of economic growth are modestly correlated with measures of the output gap. A typical cross-correlation is about 0.25, with a range from 0.03 to 0.50. Correlations among the measures of economic growth are much higher than these cross-correlations, averaging around 0.67.

4.2 Spanning

A natural modeling approach to link nominal yields with the macroeconomy is to define a vector of observables comprised of nominal yields, an inflation measure, and a measure of economic activity. The vector's dynamics are determined by the dynamics of a length- p state vector as in (1) and (2).

Since the state vector determines the persistent components of both inflation and economic activity, a natural inference is that the state vector can be rotated into p linear

combinations of bond yields. In the terminology of Section 2.3, there are p yields such that the $p \times p$ submatrix of \mathbf{B} is invertible; the yields can be treated as the state vector. Note that this argument does not require that inflation and economic activity are the *only* determinants of yields. For example, an ‘animal spirits’ factor may partially determine the compensation investors require to bear interest rate risk. Each bond yield will then have its own loading on animal spirits, and there will be linear combinations of yields that are invariant to animal spirits (the linear combination of animal spirit loadings sums to zero). Qualitatively, invertibility means that some of these linear combinations have nonzero loadings on the factors that drive inflation and economic activity.

Duffee (2011) describes a slight hole in this logic that is exploited empirically by Joslin, Priebisch, and Singleton (2010). It is possible that a shock to some macroeconomic factor has offsetting effects on expected future short rates and bond risk premia, such that bond yields do not react to the shock. But this is a knife-edge restriction of non-invertibility. If yields are truly determined by the macroeconomy, it is implausible that there is a large wedge between the factors driving the macroeconomy and the factors driving bond yields.

There is a simple, albeit informal, way to examine empirically this implication. As stated in (4), the persistent components of the observables are linear combinations of the state vector. A regression of the observables on the state vector should produce serially uncorrelated residuals. Therefore if yields span the state vector, a regression of the macro variables on bond yields should produce fitted residuals that are close to uncorrelated. The meaning of “close” is both in the eye of the beholder and dependent on the overall fit of the regression. If, for example, a regression of inflation on nominal yields produces an R^2 of 99 percent, the serial correlation of the residuals is irrelevant. Economically, we could say that yields effectively span inflation.

Table 2 reports results of these regressions. Nominal yields on six zero-coupon bonds are from the Center for Research in Security Prices (CRSP). The maturities are three months and one through five years. The yields, real GDP, and industrial production are available for

the sample period 1952Q2 through 2010Q4. Capacity utilization is available beginning with 1967Q1. Results are reported for the longest available sample, as well as the period 1986Q1 through 2007Q4. This shorter sample is arguably characterized by a single monetary policy regime.

The results for the output gap cast considerable doubt on the spanning implication. The regression residuals are not close to serially uncorrelated. Across the two samples and the five measures, the smallest estimated first-order correlation with these quarterly data is 0.70. In addition, the R^2 s are small. Their average for the full sample is less than 0.4, and their average for the shorter sample is about 0.5. In other words, much of the predictable variation in the output gap does not show up in the nominal yield curve.

The same regressions are estimated for the five measures of economic growth. These results are less conclusive, and thus less discouraging. The R^2 s are much lower than those for the output gap regressions. For log-differenced real consumption, real GDP, and industrial production, the R^2 s are around 5 percent for the full sample and around 20 percent for the shorter sample. However, the residuals are also substantially less persistent; the largest estimated first-order serial correlation is only 0.35. This suggests that much of what these regressions are missing may be variability created by white-noise shocks. The regressions for CFNAI and the Ludvigson-Ng principal component have somewhat higher R^2 s but also somewhat higher residual serial correlations. This pattern suggests that these indexes smooth out some of the noise in individual measures of economic growth.

Finally, Table 2 reports R^2 s for measures of inflation. As with output gap measures, the results are discouraging. For inflation measured with the GDP deflator and the CPI, the R^2 s are less than 0.5 for the full sample and less than 0.3 for the shorter sample. The R^2 s for the CPI excluding food and energy are higher, especially for the short sample. It is not clear why this measure is closely related to the term structure over the period 1986 through 2007, especially since its full-sample R^2 is close to the R^2 s for the GDP deflator and the CPI.

Another way to examine whether yields span the predictable component of economic growth is to look at forecasting regressions. The state vector follows a first-order autoregressive process. Therefore the state vector contains all information relevant to forecasting future values of the observables. If yields span the state vector, the same is true of yields.

To examine this implication for, say, consumption growth, first estimate a forecasting regression such as

$$\Delta\tilde{c}_{t+j} = b_0 + \sum_{i=0}^1 a_{1+i}\Delta\tilde{c}_{t-i} + \sum_{i=0}^1 a_{3+i}\tilde{\pi}_{t-i} + a'_5\tilde{y}_t^{\$} + e_{t,t+j}.$$

In this regression, the j -ahead forecast of consumption growth is predicted using current and lagged realizations of consumption growth and inflation, along with the current term structure of nominal yields. Next regress the fitted values from this regression on the term structure:

$$\hat{E}_t\Delta\tilde{c}_{t+j} = b_0 + b'_1\tilde{y}_t^{\$} + e_t.$$

In an infinite sample, the first-stage regression does not overfit predictability. Therefore the second-stage regression should have an R^2 of one. In finite samples, overfitting in the first stage (more precisely overfitting through the macro variables on the right side) lowers the R^2 of the second stage regression.

Table 3 reports regression results for two-quarter-ahead forecasts. In all of the first-stage regressions, two lags of the measure of economic growth and two lags of CPI inflation, along with nominal yields, are used to form forecasts. The table reports the R^2 s for the first-stage and second-stage regressions, as well as the serial correlation of residuals for the second-stage regression.

The results are mixed. Over the full sample 1952 through 2010, the second-stage R^2 s are fairly high—in the neighborhood of 0.75—for log-differenced per capita consumption, real GDP, and industrial production. But over the more homogeneous sample of 1986 through 2007, these R^2 s are much smaller, ranging from 0.3 to 0.5. Moreover, there is little first-

stage predictability of these measures. There is greater first-stage predictability of the two measures designed to capture economic growth. The first-stage R^2 s for CFNAI and the Ludvigson-Ng principal component are around 0.4. The corresponding second-stage R^2 s vary widely across these two measures and two sample periods, from 0.4 to close to 0.9. Hence it is possible to cherry-pick measures of growth and sample periods that are consistent with spanning, but spanning is not a robust result.

The main message of Tables 2 and 3 is that yields do not span the predictable components of economic activity and inflation. We can reverse the question and ask to what extent economic activity and inflation span yields. The analysis is less straightforward in this case. Our economic intuition tells us that yields impound information about expected future inflation and economic activity, but other factors can also drive yields. For example, yields may depend on habit levels and conditional volatilities of economic growth. In addition, we do not have precise measures of expected inflation and economic growth.

Table 4 reports results of regressing quarter- t nominal yields on quarter- t values of HP-filtered log real GDP and log industrial production, log-differenced industrial production for quarters $t - 2$, $t - 1$, and t , and ARMA-smoothed CPI inflation for quarters $t - 2$, $t - 1$, and t . The quarter- t value of this smoothed series is the one-quarter-ahead forecast of CPI inflation from an ARMA(1,1) description of the time series. Following the previous tables, results are reported for both the full sample 1952Q2 through 2010Q4 and the shorter sample 1986Q1 through 2007Q4.

The results are roughly consistent across the two samples. The R^2 s are all in the range of 0.5 to 0.7. The serial correlations of the residuals are all in the range of 0.8 to 0.9. The standard deviations of the residuals are all well above one percentage point of annualized yields. In sum, a large fraction of the variation in nominal yields is not related to current or past inflation and economic growth. These results are not necessarily bad news for models that explain variations in yields with macroeconomic fundamentals. The evidence can be reconciled with such models as long as yields contain substantial information about future

long-horizon inflation or economic growth that is not contained in current and lagged macro variables. Evaluating this possibility requires placing additional economic structure on the factor model. First, however, we examine how the lack of spanning affects estimates of dynamic joint models of yields and the macroeconomy.

4.3 A workhorse empirical example

This subsection applies the models of Section 2.1 and Section 3.2 to quarterly observations of inflation, growth in per capita real consumption of nondurables and services, and nominal Treasury bond yields. The sample period is 1952Q2 through 2010Q4. One goal of this exercise is to illustrate some implications of the lack of spanning. Another is to produce parameter estimates to be used in Section 5's discussion of bond risk compensation

Four versions of these models are estimated. All are four-factor models with the transition equation (1). The first model uses observations of inflation, consumption growth, and yields on three-month and five-year bonds. There is no reason to impose no-arbitrage, since it has no bite with four factors and only two yields. The second model adds intermediate maturities of one through four years. With a total of six bond yields and four factors, no-arbitrage imposes cross-sectional restrictions, but the second model does not apply them. The restrictions are imposed with the third estimated model. The fourth model is described in Section 4.4. Estimation is with maximum likelihood via the Kalman filter.

Table 5 contains information about the models' accuracy. Panel A reports the square root of the time series mean of $\hat{\eta}_{i,t}^2$ for each observed time series i . In other words, it reports the root mean squared cross-sectional deviation from an exact four-factor model. Panel B reports the root mean squared forecast error for one-quarter-ahead forecasts.

The first model, which uses only two bond yields, is the benchmark for evaluating the effects of including additional bond yields and imposing no-arbitrage. With four factors and four observables, the model can be parameterized such that there are no cross-sectional deviations from an exact factor model. In other words, the model nests a VAR(1). In the

model, observables are subject to two kinds of shocks: persistent shocks and shocks that die out completely by the next quarter. The latter are the deviations from an exact factor model. Maximum likelihood chooses parameters that imply large cross-sectional deviations for the macro variables. Root mean squared cross-sectional deviations for inflation and consumption growth are about two and one percentage points, respectively. The corresponding values for the three-month and five-year yields are 31 and 1 basis points. Consistent with their interpretation as white-noise shocks, the serial correlations of the fitted deviations are very close to zero, ranging from -0.01 for inflation to 0.05 for the three-month yield. (These correlations are not reported in any table.)

Visual evidence of the importance of the two kinds of shocks is in Fig. 1, which displays the time series of actual and filtered estimates of inflation and consumption growth. The figure shows that the cross-sectional deviations are fairly large relative to the overall variation in the time series. Presumably if the model were estimated at a higher frequency, say monthly, the white-noise shocks would be less important.

The first model uses four factors to describe the joint dynamics of inflation, consumption growth, the short end of the yield curve, and the long end. The second model uses four factors to describe the dynamics of the macro variables and the entire yield curve (more precisely, six points on the curve). In principle, better forecasts of inflation and consumption growth may be produced by using entire yield curve than by using just two points on the curve. In practice, the main effect of adding these data is to weaken the cross-sectional fit of the four-factor model. Table 5 reports that the root mean squared cross-sectional deviation for consumption growth increases from 1.1 to 1.7 percentage points. In addition (not reported in the table), many of the serial correlations of the fitted cross-sectional deviations are substantially different from zero. For example, they are 0.22 and 0.21 for the three-month yield and consumption growth, respectively.

This cross-sectional deterioration implies that four factors are not sufficient to capture the persistent components of the entire nominal term structure, inflation, and consumption

growth. The fitted cross-sectional deviations absorb some of these components. Litterman and Scheinkman (1991) conclude that three factors are needed to describe the cross-section of the nominal term structure. These results tell us that one additional factor cannot describe the orthogonal variation in the persistent components of inflation and consumption growth. Since we know from Tables 2 and 3 that yields do not span these persistent components, the misspecification is not surprising.

In this example, the likelihood is maximized when the model's factors are primarily devoted to fitting the cross-section of the nominal yield curve. This is a typical result in the macro-finance literature when the observable related to economic activity is a measure of economic growth. Section 6 estimates a model in which a measure of the output gap is observed instead of economic growth. In that case, the model fits the output gap at the expense of the term structure.

The conclusion that this model is misspecified has nothing to do with no-arbitrage, since the second model does not impose it. The third model includes no-arbitrage restrictions. A glance at Table 5 reveals that these restrictions have almost no effect on either cross-sectional fit or forecast accuracy. This result is consistent with the analysis of Joslin, Le, and Singleton (2011).

One potential method to reduce the persistence in the cross-sectional deviations is to add more factors, giving the model more cross-sectional flexibility. However, this also gives the model more dynamic flexibility, leading to overfitting. A five-factor version of the no-arbitrage model described in Section 3.2 has thirty parameters devoted to fitting risk premia. In a yields-only setting, Duffee (2010) concludes that the in-sample conditional maximum Sharpe ratios, defined by (28), are astronomically large (greater than 10^{20}) for the five-factor model. The next subsection considers an alternative way to handle this misspecification.

4.4 Interpreting and altering cross-sectional accuracy

We cannot infer that the second and third estimated models have too few factors simply by examining the magnitude of the estimated model’s cross-sectional deviations. These models have small cross-sectional yield deviations and large cross-sectional macroeconomic deviations, but so does the first estimated model. However, researchers working with macro-finance models typically view such large macroeconomic deviations as unacceptable. An almost universal assumption in the empirical macro-finance literature is that there are no macroeconomic cross-sectional deviations. More precisely, the variances of the cross-sectional deviations are fixed to zero. These restrictions force ML estimation to reproduce exactly the cross-section of the macro variables.

The logic behind this approach is unclear. It may reflect confusion about the role of cross-sectional deviations in a factor model. As noted earlier, the deviations are often called “noise” or “measurement error” in the literature. If these terms accurately characterize the deviations, the approach has merit. From this perspective, the estimated models in Table 5 are able to fit the cross-section of yields only by treating much of the variation in inflation and economic growth as measurement error. Imposing zero cross-sectional deviations on the macro variables is a way to override the likelihood function’s emphasis on economically small, but statistically strong, common components of the term structure, and place the emphasis on the components common to the macro variables. An extreme example of this approach is a parameterization in Joslin, Le, and Singleton (2011) that uses only three factors, and equates two of them with observed inflation and real activity.

However, Section 2.1 explains that it does not make sense to treat macroeconomic cross-sectional deviations as purely measurement error. Moreover, forcing a model to treat white-noise shocks as persistent, whether or not the shocks are measurement error, is another form of misspecification. Therefore this approach is often combined with pre-smoothing of the macro data. For example, Ang and Piazzesi (2003) use measures of inflation and real activity that are cross-sectionally smoothed, using principal components. They also smooth

across time by calculating inflation and economic growth at annual horizons rather than the horizon that corresponds to their data frequency. Many researchers, such as Rudebusch and Wu (2008) and Joslin, Le, and Singleton (2011), follow in their footsteps. Data can also be pre-filtered, such as using one-period-ahead forecasts from an ARMA model in place of the actual data. Duffee (2011) uses this approach, following an early version of Joslin, Priebisch, and Singleton (2010).

To illustrate this approach, univariate ARMA(1,1) processes are fit to inflation and consumption growth. The ARMA quarter- t predictions of inflation and consumption growth during quarter $t + 1$ are used as exactly-observed measures of quarter- t inflation and consumption growth. Then the no-arbitrage model using six bond yields, inflation, and consumption growth is reestimated. Results are displayed in the final column of Table 5. The root mean squared deviations and forecasting errors for the macro variables are calculated using actual inflation and consumption growth. For example, the cross-sectional root mean squared deviation for inflation is the square root of the mean squared difference between CPI inflation and the one-quarter-ahead forecast from the ARMA model.

The results are not particularly surprising. The most noticeable change relative to the third estimated model is the large drop in the the magnitude of cross-sectional deviations of consumption growth. The serial correlation of the fitted deviations also drops, from 0.2 to 0.12. This estimated model trades off more a more accurate fit of consumption growth for a less accurate fit of the three-month yield. Its root mean squared deviation increases from 25 to 42 basis points, and the serial correlation of the fitted residuals increases from 0.22 to 0.45.

From the perspective of forecasting actual inflation and consumption growth, the use of smoothed macro data is somewhat counterproductive. Table 5 shows that among the four estimated models, this version has the largest root mean squared forecast errors for both inflation and consumption growth. The reason is that the model parameters are chosen by ML to best predict smoothed macro data, not actual macro data. However, the RMSE

differences across the estimated models are on the order of a few basis points.

Is it sensible for researchers to assume that measures of inflation and economic growth have no cross-sectional deviations? Based on this evidence, which uses pre-smoothed macro data, the effects of adopting the assumption are small. It slightly alters the cross-sectional deviations relative to unsmoothed data, and slightly reduces forecast accuracy. If the assumption were applied to the unsmoothed macro data—an approach that is not typically taken in the literature—the effects would likely be large and undesirable.

The final point to glean from this analysis is that if our goal is to use ML to estimate a four-factor joint model of inflation, consumption growth, and yields, we are probably better off parameterizing the model with just a short-maturity yield and a long-maturity yield. It is better to throw away the information in the additional yields rather than to attempt to squeeze the information into the four-factor structure. This conclusion does not apply to more flexible estimation procedure such as GMM, since we could use economically-motivated moments and weights rather than likelihood-based moments and weights.

5 Modeling risk premia

The no-arbitrage model of Section 3.2 is silent about the sources of risk premia and the reasons why risk premia vary over time. The model simply says that the law of one price holds. This section discusses possible restrictions on risk premia dynamics. One reason for such restrictions is to limit the problem of overfitting. As mentioned in Section 4.3, highly flexible specifications of risk premia can generate outlandish results. More importantly, we want to understand the fundamental determinants of bond risk premia by linking them to utility-based models.

This section describes theories about bond risk premia and presents relevant evidence. It is a very long section with a single theme running throughout. There is little to no evidence that any existing theory explains the observed properties of bond risk premia.

5.1 Practical approaches to modeling risk premia

The pure dynamic term structure literature (i.e., the literature that ignores the macroeconomy) has long wrestled with the problem of overfitting bond risk premia. In a p -factor model, $p(p + 1)$ parameters determine how compensation for p risks varies over time. Restrictions on these parameters are almost always imposed in applied work. However, it is difficult to express a utility-based motivation for risk premia purely in terms of yields. Thus the restrictions are usually motivated on statistical grounds rather than economic principles.

This type of motivation also shows up in the macro-finance literature. For example, Christensen, Lopez, and Rudebusch (2010) and Joslin, Priebsch, and Singleton (2010) use various information criteria to decide among possible parametric restrictions. Simple parsimony also motivates some restrictions in macro-finance models, such as those in Ang and Piazzesi (2003) and Ang, Bekaert, and Wei (2008).

Although many macro-finance models have enough content to allow for economically rigorous models of risk premia, applied work almost universally avoids these models. The reason is that standard utility frameworks have great difficulty generating basic features of the nominal yield curve. We explore this difficulty in depth beginning with Section 5.3. First, however, we discuss another motivation for restrictions on risk premia. Macro-finance models can be used to calculate the compensation investors require to face macro risks, such as shocks to inflation and real activity. Even if an estimated model produces accurate estimates of the compensation for holding bonds, it does not necessarily produce accurate estimates of compensation for macro risk. As the next subsection documents, additional restrictions may be required.

5.2 A brief example

This example illustrates problems with inferring properties of risk premia from models that impose no-arbitrage but do not otherwise restrict risk premia dynamics. Two such models are estimated in Section 4.3. They are the third and four estimated models in Table 5.

Both are four-factor models that differ in their assumptions about cross-sectional deviations in observed inflation and consumption growth. Table 6 lists some of the unconditional properties of the nominal and real term structures and the real stochastic discount factor that are implied by the models' parameter estimates.

Both estimated models imply that the unconditional mean yield curve slopes up. The models agree that the mean five-year yield less the mean three-month yield is about one percent. The estimates have fairly tight standard errors. The usual interpretation of the upward-sloped nominal yield curve is that inflation risk is priced, in the sense that on average, investors require compensation to hold assets that decline in value when inflation unexpectedly increases. The models allow us to determine (up to sampling error) whether this interpretation is correct.

The intuition behind this exercise is straightforward. There are four shocks in the model. The combination of the sensitivity of log bond prices to these shocks and the mean excess returns to bonds allow calculation of the mean compensation investors require to face each shock. Since inflation is included as an observable, shocks to inflation can be decomposed into weighted averages of the four factor shocks. Therefore we can infer the compensation investors require to face this weighted average of factor shocks. Table 6 reports model-implied mean compensation for inflation risk. The units are the expected excess return, in percent per quarter, for an asset that declines in price by one percent when inflation is one percent higher than expected. If inflation shocks were permanent, this would be approximately the inflation compensation for a one-year nominal bond.

Surprisingly, the point estimates of the first no-arbitrage model imply that inflation bets are, on average, hedges. Investors pay half a percent per quarter to make the inflation bet. Stated somewhat differently, a one standard deviation positive shock to inflation from quarter t to quarter $t + 1$ lowers the marginal value of a real unit of consumption by half a percent.

How can the model generate a positively sloped nominal yield curve with a negative price

of inflation risk? The reason is that there are two main factors driving variation in the level of the term structure, one of which is inflation. (The other can be rotated arbitrarily, as long as it is not perfectly correlated with inflation.) A positive price of risk on either of these factors produces an upward-sloped term structure. Maximum likelihood assigns a negative risk premia to inflation and a large positive risk premia to the other factor. Their sum is slightly positive, which generates the modestly upward-sloped yield curve. The standard error on the inflation risk premium indicates that there is not much information in the sample about the true decomposition of risk premia between these two factors.

Since a three-month nominal bond is risky in real terms, there is a negative inflation compensation built into its mean yield. This implies a relatively high real risk-free rate. Starting with the expression for the nominal risk-free rate (15), substitute in the state dynamics (1), the functional form of the real SDF (16), and the functional form of risk compensation (18). The result is

$$r_t^{\$} = r_t + E_t(\pi_{t+1}) - B'_\pi(\lambda_0 + \lambda_1 x_t) - \frac{1}{2} B'_\pi \Sigma \Sigma' B_\pi - \frac{1}{2} \Omega_\pi, \quad (39)$$

where B_π is the loading of inflation on the state vector and Ω_π is the variance of the white-noise component of inflation. The third term on the right is the inflation compensation for holding the nominal risk-free rate. The model-implied unconditional mean inflation rate is 3.13 percent per year, thus the mean real rate implied by the Fisher equation is about one percent per year. But adjusting for the negative risk premium in the nominal risk-free rate, the model-implied mean real risk-free rate is about three percent per year. Long-term real bonds have large exposures to the non-inflation factor that drives nominal yields, thus they have large average risk premia. The model-implied slope of the real yield curve (five-year yield less three-monthly yield) is about five percent in annualized yields.

The second no-arbitrage model has an even more unreasonable price of inflation risk. The huge standard error indicates that there is almost no information in the data that

distinguishes between the price of inflation risk and the price of the other factor that drives nominal yields.

A similar problem holds for the risk premium associated with consumption growth. According to the point estimates of both models, assets exposed to consumption growth are a hedge. For the first model, a one standard deviation negative shock to aggregate consumption from quarter t to quarter $t + 1$ lowers the marginal value of a unit of consumption by 0.15 percent. For the second model, the decline in marginal utility is an astronomical 32 percent. Again, the standard errors indicate there is little information in the sample about the price of consumption risk.

It is clear from these results that without additional structure, there is not enough information in the nominal term structure to determine risk compensation for inflation and consumption growth. One way to circumvent this problem, at least for the inflation risk premium, is to use additional data. The term structure model can be estimated using yields on both nominal and indexed bonds, effectively using their spread to infer inflation compensation. D'Amico, Kim, and Wei (2008) take this approach with U.S. data. Unfortunately, they conclude that owing to liquidity problems, yields on indexed Treasury bonds are reliable only for the post-2002 period. Moreover, indexed bond yields do not help pin down compensation for consumption risk.

The most natural type of structure to impose on risk premia comes from consumption-based models of utility. However, our standard models are largely inconsistent with the behavior of nominal yields. To understand the problems, begin with a quick look at empirical properties of bond returns.

5.3 Some properties of observed bond returns

The workhorse models used in finance to explain risk premia are extensions of the classic case of a representative agent with power utility and an aggregate consumption endowment. An asset's conditional expected excess return depends on the return's conditional covariance with

consumption growth (power utility and habit formation as in Campbell and Cochrane (1999)) and with the return to total wealth (recursive utility as in Epstein and Zin (1989)). Therefore it is useful to summarize how excess bond returns covary with aggregate consumption growth and aggregate stock returns.

Excess bond returns are measured by the return to a portfolio of Treasury bonds with maturities between five and ten years. Quarter-end to quarter-end returns are simple returns cumulated from monthly returns from CRSP. The simple return to a three-month T-bill is subtracted to form excess returns. Excess returns to the aggregate stock market are constructed in the same way, using the CRSP value-weighted index.

Table 7 reports correlations among these excess returns and log aggregate consumption growth. The table also includes inflation and log growth in industrial production (IP). The latter variable is defined for quarter t as the log change in IP from the final month of quarter $t - 1$ to the final month of quarter t . The sample is 1952Q2 through 2010Q4. The most important information in the table is that at the quarterly frequency, excess Treasury bond returns are countercyclical. The correlations with consumption growth and industrial production growth are both negative.

This pattern may surprise some readers. The influential handbook chapter of Campbell (2003) reports that the correlation between excess bond returns and consumption growth is slightly positive for the U.S., although the evidence for other countries is mixed. Campbell uses the “beginning of period” assumption for consumption growth. In Table 7 that is the correlation between excess bond returns and the lead of consumption growth, which is slightly positive. The motivation behind the beginning of period assumption is that aggregate stock returns are more closely correlated with future consumption growth than contemporaneous consumption growth. The usual interpretation is that the shock underlying the aggregate stock return is not fully incorporated into measured consumption until the next quarter. But that is not the case for the shock underlying bond returns; the immediate reaction of both consumption and industrial production is much stronger than the one-quarter-ahead

reaction.

The table also reports that the correlation between aggregate stock returns and aggregate consumption growth is close to zero. This fact, combined with the negative correlation with consumption growth, spells trouble for simple consumption-based models of the term structure. On average, the nominal yield curve in the U.S. slopes up. The sign of this slope implies that expected excess returns to Treasury bonds are positive. But these correlations suggest that the risk premium should be negative.

More formal evidence is in Table 8, which reports results of regressing excess bond returns on either contemporaneous consumption growth or contemporaneous aggregate stock returns. The table also reports the sample mean of excess returns. Over the full sample 1952 through 2010, the consumption beta of excess bond returns is in the neighborhood of minus one, while the stock market beta is almost exactly zero. Neither of these coefficients is statistically different from zero at the five percent level. The same description applies to results displayed for the first and second halves of the sample. Yet the two halves differ from each other in an important dimension: the mean excess bond return in the first half is weakly negative, while the mean excess return in the second half is strongly positive. The table also reports for the shorter, more homogeneous sample period of 1986 through 2007. For this sample the mean excess return is positive and the consumption beta is negative; both are statistically different from zero at the five percent level.

There are, of course, a variety of ways to reconcile these betas with a positively-sloped term structure. For example, it is possible that conditional betas covary with conditional risk premia in a manner that generates both unconditionally positive mean excess returns and unconditionally negative consumption betas. In addition, the aggregate stock return is not the return to total wealth. It is possible that the beta with respect to the return to total wealth is positive even though the beta with respect to the stock market is zero. Moreover, there are alternative utility-based models that introduce heterogeneous agents, heterogeneous goods, and different utility specifications.

However, all of these approaches must match both the risk premia on nominal Treasury bonds and the joint dynamics of inflation and the nominal short-term interest rate. This is a significant challenge, as the next subsections discuss.

5.4 Power utility

We first consider term structure behavior in an exchange-economy setting where a representative agent has power utility. The power utility real SDF is

$$M_{t+1} = \beta \left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\gamma}. \quad (40)$$

where \tilde{C}_{t+1} is aggregate consumption. The real short rate is

$$r_t = -\log \beta + \gamma E_t(\Delta \tilde{c}_{t+1}) - \frac{\gamma^2}{2} \text{Var}_t(\Delta \tilde{c}_{t+1}). \quad (41)$$

The nominal short rate is

$$r_t^{\$} = r_t + E_t(\tilde{\pi}_{t+1}) - \gamma \text{Cov}_t(\Delta \tilde{c}_{t+1}, \tilde{\pi}_{t+1}) - \frac{1}{2} \text{Var}_t(\tilde{\pi}_{t+1}) \quad (42)$$

Both short rates move one for one with the coefficient of relative risk aversion times one-period-ahead expected consumption growth. This feature produces the risk-free rate puzzle of Weil (1989). For plausible values of the coefficient of relative risk aversion and the sample mean of aggregate consumption growth, power utility implies mean real and nominal short rates that are unrealistically large.

Since the risk-free rate puzzle is a well-known result, this discussion ignores misspecification of the average level of short rates to focus on bond risk premia. Backus, Gregory, and Zin (1989) and den Haan (1995) use simple models to argue that the power utility setting generates bond risk premia that have the wrong sign. The analysis that follows is in the spirit of Piazzesi and Schneider (2007) because the model is easily extended to the case of

recursive utility.

The intuition underlying the sign of nominal bond risk premia is straightforward. Investors fear negative shocks to consumption growth. In U.S. data, shocks to consumption growth are positively autocorrelated. They are also positively correlated with shocks to expected future inflation. Therefore a negative shock to consumption growth lowers real rates, through the second term in (41), and lowers nominal rates, through both the first and second terms in (42). Hence nominal bond prices increase when consumption growth falls; nominal bonds are a hedge.

To embed this intuition in the linear Gaussian framework, assume there is a state vector with dynamics (1). The mapping from the state vector to per capita consumption growth and inflation is

$$\begin{pmatrix} \Delta \tilde{c}_t \\ \tilde{\pi}_t \end{pmatrix} = \begin{pmatrix} A_c \\ A_\pi \end{pmatrix} + \begin{pmatrix} B'_c \\ B'_\pi \end{pmatrix} x_t + \begin{pmatrix} \eta_{c,t} \\ \eta_{\pi,t} \end{pmatrix}. \quad (43)$$

With these dynamics, consumption growth is homoskedastic with a time-varying conditional mean. Recall that tildes represent observed values, which for this analysis are assumed to be uncontaminated by measurement error. The white-noise shocks are simply components of consumption growth and inflation that do not persist.

The real short rate is

$$r_t = \delta_0 + \gamma B'_c K x_t \quad (44)$$

where δ_0 is a constant determined by the model's parameters. The nominal short rate is

$$r_t^{\$} = \delta_0^{\$} + (\gamma B'_c + B'_\pi) K x_t, \quad (45)$$

where $\delta_0^{\$}$ is a constant. Risk premia and volatilities are constant in this example, so the nominal short rate varies only with the real short rate and expected inflation.

The log of the real SDF is

$$m_{t+1} = -r_t - \frac{1}{2} \text{Var}_t(m_{t+1}) - \gamma B'_c \Sigma \epsilon_{t+1} - \gamma \eta_{c,t+1}.$$

The shock to the marginal utility of a unit of consumption moves inversely with the coefficient of relative risk aversion times the shock to consumption growth. The log of the nominal SDF is

$$m_{t+1}^{\$} = -r_t^{\$} - \frac{1}{2} \text{Var}_t(m_{t+1}^{\$}) - (\gamma B'_c + B'_\pi) \Sigma \epsilon_{t+1} - \gamma \eta_{c,t+1}. \quad (46)$$

The shock to the marginal utility of a dollar has a loading of one on the shock to the marginal utility of a unit of consumption and a loading of minus one on the inflation shock. This model fits into a standard Gaussian term structure setting. Real log bond prices satisfy

$$p_t^{(n)} = \mathbb{A}(n) + \mathbb{B}(n)' x_t,$$

where the loading of log prices on the state vector is

$$\mathbb{B}(n)' = -\gamma B'_c K (I - K)^{-1} (I - K^n). \quad (47)$$

The notation for nominal bond prices is similar. The loading of log nominal prices of nominal bonds on the state vector is

$$\mathbb{B}^{\$}(n)' = -(\gamma B'_c + B'_\pi) (I - K)^{-1} (I - K^n). \quad (48)$$

The weak form of the expectations hypothesis holds for both real and nominal bonds. A bond's yield moves one-for-one with changes in average expected short rates over the life of the bond. Average expected real short rates move one-for-one with γ times expected consumption growth over the life of the bond, and average expected nominal short rates move one-for-one with the sum of inflation and γ times expected consumption growth over

the life of the bond.

Given values of the utility parameters β and γ , both nominal and real term structure dynamics are determined by the joint dynamics of inflation and consumption growth. The relevant empirical properties of the data are examined after a discussion of a recursive utility extension to this power utility framework.

5.5 Recursive utility

The fear of stagflation may be the reason investors require a premium to hold nominal bonds. Informally, we can think of stagflation as a persistent regime characterized by high inflation and low economic activity. Piazzesi and Schneider (2007) formally motivate this mechanism in a recursive utility framework. Presumably the mechanism can also be generated by investors who have uncertainty, in a robust control or ambiguity sense, over the likelihood of a stagflation regime.

Piazzesi and Schneider (2007) adopt a representative agent model where the agent has recursive utility with a unit elasticity of intertemporal substitution (EIS). With recursive utility, the marginal value of a dollar today depends on expected log-run future consumption growth. This opens the door to a stagflation risk premium. Because the EIS is one, there is a closed-form expression for the real SDF. Continuing to denote the time rate of preference as β and the coefficient of relative risk aversion as γ , the log of the real SDF is

$$m_{t+1} = \log \beta - \Delta \tilde{c}_{t+1} - (\gamma - 1)\zeta_{t+1} - \frac{(\gamma - 1)^2}{2} \text{Var}_t(\zeta_{t+1}) \quad (49)$$

where ζ_{t+1} is news about future consumption. More precisely, it is the innovation in the expectation of the infinite sum of future consumption growth discounted at the time rate of preference. Formally,

$$\zeta_{t+1} = (E_{t+1} - E_t) \sum_{i=0}^{\infty} \beta^i \Delta c_{t+1+i}.$$

When the coefficient of relative risk aversion is one, this SDF collapses to the power utility SDF.

Again assume that the joint dynamics of consumption growth and inflation are described by state dynamics (1) and the mapping (43). Then the variations in real and nominal risk-free rates implied by this recursive utility framework are identical to the variations implied by power utility with a unit elasticity of intertemporal substitution. The only differences between the two utility frameworks are the constant terms in the risk-free rates and the shocks to real marginal utility. With power utility there is only the shock to immediate consumption growth. Recursive utility adds the shock to expected future consumption growth. The log of the real SDF can be written as

$$m_{t+1} = -r_t - \frac{1}{2}\text{Var}_t(m_{t+1}) - B'_c \Sigma \epsilon_{t+1} - (\gamma - 1)B'_c (I - \beta K)^{-1} \Sigma \epsilon_{t+1} - \gamma \eta_{c,t+1}$$

where the fourth term on the right does not appear with power utility.

5.6 The empirical performance of power and recursive utility

One way to take these models to data is to estimate the joint dynamics of per capita consumption growth and inflation, then infer term structure properties from the parameter estimates. To illustrate the stagflation risk premium, I estimate a linear Gaussian model with two factors and these two observables for the sample 1952Q2 through 1985Q4. The sample includes the stagflation of the 1970s and the Fed's reaction to the stagflation. The model's estimated parameters are then used to calculate model-implied covariances among shocks to consumption growth, inflation, and the discounted infinite sum of consumption growth. The discount rate is fixed at $\beta = 0.999$. Results are in Panel A of Table 9.

The columns of the table are the stochastic components of the real and nominal SDFs. The first two rows of the table are the building blocks of real and nominal risk-free rates: one-period-ahead expected consumption growth and inflation. The next two rows of the

table are building blocks of five-year real and nominal yields: average expected consumption growth and inflation over the next twenty quarters. In particular, nominal five-year bond yields move one-for-one with the sum of average expected consumption growth and average expected inflation.

With power utility, the signs of the covariances point to negative risk premia for long-term nominal bonds. The power utility nominal SDF (46) indicates that the marginal value of a dollar is unexpectedly low when either consumption growth or inflation are unexpectedly high. According to Table 9, both of these shocks positively covary with the sum of 20-quarter average expected consumption growth and 20-quarter average expected inflation. Therefore prices of five-year nominal bonds negatively covary with consumption and inflation shocks.

The stagflation story largely hinges on the covariance between shocks to long-run inflation and long-run consumption growth. The table reports that the covariance between average expected inflation over the next five years and discounted future consumption growth is negative. With recursive utility and with a sufficiently high coefficient of relative risk aversion, the effect of this negative covariance dominates the positive covariances between nominal yields and contemporaneous consumption growth and inflation. Fig. 2 displays unconditional mean real and nominal yield curves for the estimated parameters of the factor model, combined with a time rate of preference of 0.999 and a coefficient of relative risk aversion of 10. Panel B of Table 9 reports that the model-implied mean excess return to a five-year bond is about one percent a year.

This evidence is encouraging, at least at first glance. However, there are three problems with this stagflation-based explanation. One is shown in Fig. 2. The model implies a negatively-sloped real yield curve. As Piazzesi and Schneider note, when investors fear stagflation, real bonds are desirable assets. The U.S. empirical evidence suggests a positively sloped real yield curve, although U.K. evidence points in the other direction. Given the short sample of real yields, we will not worry much about this problem. A more important problem is that the data do not speak clearly about the importance of stagflation in the data. An

even more important problem is that nominal yields do not behave in the way that the recursive utility model implies they behave.

Based on the empirical exercise summarized in Table 9, the statistical reliability of the stagflation story is a little weak. For example, the asymptotic t -statistic for the mean excess return per quarter is about 1.7. However, this exercise ignores data after 1985. It also ignores any information that other observables, such as bond yields, may have about the dynamics of consumption growth and inflation. We should consider different data samples and expanded factor models to evaluate the robustness of these results.

Conveniently, we already have an additional set of parameter estimates in hand. Recall the first of the factor models estimated in Section 4.3. This model uses four factors to capture the joint dynamics of inflation, consumption growth, and two nominal bond yields. The estimated measurement equation is

$$\begin{pmatrix} \Delta\tilde{c}_t \\ \tilde{\pi}_t \\ \tilde{y}_t^{\$(1)} \\ \tilde{y}_t^{\$(20)} \end{pmatrix} = \begin{pmatrix} A_c \\ A_\pi \\ A^{\$(1)} \\ A^{\$(20)} \end{pmatrix} + \begin{pmatrix} B'_c \\ B'_\pi \\ B^{\$(1)'} \\ B^{\$(20)'} \end{pmatrix} x_t + \begin{pmatrix} \eta_{c,t} \\ \eta_{\pi,t} \\ \eta_t^{\$(1)} \\ \eta_t^{\$(20)} \end{pmatrix}. \quad (50)$$

This measurement equation overidentifies the recursive utility model. In that model, the $A^{\$}$ and $B^{\$}$ coefficients for the nominal yields are functions of the model's other parameters. For this empirical exercise we ignore the estimated values of $A^{\$}$ and $B^{\$}$ for the bonds to focus on the estimated joint dynamics of inflation and consumption growth. (Overidentifying implications are discussed below.) I also estimate a three-factor model with consumption growth, inflation, and the three-month yield as observables. Again, the measurement equation overidentifies the recursive utility model, and the overidentifying restrictions are ignored.

Table 10 reports model-implied risk premia on a five-year nominal bond for six estimated models. Two of the models use the same 1952Q2 through 1985Q4 sample used in Table 9. They differ from the model of Table 9 in the dimension of the state vector and the choice of

observables used to estimate the model. Three other models are estimated using the entire sample 1952Q2 through 2010Q4. This includes the financial crisis period, for which nominal bond risk premia may be reversed. Nominal bonds hedge the risk that the central bank cannot push interest rates off the lower bound of zero. Finally, a four-factor model using four observables is estimated over the homogeneous data sample of 1986Q1 through 2007Q4.

For each of the estimated models, the statistical reliability of positive risk premia is substantially weaker than reported in Table 9. (Of course, this is not an accident. I chose the estimated model underlying Table 9 because it puts the best face on the stagflation story.) None of the risk premia has an asymptotic t -statistic greater than 0.9, and two of the six estimates are negative. There is simply not enough information in the estimated joint dynamics of consumption growth and inflation to support the stagflation interpretation of risk premia.

The other critical problem with this consumption-based explanation of risk premia is that the actual dynamics of bond yields are not close to model-implied dynamics of bond yields. According to the model, the real risk-free rate moves one-for-one with expected growth in aggregate consumption. A long line of research rejects this description. Empirically, there is substantial variation over time in the real risk-free rate (as defined by the Fisher equation), but its covariance with expected consumption growth is close to zero.³ Thus the elasticity of intertemporal substitution with respect to aggregate consumption appears to be much less than one. For the purposes of this discussion, a more important conclusion is that much of the variation in the real risk-free rate is orthogonal to expected consumption growth.

The bond risk premia calculated in Tables 9 and 10 are based on the assumption that long-term yields vary one-for-one with expected average future short rates, which in turn are mechanically determined by expected inflation and consumption growth. One way to test these restrictions is with the overidentifying restrictions of the measurement equation (50). In other words, are the estimated unrestricted mappings from the state vector to the yields

³Campbell (2003) summarizes the literature that estimates the elasticity of intertemporal substitution. A more recent discussion is in Beeler and Campbell (2011).

equal to the model-implied mappings?

Such a test implicitly requires that investors know the true dynamics of consumption growth and inflation. Piazzesi and Schneider (2007) note that bond yields will not vary in the way their model predicts unless investors understand the joint dynamic properties of inflation and consumption growth. It takes time for investors to infer these dynamics. The following exercise is designed to informally check the overidentifying restrictions of the model while taking into account the possibility of learning.

Begin with the model summarized in Table 9, which uses two factors to describe the joint dynamics of consumption growth and inflation over the sample 1952Q2 through 1985Q4. Given the model's parameters, construct Kalman-smoothed fitted values of shocks to consumption growth, inflation, and the discounted sum of expected future consumption growth. Do this for the entire sample 1952Q2 through 2010Q4. Then compute contemporaneous covariances between these fitted shocks and changes in nominal bond yields. If we assume that investors have always known that the estimated model is the true model, the full-sample covariances should line up with the model-implied covariances in Table 9. If investors are learning during the period, the fitted covariances for the post-1985 sample should line up with the model-implied covariances estimated for the pre-1985 sample.

Results are in Panel A of Table 11. The key covariances are those between bond yields and shocks to the discounted sum of future consumption growth. The covariances are not particularly sensitive to the sample period. They are typically about one-tenth the magnitude of the model-implied covariances in Table 9.⁴ In other words, long-term bond yields are much less responsive to news about future consumption growth than the model implies. Similar results are in Panel B of Table 11, which uses a three-factor model and includes the nominal short rate among the observables. For this model, long-term bond yields actually positively covary with news about future consumption growth.

One possible interpretation of this evidence is that investors have recursive utility pref-

⁴For Table 9, add the covariances involving expected future consumption growth and expected future inflation.

erences described by this model, but their beliefs about the likelihood of stagflation differ substantially from the likelihood implied by the estimated joint dynamics. Alternatively, investors do not have the preferences described here. In either case, the evidence in Table 11 shows that even if an investor is concerned about persistently low consumption growth, long-term nominal yields are not particularly risky assets. Thus we must venture beyond recursive utility if we want to explain an upward-sloped nominal yield curve. But before considering alternate models, we study some additional empirical properties of expected excess returns.

5.7 Predictable variation of excess bond returns

Tests of the expectations hypothesis document conclusively that conditional expectations of excess returns to Treasury bonds vary with the shape of the term structure. Excess bond returns can be predicted with spreads, including both the spread between forward rates and short-maturity yields as in Fama and Bliss (1987) and yield spreads as in Campbell and Shiller (1991).

Campbell (1987) notes that spreads are powerful instruments for predicting excess returns because changes in expected excess returns to long-term bonds are automatically compounded in the prices of these bonds, and thus in the spread between long-term and short-term bond yields. Put differently, there is an accounting relation linking expected excess returns to forward rates. Yet the same accounting relation that makes spreads powerful instruments also makes them, in a sense, uninformative. Variations in expected excess returns can be detected with spreads regardless of the reasons for the variation, hence this evidence says nothing about the underlying determinants of the variations.

Beginning with Kessel (1965) and Van Horne (1965), economists have proposed various theories of time-varying expected excess bond returns. Naturally, many theories imply that this variation is correlated with the state of the economy, some of which are discussed in Sections 5.8 and 5.9. In this subsection we look at some empirical evidence to help us

evaluate the theories.

For this exercise, monthly excess returns to nominal Treasury bonds are defined as the simple return to a portfolio of bonds with maturities between five and ten years less the return to a portfolio with maturities under six months. The data are from CRSP. The return from the end of month t to the end of month $t + 1$ is regressed on a variety of variables dated t or earlier. The focus on monthly returns is somewhat at odds with much of the recent literature that follows Cochrane and Piazzesi (2005) by studying annual returns. But the goal here is not to generate large R^2 s, but rather to judge the statistical significance of predictability. When necessary, we follow the intuition of Hodrick (1992) by predicting monthly returns with annual averages of predictors rather than predicting annual returns with month- t values of the predictors.

Panel A of Table 12 reports the results of a variety of forecasting regressions. Over the sample period 1952 through 2010, the statistical evidence for predictability using information in the term structure is overwhelming. The slope of the term structure, measured by the five-year yield less the three-month yield, is positively correlated with future excess returns. As first documented by Cochrane and Piazzesi (2005), forward rates constructed with yields on bonds with maturities of one through five years also forecast excess returns.

Although excess returns are clearly predictable, a couple of obvious choices of macroeconomic determinants of risk premia have no predictive power. A measure of the conditional standard deviation of long-term bond yields is constructed using squared daily changes in yields. This measure has no ability to forecast excess bond returns; the p -value is about 0.5. (Owing to the availability of daily observations of long-term yields, the sample period for this regression begins with 1962.) Many theories imply that risk premia are countercyclical. Yet over the period 1952 through 2010, lagged changes in log industrial production have no ability to forecast excess returns, whether the changes are from $t - 1$ to t or from $t - 12$ to t .

Ludvigson and Ng (2009) and Cooper and Priestly (2009) appear to be more successful at uncovering predictability with measures of macroeconomic activity. In particular, Ludvigson

and Ng (2009, 2010), construct principal components of more than 130 macroeconomic and financial time series. They use the first eight principal components to predict excess bond returns. The principal components studied in Ludvigson and Ng (2010) are available for 1964 through 2007. Table 12 reports that when these eight principal components are used to forecast monthly excess returns, the adjusted R^2 is more than eight percent. The hypothesis that the coefficients are jointly zero is rejected at any conceivable confidence level.

Although this evidence is certainly encouraging to researchers searching for links between risk premia and the macroeconomy, there are three important caveats. First, it is not clear how to map much of this predictability to specific macroeconomic concepts. Ludvigson and Ng call the first principal component “real activity” because it is highly correlated with standard measures of real activity. For example, its correlation with log-differenced industrial production exceeds 0.8. But relatively little of the forecast power of the principal components comes from the measure of real activity. Table 12 reports that the adjusted R^2 using only the first principal component is about 1.6 percent. Interpreting the other principal components is more problematic.

Second, the properties of the sample period 1964 through 2007 may be unusual. Table 12 uses this shorter sample to redo the forecasting regressions with five forward rates and with industrial production growth. With the shorter sample, the adjusted R^2 for the five forward rates more than doubles. The p -value for growth in industrial production drops from 0.5 to 0.02. Since the first principal component is highly correlated with the growth in industrial production, this result suggests that the predictability associated with Ludvigson and Ng’s real activity factor may be sample-specific.

Third, the predictability of excess bond returns is not accompanied by similar predictability of excess stock returns. In models that link time-varying bond risk premia to the macroeconomy, this variation is typically accompanied by varying equity risk premia. But Panel B of Table 12 reports that the shape of the term structure contains little information about future aggregate equity excess returns. The eight principal components of macroeconomic

and financial variables collectively have substantial forecasting power, but the coefficients on the principal components do not line up with those that forecast bond returns. The correlation between the fitted values of the regressions is only 0.39. For the aggregate stock excess return, the coefficient on real activity is statistically indistinguishable from zero.

5.8 Extensions to power utility and recursive utility

The subsection briefly two extensions to the power utility and recursive utility approaches discussed above. They both depart from the Gaussian setting, and require either approximations or numerical methods to solve for bond prices.

Wachter (2006) argues that habit formation can explain both a nominal yield curve that slopes up on average and time-varying expected excess returns to nominal bonds. In the habit formation model of Campbell and Cochrane (1999), surplus consumption has two effects on the risk-free real rate. The first is an intertemporal smoothing effect induced by mean-reverting surplus. If, say, surplus is below its mean, investors expect it to increase over time. Thus they expect the marginal value of a unit of consumption to be lower in the future than today, which pushes up the real rate. Offsetting this is the second effect of precautionary savings. If surplus is below its mean, the conditional variance of real marginal utility is high, and thus the desire for precautionary savings is high. This pushes down the real rate. Campbell and Cochrane simplify their model by equating these effects, producing a constant real rate.

Wachter generalizes this approach such that the sum of the two effects is unsigned. In her model, the real rate has an extra term of a constant times surplus consumption. She then uses numerical methods to solve for real bond prices. Finally, she adds an exogenous inflation process and solves for nominal bond prices.

To take her model to the data, Wachter constructs a proxy for surplus consumption equal to an exponentially declining sum of the previous 40 quarters of inflation. She estimates a negative relation between real rates and this proxy. In her model, this implies that real bonds

command a positive risk premium. When consumption, and thus surplus, unexpectedly fall, real rates rise and real bond prices fall. Her estimated inflation process implies that current and expected inflation are negatively correlated with consumption growth, hence nominal bonds command a higher risk premium than real bonds. In addition, both real and nominal bonds have time-varying risk premia that are higher when surplus is relatively low.

Here we revisit some of Wachter’s empirical analysis. Define a proxy for surplus consumption as

$$\hat{s}_t \equiv \sum_{i=1}^{40} \phi^{i-1} \Delta c_{t-i+1}$$

where $\phi = 0.97$, following Wachter.⁵ The first available observation is 1962Q2. Following Wachter, regress the ex-post real return to a one-quarter bond on surplus consumption, or

$$r_t^{\$} - \pi_{t+1} = b_{r,0} + b_{r,1} \hat{s}_t + e_{r,t+1}.$$

Also use this proxy to test the implication of her model that expected excess returns are negatively related to surplus. Quarterly excess returns to a portfolio of Treasury are also regressed on surplus. The bonds have maturities between five and ten years.

$$\text{excess bond return}_{t+1} = b_{b,0} + b_{b,1} \hat{s}_t + e_{b,t+1}.$$

These regressions are estimated over the sample 1962Q2 through 2004Q4, which is similar to the one used by Wachter, and over the longer sample 1962Q2 through 2010Q4. Results are in Table 13. Wachter uses aggregate consumption growth to construct her proxy. Here, results are presented for both aggregate consumption growth and per capita consumption growth.

Using Wachter’s proxy and the shorter sample period, the ex-post real rate regression results are close to those reported by Wachter. High surplus corresponds to a low real

⁵This definition differs by a lag from the measure of Wachter, who begins the sum with the first lag. In practice, this is irrelevant.

rate, and the evidence appears statistically significant. But over the full sample, the result is much weaker, both economically and statistically. The hypothesis that the coefficient is zero cannot be rejected. Moreover, when per capita consumption is used to construct the surplus consumption proxy, the evidence for a real-rate channel disappears even in the shorter sample. Finally, there is at best weak evidence that surplus predicts excess bond returns. Only one of the reported t statistics is significant at the ten percent level, and none are significant at the five percent level.

There are two other empirical problems with her story. First, it implies that shocks to aggregate consumption are negatively correlated with long-term nominal yields, which is counterfactual (Table 7). Second, it implies that risk premia on stocks and bonds move together, which is also counterfactual (Table 12).

An alternative theoretical approach is taken by Bansal and Shaliastovich (2010). They rely on stochastic volatility in the long-run risks model of Bansal and Yaron (2004). A state variable drives the conditional standard deviation of expected consumption growth. Expected inflation loads on expected consumption growth, so the conditional standard deviation of expected inflation also varies through time. This approach adds considerable flexibility to a term structure model. It breaks the tight link between conditional expectations of consumption growth and real risk-free rates. It also adds another state variable to the determinants of expected excess real and nominal bond returns.

Like Wachter's model, this model has strong, testable implications. Since conditional volatilities affect real rates, it implies that the shape of the real yield curve should be valuable conditioning information in constructing conditional standard deviations of expected consumption growth and expected inflation. It also implies that conditional standard deviations of real yields should forecast excess returns to real and nominal bonds. To date, neither of these implications has been confronted with the data.

Their model also implies that conditional standard deviations of nominal yields should forecast excess returns to real and nominal bonds. From Table 12 we know this implication

is counterfactual, at least for nominal bonds. However, as Bansal and Shaliastovich note, this implication is primarily a matter of modeling convenience. The implication need not carry over to a richer model.

Bansal and Shaliastovich’s model has one important advantage over all of the other models discussed in this section: it has not been shown to be inconsistent with the data. But given the string of failures that precedes it, it would be imprudent to give it much credence until some of its strong overidentifying restrictions are tested. With high-frequency data to construct estimates of conditional standard deviations, it should be possible to construct powerful tests of the model’s joint implications for conditional standard deviations, real yields, and expected excess returns.

5.9 Moving away from endogenous risk premia

Both Wachter (2006) and Bansal and Shaliastovich (2010) add a state variable to simple power utility and recursive utility frameworks. These state variables are grounded in the behavior of aggregate consumption, hence their models continue to tie risk premia tightly to the dynamic behavior of aggregate consumption. Alternative approaches weaken this link. Two of them are briefly discussed in this subsection.

Gallmeyer, Hollifield, Palomino, and Zin (2008) generalize the habit formation setting of Campbell and Cochrane (1999) by replacing surplus consumption in the utility function with an exogenous state variable, which we can call a “price of risk” state variable. Like surplus consumption, it affects asset prices by magnifying the price of consumption risk. They avoid some modeling complications by specifying the process to be a martingale, thus avoiding an intertemporal smoothing effect in the risk-free real rate. The only effect of the state variable on the real risk-free rate is to vary the precautionary savings motive over time.

The nominal term structure also depends on this variable through its effect on the real risk-free rate. The model adds another channel through which time-varying risk premia can affect the nominal term structure. A monetary authority that follows a Taylor rule will

induce a correlation between shocks to inflation and shocks to the price of risk, generating an additional source of risk premia for nominal bonds.

Like Gallmeyer, Hollifield, Palomino, and Zin (2008), Lettau and Wachter (2010) introduce an exogenous “price of risk” state variable. But they take one more step away from a consumption-based model by specifying an exogenous stochastic process for the real risk-free rate. These models avoid many of the counterfactual implications of the models discussed earlier in this section by simply not taking a testable stand on the magnitude of risk premia or the source(s) of their variation through time. Nonetheless, they share two testable implications. First, only aggregate consumption risk is priced.⁶ Second, this risk must be priced consistently across stocks and nominal bonds; if the price of risk rises, risk premia on both sets of assets must increase.

These two implications are inconsistent with the results of Tables 8 and 12. Table 8 tells us that the consumption beta of nominal Treasury bonds is modestly negative. Table 12 tells us that the variables that predict excess nominal bond returns, such as the slope of the term structure and average forward rates, are not significant predictors of excess stock market returns.⁷

The conclusion to draw from this discussion is that these models of exogenous risk premia, like the models of endogenous premia discussed previously in this section, are unsuccessful at explaining basic features of returns to nominal bonds. Whether this is good or bad news depends on your perspective. For example, new PhDs can be confident that some of the most important questions concerning the joint behavior of the term structure and the macroeconomy remain unanswered.

⁶For Lettau and Wachter, the priced risk is aggregate dividend risk because consumption does not appear in their model.

⁷Lettau and Wachter regress excess stock returns on the slope of the term structure. For a five-year holding period (but not a one-year holding period), they uncover a statistically significant positive relation. However, with only ten non-overlapping observations, the use of asymptotic standard errors is difficult to justify.

6 New Keynesian models

Dynamic models play a central role in macroeconomic policy and research. However, models used by macroeconomists have little in common with those discussed in Section 5. The previous section focused on restrictions based on attitudes towards risk. Dynamic macroeconomic models typically emphasize restrictions derived from New Keynesian logic.

In principle, New Keynesian models help solve two of the most important problems with macro-finance dynamic term structure models. First, they impose restrictions on the joint dynamics of the macroeconomic variables, for which no-arbitrage restrictions are meaningless. Second, they replace the unspecified, arbitrary “factors” of Section 2.3 with economically meaningful concepts.

Given these advantages, it is not surprising that research into New Keynesian term structure models is expanding rapidly. As of this writing, the models do not stack up well empirically, in the sense that many of the New Keynesian restrictions appear to be at odds with the data. This section uses a particular model to illustrate some of the empirical challenges that future research may well overcome.

6.1 A reduced-form New Keynesian model

This model is a simplified version of the model of Bekaert, Cho, and Moreno (2010). Similar models appear in Hördahl, Tristani, and Vestin (2006) and Rudebusch and Wu (2008). New Keynesian models have optimizing agents and firms in an economy in which prices are sticky. They also have a monetary authority that conducts monetary policy. A key component to monetary policy is the monetary authority’s long-run target for inflation. This target is one of the determinants of the current inflation. Presumably, it is also an important determinant of long-term bond yields. Thus a New Keynesian model puts an intuitive label on an otherwise latent factor in a macro-finance model.

The equations are equilibrium conditions. Aggregate demand is described by the IS

equation

$$g_t = \alpha_g + \beta_g E_t g_{t+1} + (1 - \beta_g) g_{t-1} - \gamma_r (r_t^{\$} - E_t(\pi_{t+1})) + \zeta_{g,t}, \quad \zeta_{g,t} \sim N(0, \sigma_t^2). \quad (51)$$

In (51) g_t is either the output gap or a variable perfectly correlated with the output gap. An intertemporal savings/consumption choice suggests that $\gamma_r > 0$, as higher real rates encourage saving. When β_g is greater than zero, aggregate output growth is forward-looking, which can be justified by habit formation. The output shock $\zeta_{g,t}$ is independent of all other shocks in the model.

Inflation is determined by the AS equation

$$\pi_t = \beta_\pi E_t \pi_{t+1} + (1 - \beta_\pi) \pi_{t-1} + \gamma_\pi g_t + \zeta_{\pi,t}, \quad \zeta_{\pi,t} \sim N(0, \sigma_\pi^2). \quad (52)$$

In the AS equation, inflation depends on the output gap through marginal cost pricing. Inflation depends on expected future inflation because firms recognize they adjust prices infrequently. The inflation shock is independent of all other shocks in the model.

The monetary authority follows a forward-looking Taylor rule

$$r_t^{\$} = \alpha_r + \beta_r r_{t-1}^{\$} + (1 - \beta_r) [\gamma_{r\pi} (E_t \pi_{t+1} - \pi_t^L) + \gamma_{rg} g_t] + \zeta_{r,t}, \quad \zeta_{r,t} \sim N(0, \sigma_r^2). \quad (53)$$

With $\gamma_{r\pi}$ and γ_{rg} both greater than zero, the central bank raises nominal rates when output increases and when expected future inflation exceeds the central bank's long-run inflation target π_t^L . The monetary policy shock is independent of all other shocks in the model.

The dynamics of the inflation target are necessary to complete the model. One plausible approach is to derive them from the central bank's objective function. However, the reason to embed this New Keynesian model into a dynamic term structure setting is to use the behavior of long-term bond yields to infer information about long-run inflation. Therefore

this literature uses a variety of ad hoc specifications that all imply high persistence of the long-run inflation target. For example, Hördahl, Tristani, and Vestin (2006) use a simple AR(1). Bekaert, Cho, and Moreno (2010) motivate informally the process

$$\pi_t^L = \beta_{L1} E_t \pi_{t+1}^L + \beta_{L2} \pi_{t-1}^L + (1 - \beta_{L1} - \beta_{L2}) \pi_t + \zeta_{L,t}, \quad \zeta_{L,t} \sim N(0, \sigma_L^2). \quad (54)$$

The shock to the inflation target is independent of all other shocks in the model.

A quick examination of these equations reveals that the unconditional mean of inflation equals the unconditional mean of the long-run inflation target. Neither this mean, nor the mean of the short-term nominal rate, is determined by these equations. The unconditional mean of g_t is determined by the free parameter α_g . To complete the model we add the unconditional means of inflation and the short-term nominal rate as free parameters.

Armed with these New Keynesian equations, the next step is to map the model into the macro-finance framework. The equations fit into the linear framework

$$\mathcal{A}x_t = \alpha + \mathcal{B}E_t x_{t+1} + \mathcal{J}x_{t-1} + \mathcal{C}\zeta_t, \quad (55)$$

where ζ_t is a vector that stacks the shocks to (51), (52), (53), and (54). the state vector is

$$x_t = \begin{pmatrix} g_t & \pi_t & r_t^{\$} & \pi_t^L \end{pmatrix}'. \quad (56)$$

The state vector has three observable factors (up to measurement error) and one latent factor.

Under the assumption that the expectations in (55) are rational, the dynamic solution to (55) takes the form

$$\begin{aligned} x_{t+1} &= \mu(\Theta) + K(\Theta)x_t + \Sigma(\Theta)\epsilon_{t+1}, \\ \epsilon_{t+1} &= \epsilon(\Theta, \zeta_{t+1}), \quad \epsilon_{t+1} \sim MVN(0, I). \end{aligned} \quad (57)$$

The vector Θ contains the parameters of the model, including the unconditional means of inflation and the short-term nominal rate. In general, analytic solutions do not exist. The literature contains a variety of numerical techniques to solve systems of equations such as these. The choice of technique is typically dictated by the specifics of the model. Bekaert, Cho, and Moreno (2010) describe how to solve (55) in the form (57).⁸

6.2 Nesting the model in a general factor structure

Assume that we observe the nominal short rate, inflation, and the output gap. We can write this New Keynesian model in a form suitable for the Kalman filter. The measurement equation is

$$\begin{pmatrix} \tilde{g}_t \\ \tilde{\pi}_t \\ \tilde{r}_t^{\$} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} x_t + \eta_t, \quad \eta_t \sim N(0, \Omega). \quad (58)$$

The transition equation is (57).

This four-factor New Keynesian model is a restricted version of a general four-factor model, which we can write using the same measurement equation (58) and with a transition equation

$$x_{t+1} = \begin{pmatrix} \underbrace{\mu_1}_{3 \times 1} \\ \underbrace{\mu_2}_{1 \times 1} \end{pmatrix} + \begin{pmatrix} \underbrace{K_{11}}_{3 \times 3} & \underbrace{K_{12}}_{3 \times 1} \\ \underbrace{K_{21}}_{1 \times 3} & \underbrace{K_{22}}_{1 \times 1} \end{pmatrix} x_t + \begin{pmatrix} \underbrace{\Sigma_{11}}_{3 \times 3} & \underbrace{0}_{3 \times 1} \\ \underbrace{\Sigma_{21}}_{1 \times 3} & \underbrace{\Sigma_{22}}_{1 \times 1} \end{pmatrix} \epsilon_{t+1}, \quad (59)$$

with Σ_{11} lower triangular. In this general factor model without New Keynesian restrictions, the latent factor is (arbitrarily) identified with the restrictions

$$\mu_2 = 0, \quad K_{21} = 0, \quad \Sigma_{22} = 1. \quad (60)$$

⁸Thanks to Seonghoon Cho for sharing their code.

These restrictions identify respectively a translation, rotation, and scaling of the latent factor.

There are two differences between the transition equations (57) and (59). The most obvious is that the New Keynesian version is more restrictive. The general transition equation (59) has 25 free parameters. The New Keynesian transition equation (57) has only 16 free parameters.

The other difference is that the general transition equation has a factor without a label. Equation (59) defines the factor by normalizing its dynamics. The factor has an unconditional mean of zero, it follows a univariate AR(1) process, and the component of its innovation that is orthogonal to the innovations of the other factors has a unit standard deviation. Because K_{12} is nonzero, the latent factor may contain information about future values of the output gap, the short rate, and inflation that is not contained in the current values of these other factors. These restrictions uniquely identify the properties of the factor but do not correspond to an intuitive economic concept.

The latent factor of the New Keynesian model has a label: the long-run inflation target. But, what, precisely, is a long-run inflation target? In other words, what is the economic content of the label? At a minimum, such a label suggests that our best guess of inflation at some long-distant date equals our best guess of the inflation target at the long-distant date, or

$$\lim_{\tau \rightarrow \infty} E_t (\pi_{t+\tau} - \pi_{t+\tau}^L) = 0.$$

This restriction is satisfied in the model, but it has almost no content. It is satisfied by any two arbitrary stationary random variables as long as they have the same mean.

There are other plausible restrictions associated with the label of long-run inflation target. A natural restriction is that a high target relative to current inflation should predict an increase in inflation. Formally,

$$\text{Cov} (E_t (\pi_{t+j}) - \pi_t, \pi_t^L - \pi_t) > 0, \quad \forall j > 0. \tag{61}$$

Similarly, a positive shock to the fundamental inflation-target shock $\zeta_{L,t}$ should raise expectations of future inflation, or

$$E(\pi_{t+j} \mid \zeta_{L,t} > 0) - E(\pi_{t+j}) > 0, \quad \forall j > 0. \quad (62)$$

Part of the variance of inflation is attributable to variations in the target. Therefore another plausible restriction is

$$\text{Var}(\pi_t - \pi_t^L) < \text{Var}(\pi_t). \quad (63)$$

Similarly, the intuition of a long-run inflation target suggests that it is less volatile than actual inflation. The restriction is

$$\text{Var}(\pi_t^L) < \text{Var}(\pi_t). \quad (64)$$

This New Keynesian model can be parameterized to satisfy restrictions (61) through (64). However, there are also parameter choices for which they do not hold. Even if all of the parameters of the IS, AS, and Taylor-rule equations have signs consistent with New Keynesian logic, each of (61) through (64) can be violated. In a nutshell, the label “long-run inflation target” as used in this model has no specific falsifiable implications. This statement does not mean the factor π_t^L is unidentified. The New Keynesian model pins it down, in the sense that there are no translations, rotations, or scaling transformations that can be applied to π_t^L without changing the log-likelihood of the model. Instead, the point here is that the identification implied by the model does not inherently correspond to any economic notion of a long-run inflation target.

Of course, this New Keynesian model could be augmented with additional restrictions to force the factor π_t^L to behave in a way consistent with our intuition about an inflation target. An alternative approach is to estimate the model without any additional restrictions, then check whether (61) through (64) are satisfied. A more informal check on the adequacy of

the label is to look at filtered estimates of π_t^L to see if the time series is plausibly a long-run inflation target.

6.3 Adding nominal bonds

Both the New Keynesian model and its general factor counterpart can be estimated using a panel of observations on economic activity, inflation, and the nominal short rate. The next subsection presents parameter estimates for both models using such a sample. Since there are four factors and only three observables, one of the factors cannot be inferred from the cross-section of observables. It is filtered out of their dynamics. However, the intuition of the New Keynesian model suggests that the long-run inflation target will be impounded into yields of long-term nominal bonds. Adding such yields to the estimation should both allow the factor to be inferred from the cross-section and improve forecasts of inflation.

There are three ways to incorporate long-term yields into this framework. The first method follows Section 2.1, in which no-arbitrage restrictions are ignored. Yields are simply added to the measurement equation (58) with unrestricted constant terms and loadings on the state vector. The empirical analysis in the next subsection adopts this approach. The second method follows Section 3 by imposing the law of one price. The sensitivity of the real SDF to shocks is an arbitrary dynamic functions of the state vector, parameterized as in (18). Ad hoc restrictions can be imposed on (18). Hördahl, Tristani, and Vestin (2006) and Rudebusch and Wu (2008) are examples of this approach in the New Keynesian term structure literature.

The third method is to take the microfoundations of New Keynesian models very seriously, using them to determine risk premia. However, two problems limit the practical applicability of this approach. The first is tractability. The convenient linear structure in (51) through (54) is a first-order approximation to nonlinear dynamic stochastic general equilibrium (DSGE) model around the nonstochastic solution. This approximation eliminates any role for risk aversion, since risk aversion depends on shocks that alter the marginal

utility of consumption. Put differently, the existence of bond risk premia is inconsistent with (51) through (54). A DSGE model with nontrivial risk premia requires a second-order approximation as in Hördahl, Tristani, and Vestin (2008). Time-varying risk premia requires a third-order approximation, as in Rudebusch and Swanson (2012), van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez (2010), and Andreasen (2011).⁹

One workaround is to ignore the fact that these equations are derived through a log-linearization. Bekaert, Cho, and Moreno (2010) treats them as exact solutions and takes the properties of risk aversion from the IS equation (51). The parameters of (51) are determined by the representative agent's desire to smooth consumption over time. With power utility the same parameters determine risk aversion. The result is a restricted form of risk compensation (18).

The second problem is empirical. Section 5 finds that bond yields do not behave in the way that consumption-based models imply they should behave. Once this discrepancy is recognized, there is little support for models of risk premia based on aggregate consumption. New Keynesian models fall squarely into this model class. Accordingly, the empirical analysis in the next subsection does not attempt to reconcile New Keynesian risk premia with expected excess returns to Treasury bonds. Instead, it asks whether the model's dynamic restrictions are consistent with the data.

6.4 An empirical application

This subsection applies the New Keynesian and general factor models to quarterly U.S. data. The output gap is proxied by detrended log industrial production. Table 2 lists various other possible proxies for the output gap. The qualitative conclusions of this exercise do not hinge on the choice. Inflation is measured using the CPI and the short-term nominal rate is the three-month Treasury bill yield. The sample period is 1960Q1 through 2010Q4. The pre-1960 sample is excluded because of the erratic behavior of industrial production during the

⁹Time-varying risk premia can also be generated with a second-order approximation of a regime-switching model, as in Amisano and Tristani (2009).

1950s.

Seven different combinations of models and data are examined. The simplest version is a general three-factor model, estimated using only data on the nominal short rate, inflation, and the output gap proxy. Since this model has only three factors, it does not nest the New Keynesian model. The same data are then used to estimate a general four-factor model. Two other versions of the general four-factor model are estimated, which differ in the observed data. One adds the yield on a five-year bond, while the other adds yields on bonds with maturities of one through five years.

Estimated features of these four combinations are contrasted with three versions of the New Keynesian model of Section 6.1. The first uses only data on the nominal short rate, inflation, and the output gap. The second adds a five-year bond yield and the third adds yields on bonds with maturities of one through five years. Table 14 reports cross-sectional and one-quarter-ahead forecast root mean squared errors for all estimated versions. Before discussing the results in detail, note that the cross-sectional fitting error for the output gap proxy is zero (to three decimal places) for all models. Recall that the factor models estimated in Section 4.3 are characterized by substantial cross-sectional fitting errors for consumption growth. The difference in these results is driven by the high persistence of the output gap measure relative to that of consumption growth. Maximum likelihood attributes only a tiny portion of the output gap's variation to a white-noise component.

The estimated general three-factor is a useful benchmark from which to evaluate the other models. Relative to this model, the main effect of adding another factor is to improve substantially the forecast of the output gap. The one-quarter-ahead RMSE falls from 6.6 percentage points to 5.9 percentage points. This is less interesting than it appears. The three-factor model underpredicts industrial production in the first half of the sample and overpredicts it in the second half. The fourth factor in the four-factor model helps reduce the serial correlation in the forecast, thus raising forecasts in the first half and lowering them in the second.

Imposing the New Keynesian restrictions on this four-factor model with three observables lowers the log-likelihood by about 13.9. A log-likelihood test rejects the hypothesis that the restrictions are correct at the 0.1 percent level. (The log-likelihoods are not reported in any table.) From an economic perspective, the deterioration in fit is not large. The root mean squared cross-sectional and forecasting errors are not substantially different across these two four-factor models. The largest difference is for the output gap forecast. The RMSE rises from 5.9 to 6.1 percentage points.

We now add a five-year bond yield to the observables that are to be explained by the model. For the general factor model, the most noticeable effect is the increase in the forecast RMSE for the output gap. Maximum likelihood is no longer able to use the fourth factor to pick up serial correlation in output gap forecast errors. Instead, it must fit the five-year yield. Therefore the RMSE is almost identical to the RMSE for the three-factor, three observables model.

In a statistical sense, the New Keynesian counterpart to this model is successful. The likelihood ratio test cannot reject the restrictions at the 10 percent level. However, the economic deterioration associated with the restrictions is precisely along the dimension that the model should be successful. Both the cross-sectional and forecasting errors for inflation are substantially larger with the New Keynesian restrictions than without them. The cross-sectional RMSE rises from 1.8 to 2.4 percentage points and the forecasting RMSE rises from 2.3 to 2.6 percentage points. Using this New Keynesian model to extract a long-run inflation factor from the five-year yield appears to be counterproductive, both relative to a less-restrictive model and a model that imposes the restrictions but ignores information in long-maturity yields.

Finally, we add another four bond yields as observables. Thus these four-factor models must now explain the behavior of the short rate, inflation, detrended log industrial production, and five longer-maturity bond yields. Given the evidence in Section 4 that four factors cannot capture the joint dynamics of nominal yields, inflation, and the term struc-

ture, the results in Table 14 are unsurprising. Cross-sectional and forecasting errors are higher when the additional yields are included. Statistically, the New Keynesian restrictions are overwhelmingly rejected. The difference in log-likelihoods exceeds 60.

What are the properties of the factor labeled as the long-run inflation target? Fig. 3 displays filtered estimates of the factor for the model estimated using the five-year bond yield (Panel A) and the model estimated using one-year through five-year bond yields (Panel B). There is no way to reconcile the time series in Panel B with this label. The factor ranges from -10 percent to almost 40 percent over the sample period. The range of the time series in Panel A is not as ridiculous. However, the parameter estimates for this model violate the restrictions (61) and (62). Maximum likelihood uses this factor to fit the five-year bond yield, not to help predict future inflation. In many of our models, this distinction is nonsensical. But in the data, it is critical.

7 Concluding comments

The idea that Treasury bond prices should be determined primarily by the macroeconomy—current and expected inflation, output, and consumption—is grounded in both casual economic intuition and state-of-the-art models. It is such an obvious concept that there is a danger we may take it for granted. Imagine that future advances in dynamic macro theory produce a model that generates an upward-sloped nominal yield curve on average, implies that excess bond returns vary predictably with the shape of the term structure, and decouples bond return predictability from stock return predictability. An implicit message of this chapter is that we should resist the strong temptation to conclude the model is successful.

Instead, we must take a close, skeptical look at the mechanisms that drive these results. The example of the fear of stagflation is instructive. The model reproduces observed features of the nominal term structure, but does so through a sequence of critical logical steps. If expected future nominal short rates do not have a large negative covariance with news about

future long-horizon consumption growth, or if nominal long-term yields do not closely track expected future short rates, then nominal bond prices do not necessarily have a large negative covariance with this news. If not (and in the data, they do not), support for the model's conclusions disappears.

We can afford to treat macro-finance models skeptically for the same reason that we are tempted to believe them: bond yields and the macroeconomy must somehow be closely connected. There is almost certainly a macroeconomic model that reproduces the behavior of nominal bond yields through mechanisms that withstand close scrutiny. We just haven't discovered it yet.

References

- Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* 15, 243-288.
- Amisano, Gianni, and Oreste Tristani, 2009, A DSGE model of the term structure with regime shifts, Working paper, ECB.
- Andreasen, Martin M., 2011, Explaining macroeconomic and term structure dynamics jointly in a non-linear DSGE model, Working paper, Bank of England.
- Ang, Andrew, Geert Bekaert, and Min Wei, 2008, The term structure of real rates and expected inflation, *Journal of Finance* 63, 797-849.
- 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.
- Ang, Andrew, Sen Dong, and Monika Piazzesi, 2007, No-arbitrage Taylor Rules, Working paper, Columbia University.
- Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680-692.
- Backus, David K., Allan W. Gregory, and Stanley E. Zin, 1989, Risk premiums in the term structure: Evidence from artificial economies, *Journal of Monetary Economics* 24, 371-399.
- Backus, David K., and Stanley E. Zin, 1994, Reverse engineering the yield curve, NBER Working paper 4676.
- Bansal, Ravi, and Ivan Shaliastovich, 2010, A long-run risks explanation of predictability in bond and currency markets, Working paper, Duke University.

- Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: a potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481-1509.
- Beeler, Jason, and John Y. Campbell, 2011, The long-run risks model and aggregate asset prices: an empirical assessment, Working paper, Harvard University.
- Bekaert, Geert, Seonghoon Cho, and Antonio Moreno, 2010, New Keynesian macroeconomics and the term structure, *Journal of Money, Credit, and Banking* 42, 33-62.
- Bernanke, Ben S., and Alan S. Blinder, 1991, The Federal funds rate and the channels of monetary transmission, *American Economic Review* 82, 901-921.
- Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373-399.
- Campbell, John Y., Consumption-based asset pricing, in George Constantinides, Milton Harris, and Rene Stulz., Eds., *Handbook of the Economics of Finance Vol. 1B*, North-Holland, Amsterdam, 803-887.
- Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.
- Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495-514.
- Cheridito, Patrick, Damir Filipović, and Robert L. Kimmel, 2007, Market price of risk specifications for affine models: Theory and evidence, *Journal of Financial Economics* 83, 123-170.
- Christensen, Jens H.E., Jose Lopez, and Glenn D. Rudebusch, 2010, Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields, *Journal of Money, Credit, and Banking* 42, 143-178.

- Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138-160.
- Cooper, Ilan, and Richard Priestly, 2009, Time-varying risk premiums and the output gap, *Review of Financial Studies* 22, 2801-2833.
- Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-407.
- Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943-1978.
- Dai, Qiang, and Kenneth J. Singleton, 2003, Fixed-income pricing, in G.M. Constantinides, M. Harris and R. Stulz, Eds., *Handbook of the Economics of Finance* Volume 1B, 1208-1246.
- Dai, Qiang, Kenneth J. Singleton, and Wei Yang, Regime shifts in a dynamic term structure model of U.S. Treasury bond yields, *Review of Financial Studies* 20, 1669-1706.
- D'Amico, Stefania, Don H. Kim, and Min Wei, 2008, Tips from TIPS: The informational content of Treasury Inflation-Protected Security prices, Federal Reserve Board Discussion Paper 2008-30.
- den Haan, Wouter J., 1995, The term structure of interest rates in real and monetary economies, *Journal of Economic Dynamics and Control* 19, 909-940.
- Duarte, Jefferson, 2004, Evaluating an alternative risk preference in affine term structure models, *Review of Financial Studies* 17, 379-404.
- Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405-443.
- Duffee, Gregory R., 2010, Sharpe ratios in term structure models, Working paper, Johns Hopkins.

- Duffee, Gregory R., 2011, Information in (and not in) the term structure, *Review of Financial Studies* 24, 2895-2934.
- Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.
- Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937-969.
- Estrella, Arturo, and Gikas Hardouvelis, 1991, The term structure as a predictor of real economic activity, *Journal of Finance* 46, 555-576.
- Gallmeyer, Michael, Burton Hollifield, Francisco Palomino, and Stanley Zin, 2008, Term premium dynamics and the Taylor rule, Working paper, University of Michigan.
- Gray, Stephen F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27-62.
- Hodrick, Robert J., 1992, Dividend yields and expected stock returns: alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357-386.
- Hördahl, Peter, Oreste Tristani, and David Vestin, 2006, A joint econometric model of macroeconomic and term structure dynamics, *Journal of Econometrics* 131, 405-444.
- Hördahl, Peter, Oreste Tristani, and David Vestin, 2008, The yield curve and macroeconomic dynamics, *Economic Journal* 118, 1937-1970.
- Joslin, Scott, Ahn Le, and Kenneth J. Singleton, 2011, Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs, Working paper, Stanford GSB.
- Joslin, Scott, Marcel Pribsch, and Kenneth J. Singleton, 2010, Risk premiums in dynamic term structure models with unspanned macro risks, Working paper, Stanford GSB.

- Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, 2011, A new perspective on Gaussian dynamic term structure models, *Review of Financial Studies* 24, 926-970.
- Kessel, Reuben A., 1965, The cyclical behavior of the term structure of interest rates, NBER occasional paper 91.
- Le, Ahn, Kenneth J. Singleton, and Qiang Dai, 2010, Discrete-time affine^Q term structure models with generalized market prices of risk, *Review of Financial Studies* 23, 2184-2227.
- Leippold, Markus, and Liuren Wu, 2002, Asset pricing under the quadratic class, *Journal of Financial and Quantitative Analysis* 37, 271-295.
- Lettau, Martin, and Jessica Wachter, 2010, The term structures of equity and interest rates, Working paper, Wharton.
- Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54-61.
- Ludvigson, Sydney C., and Serene Ng, 2009, Macro factors in bond risk premia, *Review of Financial Studies* 22, 5027-5067.
- Ludvigson, Sydney C., and Serene Ng, 2010, A factor analysis of bond risk premia, *Handbook of Applied Econometrics*, forthcoming.
- Piazzesi, Monika, and Martin Schneider, 2007, Equilibrium yield curves, *NBER Macroeconomics Annual 2006*, Cambridge, MA: MIT Press, 389-442.
- Rudebusch, G.D., Swanson, E. (2012). "The bond premium in a DSGE model with long-run real and nominal risks". *American Economic Journal: Macroeconomics* 4, 105-143.
- 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.

- Taylor, John B., 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177-188.
- van Binsbergen, Jules H., Jesús Fernández-Villaverde, Ralph S.J. Koijen, and Juan F. Rubio-Ramírez, 2010, The term structure of interest rates in a DSGE model with recursive preferences, Working paper, Stanford GSB.
- Van Horne, James, 1965, Interest-rate risk and the term structure of interest rates, *Journal of Political Economy* 73, 344-351.
- Wachter, Jessica A., 2006, A consumption-baese model of the term structure of interest rates, *Journal of Financial Economics* 79, 365-399.
- Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401-421.

Table 1. Correlations among quarterly measures of economic activity

Five measures capture the output gap. They are [1] detrended log real GDP, [2] detrended log industrial production, [3] Hodrick-Prescott filtered log real GDP, [4] Hodrick-Prescott filtered log industrial production, and [5] demeaned capacity utilization. Trends are removed with linear regressions on a constant and a time index. Five measures capture current economic growth. They are [1] log-differenced per capita real consumption on nondurables and services, [2] log-differenced real GDP, [3] log-differenced industrial production, [4] the Federal Reserve Bank of Chicago's National Activity Index (CFNAI), and [5] the Ludvigson and Ng (2010) macroeconomic activity principal component. The first observations of capacity utilization, CFNAI, and Ludvigson-Ng are 1967Q1, 1976Q2, and 1964Q1 respectively. All other series begin in 1952Q2. The last observation is 2010Q4 for all except the Ludvigson-Ng series, which ends in 2007Q4. The reported correlation between two series uses all non-missing data available for the two series.

	Output gap measures					Growth measures				
	[1]	[2]	[3]	[4]	[5]	[1]	[2]	[3]	[4]	[5]
Cycle										
[1]	1.00									
[2]	0.95	1.00								
[3]	0.47	0.42	1.00							
[4]	0.41	0.45	0.90	1.00						
[5]	0.74	0.75	0.65	0.70	1.00					
Growth										
[1]	0.08	0.06	0.03	0.06	0.22	1.00				
[2]	0.18	0.18	0.27	0.24	0.26	0.52	1.00			
[3]	0.07	0.13	0.13	0.28	0.32	0.50	0.69	1.00		
[4]	0.39	0.40	0.30	0.37	0.50	0.57	0.74	0.90	1.00	
[5]	0.22	0.22	0.25	0.27	0.34	0.46	0.58	0.82	0.91	1.00

Table 2. Projections of economic activity and inflation on nominal yields

Five measures capture the output gap. Five others capture current economic growth. They are all described in the notes to Table 1. Three quarterly inflation measures are the quarter-to-quarter log change in the GDP deflator, the log change in the CPI (final month in previous quarter to final month in current quarter), and the log change in the CPI excluding food and energy. This table reports R^2 s and first-order serial correlation of residuals for cross-sectional regressions on six Treasury bond yields. The bond maturities range from three months to five years. There are missing observations for parts of the sample periods for capacity utilization, CFNAI, and the Ludvigson-Ng principal component. The available data ranges for the economic activity data are described in the notes to Table 1. Inflation measured with the CPI excluding food and energy begins with 1957Q2.

Measure	1952Q2–2010Q4		1986Q1–2007Q4	
	R^2	AR(1) of residuals	R^2	AR(1) of residuals
Output gap				
log real GDP, detrended	0.38	0.94	0.41	0.82
log ind prod, detrended	0.44	0.91	0.33	0.89
log real GDP, HP-filtered	0.33	0.73	0.54	0.70
log ind prod, HP-filtered	0.34	0.70	0.62	0.70
demeaned capacity use	0.42	0.82	0.64	0.78
Growth				
log-diff per capita C	0.09	0.30	0.07	0.01
log-diff real GDP	0.02	0.34	0.19	0.15
log-diff ind prod	0.03	0.35	0.26	0.11
CFNAI	0.07	0.71	0.39	0.49
Ludvigson-Ng PC	0.13	0.60	0.41	0.32
Inflation				
GDP deflator	0.47	0.71	0.29	0.47
CPI	0.49	0.25	0.19	-0.23
CPI ex food & energy	0.57	0.53	0.69	0.22

Table 3. Projections of expected future economic growth on the nominal term structure

Measures of economic growth in quarter $t+2$ are regressed on economic growth in quarters t and $t-1$, CPI inflation in quarters t and $t-1$, and six quarter- t Treasury bond yields. The fitted values of these forecasting regressions are regressed on the six Treasury quarter- t bond yields. The R^2 of the second regression is the fraction of the regression-based expectations that is spanned by the nominal term structure. The table reports the R^2 s of both regressions. It also reports the serial correlation of the residuals from the second-stage regression. The five measures of economic growth and their available data ranges are described in the notes to Table 1.

	Forecasting regression R^2	Fraction of forecast spanned by nominal yields	AR(1) of unspanned component
1952Q2–2010Q4			
log-diff per capita C	0.24	0.68	0.44
log-diff real GDP	0.15	0.71	0.15
log-diff ind prod	0.15	0.80	0.40
CFNAI	0.39	0.39	0.73
Ludvigson-Ng PC	0.41	0.72	0.61
1986Q1–2007Q4			
log-diff per capita C	0.20	0.30	0.19
log-diff real GDP	0.17	0.43	0.04
log-diff ind prod	0.19	0.53	0.14
CFNAI	0.44	0.67	0.54
Ludvigson-Ng PC	0.36	0.86	0.42

Table 4. Projections of nominal yields on measures of economic activity and inflation

Quarter-end yields on Treasury bonds with maturities ranging from three months to five years are regressed on two contemporaneous measures of the output gap, lags zero through two of ARMA-smoothed CPI inflation, and lags zero through two of the log-change in industrial production. This table reports R^2 s, standard deviations of the fitted residuals, and first-order serial correlation of residuals.

Maturity	1952Q2–2010Q4			1986Q1–2007Q4		
	R^2	Std dev of residuals	AR(1) of residuals	R^2	Std dev of residuals	AR(1) of residuals
Three months	0.66	1.77	0.83	0.59	1.30	0.88
One year	0.64	1.83	0.86	0.60	1.32	0.84
Two years	0.62	1.86	0.88	0.57	1.32	0.80
Three years	0.59	1.86	0.89	0.54	1.30	0.80
Four years	0.58	1.85	0.90	0.54	1.26	0.79
Five years	0.58	1.83	0.90	0.54	1.23	0.80

Table 5. Root mean squared cross-sectional and forecasting errors for three models

Four-factor Gaussian linear models are used to explain the joint dynamics of quarterly inflation, aggregate per capita consumption growth, and nominal Treasury bond yields. The first model uses only inflation, consumption growth, and the three-month and five-year yields. The second model adds yields on bonds with maturities of one, two, three, and four years. The third model uses the same data as the second model. It also imposes no-arbitrage. For each of these models, all observations are assumed to contain white-noise shocks that account for deviations from an exact four-factor model. The final model replaces inflation and consumption growth with ARMA-smoothed versions, and assumes that these smoothed versions have no white-noise shocks. The data sample is 1952Q2 through 2010Q4. Estimation is maximum likelihood using the Kalman filter. Panel A reports root mean squared errors of each models model's cross-sectional fit. Panel B reports root mean squared errors of one-quarter-ahead forecasts. The units are annualized percentage points.

Variable	Factor model, two bonds	Factor model, six bonds	No-arbitrage, six bonds	No-arbitrage, ARMA smoothed
Panel A. Cross-sectional deviations				
3 month yield	0.31	0.26	0.25	0.42
1 year yield		0.08	0.08	0.12
2 year yield		0.05	0.06	0.05
3 year yield		0.04	0.04	0.06
4 year yield		0.07	0.07	0.06
5 year yield	0.01	0.05	0.05	0.07
Inflation	1.69	1.60	1.63	1.61
Consumption growth	1.10	1.67	1.64	1.28
Panel B. One-quarter-ahead forecasting errors				
3 month yield	0.93	0.99	0.99	0.98
1 year yield		0.91	0.91	0.88
2 year yield		0.82	0.82	0.80
3 year yield		0.75	0.75	0.73
4 year yield		0.70	0.70	0.70
5 year yield	0.65	0.65	0.66	0.66
Inflation	2.23	2.28	2.28	2.30
Consumption growth	1.73	1.73	1.73	1.76

Table 6. Properties of the stochastic discount factor implied by an estimated model

Four-factor Gaussian no-arbitrage models are used to explain the joint dynamics of quarterly inflation, aggregate consumption, and nominal Treasury bond yields. One model assumes that all observations have components related to the factors and a white-noise shock that accounts for deviations from the factor model. The second model replaces inflation and consumption growth with ARMA-smoothed measures and assumes that these measures have no white-noise components. The data sample is 1952Q2 through 2010Q4. Estimation is maximum likelihood using the Kalman filter. The table reports unconditional means of various features of the implied stochastic discount factor. Standard errors are in parentheses.

Feature	All with white-noise components	No macro white-noise components
Nominal three-month yield (%/year)	4.09 (1.92)	4.21 (2.01)
Nominal five-year yield less three-month yield (%/year)	1.03 (0.18)	1.03 (0.29)
Compensation for a log-price loading of minus one on inflation (%/quarter)	-0.52 (0.33)	-23.03 (1000.12)
Compensation for a log-price loading of one on consumption growth (%/quarter)	-0.15 (0.12)	-31.81 (1367.06)
Real three-month yield (%/year)	3.05 (1.64)	93.14 (4000.53)
Real five-year yield less three-month yield (%/year)	5.17 (5.06)	337.57 (11874.81)

Table 7. Sample correlations, 1952 through 2010

Excess returns to a portfolio of long-term Treasury bonds and the aggregate stock market are measured from the end of quarter $t - 1$ to the end of quarter t . Contemporaneous consumption growth is the log change in per capita aggregate consumption from quarter $t - 1$ to quarter t , Δc_t . Industrial production growth and inflation are both log changes from the final month of quarter $t - 1$ to the final month of quarter t , ΔIP_t and π_t respectively. The sample period is 1952Q2 through 2010Q4.

	Excess bond ret	Excess stock ret	Δc_t	Δc_{t+1}	ΔIP_t	ΔIP_{t+1}	π_t
Excess bond return	1.00						
Excess stock return	0.06	1.00					
Δc_t	-0.13	0.20	1.00				
Δc_{t+1}	0.07	0.26	0.38	1.00			
ΔIP_t	-0.28	0.04	0.50	0.26	1.00		
ΔIP_{t+1}	-0.13	0.36	0.42	0.52	0.38	1.00	
π_t	-0.29	-0.15	-0.25	-0.30	-0.13	-0.19	1.00

Table 8. Regressions of excess bond returns on consumption growth and stock returns

Excess returns to a portfolio of long-term Treasury bonds and the aggregate stock market are measured from the end of quarter $t-1$ to the end of quarter t . Contemporaneous consumption growth is the log change in aggregate consumption from quarter $t-1$ to quarter t , Δc_t . The table reports coefficients of regressions of excess bond returns on either consumption growth or stock returns. Asymptotic t -statistics, in parentheses, use a Newey-West adjustment with two lags. The units are percent per quarter.

Sample	Constant	Consumption growth	aggregate stock return
1952Q2–2010Q4	0.40 (1.91)		
	0.77 (2.41)	-0.86 (-1.81)	
	0.36 (1.64)		0.02 (0.63)
1952Q2–1980Q4	-0.15 (-0.67)		
	0.29 (0.66)	-0.88 (-1.40)	
	-0.27 (-1.14)		0.07 (1.93)
1981Q1–2010Q4	0.92 (2.90)		
	1.11 (2.52)	-0.51 (-0.65)	
	0.94 (2.88)		-0.01 (-0.27)
1986Q1–2007Q4	0.74 (2.64)		
	1.47 (3.06)	-1.76 (-1.97)	
	0.85 (2.99)		-0.06 (-1.38)

Table 9. Properties of an estimated factor model

A two-factor Gaussian linear model is used to explain the joint dynamics of quarterly inflation π_t and per-capita aggregate consumption growth Δc_t . The data sample is 1952Q2 through 1985Q4. Estimation is maximum likelihood using the Kalman filter. Given the estimated parameters, model-implied covariances conditioned on quarter- t information can be calculated between the variables listed in Panel A's rows and columns. The notation ζ_{t+1} refers to news about the infinite sum of consumption growth beginning with Δc_{t+1} , discounted at a time rate of preference $\beta = 0.999$. In Panel B, the mean return to a nominal bond in excess of the nominal short rate is calculated using a recursive utility model with a unit elasticity of substitution, a coefficient of relative risk aversion of ten, and a time rate of preference of β . All variables are measured in percent per quarter. Standard errors are in parentheses.

A. Covariances among innovations

	Δc_{t+1}	π_{t+1}	ζ_{t+1}
$E_{t+1}(\Delta c_{t+2})$	0.036 (0.014)	-0.020 (0.006)	-0.025 (0.046)
$E_{t+1}(\pi_{t+2})$	0.012 (0.017)	0.023 (0.012)	-0.197 (0.112)
$E_{t+1}(\frac{1}{20} \sum_{h=1}^{20} \Delta c_{t+1+h})$	-0.007 (0.002)	-0.003 (0.004)	0.049 (0.027)
$E_{t+1}(\frac{1}{20} \sum_{h=1}^{20} \pi_{t+1+h})$	0.033 (0.008)	0.005 (0.013)	-0.172 (0.097)

B. Nominal bond returns implied by recursive utility

Maturity	Mean excess return per quarter (%)
5 years	0.210 (0.125)

Table 10. Nominal bond risk premia in a recursive utility model

The joint dynamics of per capita consumption growth and inflation are estimated with various factor models. The estimated models differ in the number of factors, whether nominal bond yields are also included in estimation, and the sample period used to estimate the model. The estimated joint dynamics of consumption growth and inflation are then plugged into a recursive utility model with a unit elasticity of substitution, a coefficient of relative risk aversion of ten, and a time rate of preference of β . The table reports the model-implied unconditional mean quarterly return to a five-year nominal bond in excess of the nominal short rate. Standard errors are in parentheses.

Factors	Nominal yields among observables	Sample	Mean excess return per quarter (%)
3	3 month	1952Q2 – 1985Q4	0.075 (0.109)
4	3 month, 5 year	1952Q2 – 1985Q4	0.284 (0.684)
2	-	1952Q2 – 2010Q4	0.068 (0.075)
3	3 month	1952Q2 – 2010Q4	-0.012 (0.095)
4	3 month, 5 year	1952Q2 – 2010Q4	0.006 (0.110)
4	3 month, 5 year	1986Q1 – 2007Q4	-0.014 (0.030)

Table 11. Sample covariances for evaluation of risk premia models

The joint dynamics of per capita consumption growth and inflation are estimated with two different factor models. The sample period used to estimate the models is 1952Q2 through 1985Q4. The estimated models are then used to construct Kalman-smoothed estimates of quarterly shocks to consumption growth, inflation, and a discounted infinite sum of expected future consumption growth over the longer sample 1952Q2 through 2010Q4. The table reports sample covariances between quarterly changes in nominal bond yields and the fitted shocks. The units of yields and shocks are percent per quarter.

A. Two factors, no additional observables

Sample	Consumption	Inflation	Discounted sum
1952Q2 to 1985Q4			
Three month	0.011	0.003	-0.066
Five year	0.002	0.003	-0.029
1986Q1 to 2010Q4			
Three month	0.004	0.000	-0.016
Five year	-0.001	0.003	-0.015
1952Q2 to 2010Q4			
Three month	0.009	0.001	-0.044
Five year	0.002	0.002	-0.022

B. Three factors, three-month yield is also observed

Sample	Consumption	Inflation	Discounted sum
1952Q2 to 1985Q4			
Three month	0.033	0.029	0.010
Five year	0.013	0.014	0.008
1986Q1 to 2010Q4			
Three month	0.015	0.004	0.027
Five year	0.011	0.005	0.005
1952Q2 to 2010Q4			
Three month	0.027	0.018	0.019
Five year	0.013	0.010	0.009

Table 12. Predictability of excess returns

Monthly excess returns to long-term nominal Treasury bonds and the aggregate stock market are regressed on a variety of predetermined variables. The column labeled “Test Stat” is the p -value of the test that the coefficient(s) are all zero. The covariance matrix of the estimated coefficients uses the Newey-West adjustment for five lags of moving average residuals. For regressions with a single predictor, the sign of the estimated coefficient is reported.

Predictors	Sample	Adj. R^2	Test stat	Sign
A. Excess bond returns				
Slope	1952:7 – 2010:12	0.0256	0.002	Pos
12-month averages of 5 forward rates	1952:7 – 2010:12	0.0203	0.005	–
Conditional SD of yields	1962:3 – 2010:12	0.0007	0.486	Pos
Log change of industrial prod.	1952:7 – 2010:12	0.0016	0.111	Neg
12-month average of log change of industrial prod.	1952:7 – 2010:12	0.0007	0.225	Neg
8 principal components of of macro/financial variables	1964:2 – 2008:1	0.0864	0.000	—
12-month averages of 5 forward rates	1964:2 – 2008:1	0.0463	0.000	–
“Real activity” principal component of macro/financial variables	1964:2 – 2008:1	0.0161	0.002	Neg
Log change of industrial prod.	1964:2 – 2008:1	0.0066	0.017	Neg
B. Excess stock returns				
Slope	1952:7 – 2010:12	0.0047	0.066	Pos
12-month averages of 5 forward rates	1952:7 – 2010:12	0.0031	0.378	–
Log change of industrial prod.	1952:7 – 2010:12	–0.0012	0.774	Pos
8 principal components of of macro/financial variables	1964:2 – 2008:1	0.0408	0.000	—
“Real activity” principal component of macro/financial variables	1964:2 – 2008:1	0.0018	0.245	Neg

Table 13. Predicting real rates and excess bond returns with surplus consumption

A discounted sum of 40 quarters of log real consumption growth is a proxy for surplus consumption. The proxy is constructed in two ways. One follows Wachter (2006) by using aggregate consumption on nondurables and services. The other uses per capita consumption on these components. Each quarter- t proxy is used to predict both the real return to a one-quarter bond in quarter $t + 1$ (the ex-post real rate) and the excess return in quarter $t + 1$ to a portfolio of Treasury bonds. The regressions are estimated using the short sample 1962Q2 through 2004Q4 and the full sample 1962Q2 through 2010Q4. The table reports parameter estimates and asymptotic t -statistics. The covariance matrices of the parameter estimates are adjusted for generalized heteroskedasticity. For the real-rate regressions, they are also adjusted for six lags of moving-average residuals.

Dependent variable	Consumption	1962Q2 – 2004Q4	1962Q2 – 2010Q4
Ex-post real rate	Aggregate	–0.084 (–2.64)	–0.022 (–0.72)
Excess bond return	Aggregate	–0.152 (–1.75)	–0.126 (–1.89)
Ex-post real rate	Per capita	–0.042 (–0.97)	–0.003 (–0.08)
Excess bond return	Per capita	–0.049 (–0.41)	–0.072 (–0.83)

Table 14. Root mean squared cross-sectional and forecasting errors for various models

Three-factor and four-factor Gaussian linear models are used to explain the joint dynamics of quarterly inflation, detrended log industrial production, and the yield on a three-month Treasury bill. Some of the estimated models impose no restrictions, while others impose restrictions based on New Keynesian models. In addition, some of the models include extra long-term Treasury bond yields. For each model, all observations are assumed to contain white-noise shocks that account for deviations from an exact factor model. The data sample is 1960Q1 through 2010Q4. Estimation is maximum likelihood using the Kalman filter. The table reports root mean squared errors of each model's cross-sectional fit. It also reports root mean squared errors of one-quarter-ahead forecasts. The units are annualized percentage points.

A. Unrestricted models

	Three factors, no long bonds	Four factors, no long bonds	Four factors, one long bond	Four factors, five long bonds
Cross-sectional fitting errors				
3 month yield	0.304	0.297	0.232	0.310
Inflation	1.778	1.750	1.792	2.329
log IP	0.000	0.000	0.000	0.000
One-quarter-ahead forecasting errors				
3 month yield	0.976	0.972	0.986	1.015
Inflation	2.276	2.243	2.261	2.445
log IP	6.584	5.914	6.582	6.548

B. Restricted four-factor models

	No long bonds	One long bond	Five long bonds
Cross-sectional fitting errors			
3 month yield	0.303	0.259	0.302
Inflation	1.660	2.423	2.487
log IP	0.000	0.000	0.000
One-quarter-ahead forecasting errors			
3 month yield	0.979	0.992	1.019
Inflation	2.289	2.550	2.567
log IP	6.057	6.006	6.885

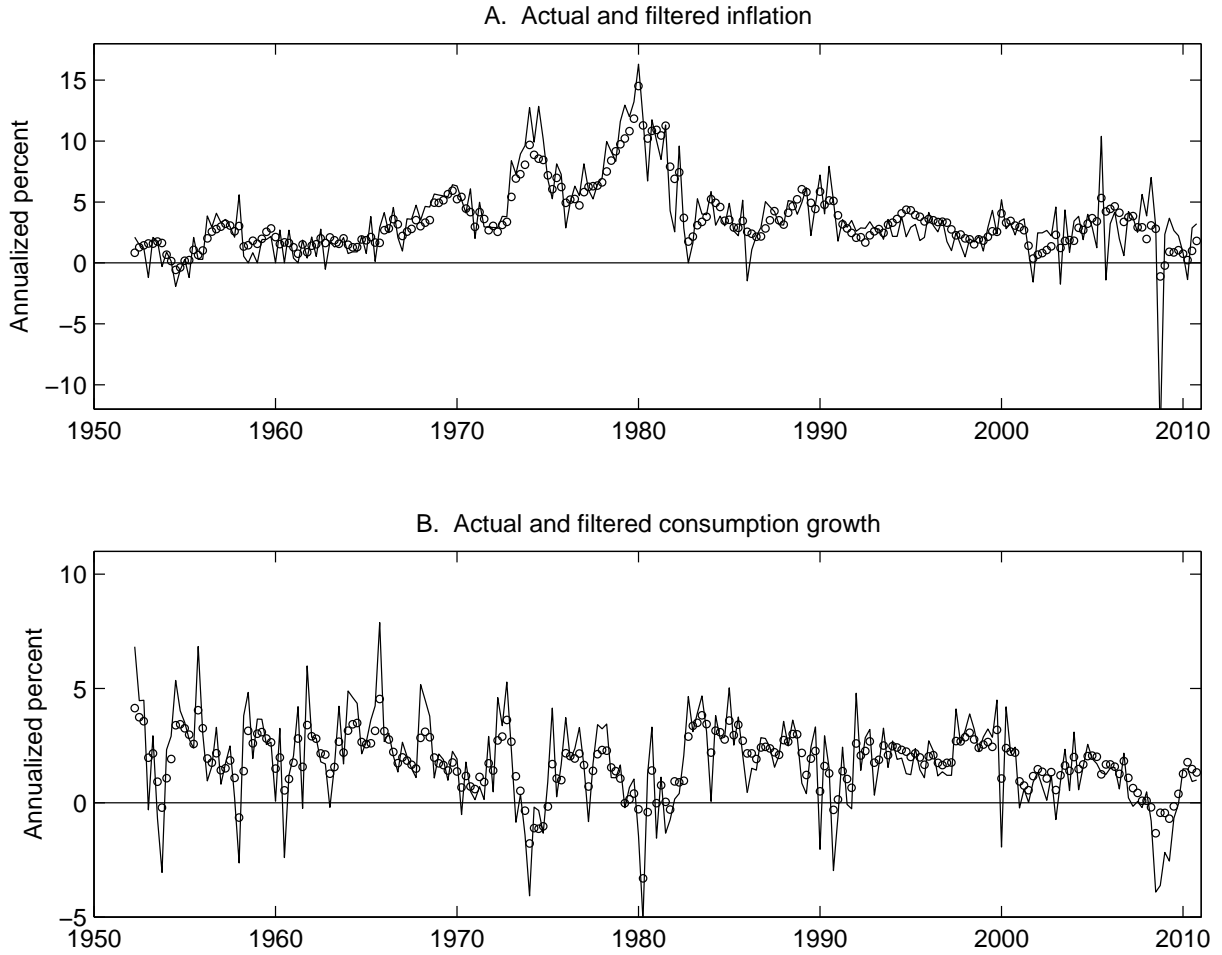


Figure 1. Filtered estimates of inflation and consumption growth from a factor model. A four-factor Gaussian linear model is fit to the joint dynamics of quarterly inflation, per capita consumption growth, and yields on three-month and five-year Treasury bonds. All observations are assumed to contain white-noise shocks that account for deviations from an exact four-factor model. The panels display actual (black lines) and Kalman filtered (circles) values of inflation and consumption growth. The data sample is 1952Q2 through 2010Q4.

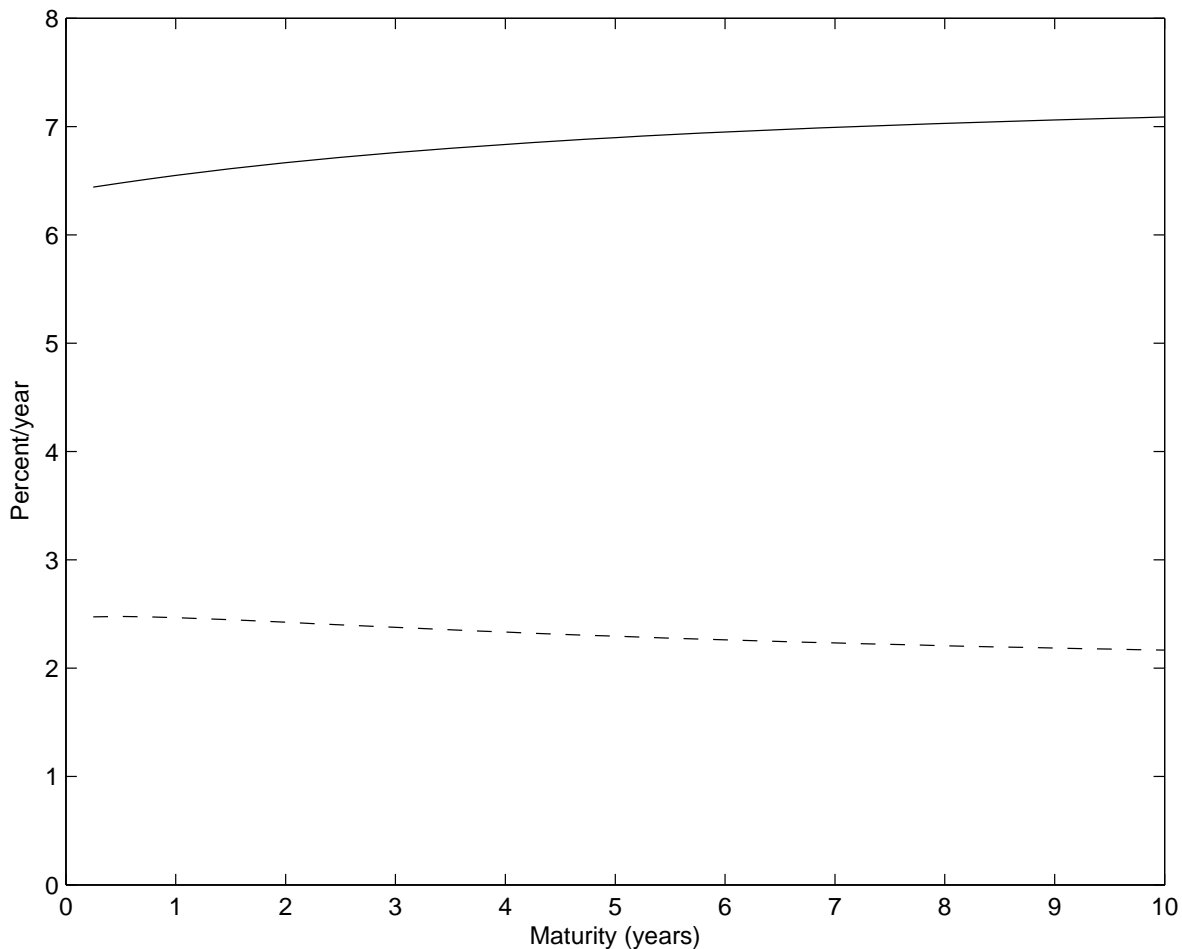


Figure 2. Mean real and nominal yield curves implied by recursive utility. The lower (upper) line is the unconditional mean real (nominal) yield curve implied by a recursive utility model and estimated joint dynamics of consumption growth and inflation. The time rate of preference is 0.999 per quarter, the coefficient of relative risk aversion is ten, and the elasticity of intertemporal substitution is one. The data sample is 1952Q2 through 1985Q4.

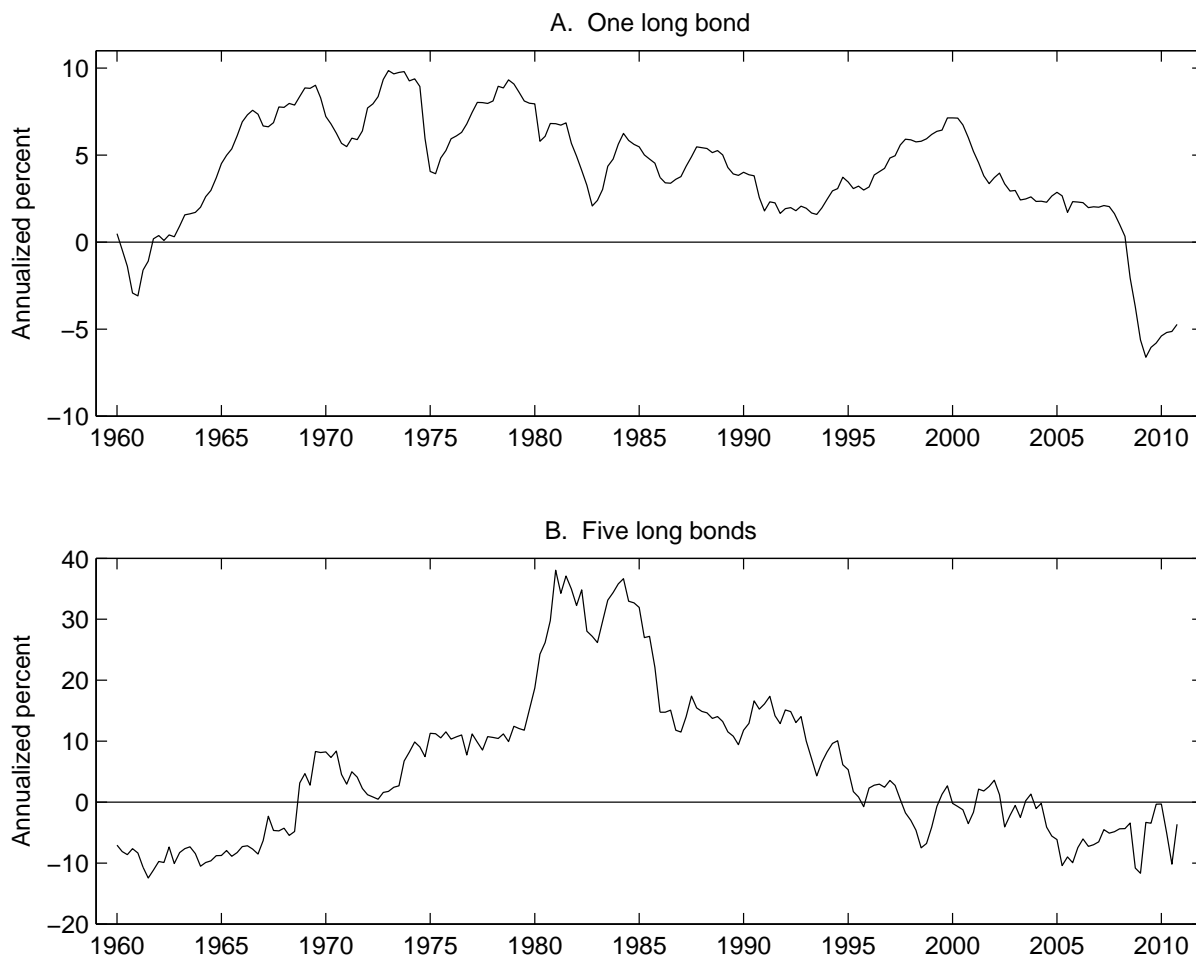


Figure 3. Filtered estimates of a “long-run inflation target” factor from a New Keynesian model. A four-factor Gaussian model, with restrictions implied by a New Keynesian macro model, is used to estimate the monetary authority’s long-run inflation target. The model underlying Panel A is estimated using only the three-month Treasury bill yield, quarter-to-quarter inflation, detrended log industrial production, and the yield on a five-year Treasury bond. The model underlying Panel B adds yields on Treasury bonds with maturities from one to five years. The data sample is 1960Q1 through 2010Q4.